TRANSCRIPTIONAL AND TRANSLATIONAL CONTROL OF GENE EXPRESSION DURING POLLEN DEVELOPMENT

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Neil Bate BSc. (Leicester)
Department of Botany
University of Leicester

October 1997

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

For Reginald William Bate

Transcriptional and translational control of gene expression during pollen development

Ph.D. Thesis by Neil Bate

The tomato *lat*52 gene encodes an essential cysteine-rich protein required for normal pollen tube growth which is preferentially transcribed in the vegetative cell during pollen maturation. The identity, organisation and role of *cis*-regulatory elements which control the developmental and cell-specific expression of *lat*52 were investigated in detail. A series of 5' promoter deletion mutants fused to the *E. coli* ß-glucuronidase gene *uidA* (*gus*) were stably introduced into tobacco. Detailed analysis of the accumulation of GUS activity during anther development demonstrated that *lat*52 promoter activity during early to mid-pollen maturation was regulated specifically by the region -492 to -101.

Transgenic studies and transient expression analysis demonstrated that the major *cis*-regulatory elements responsible for high-level pollen-specific transcriptional activity were located between -492 and -42. Fusion of the region -492 to -52 to a minimal *CaMV35S* promoter enhanced reporter gene expression in a pollen-specific manner. This gain of function approach was used to show that the previously defined sequence motifs PBII and PBIII did not contribute to pollen-specific transcription in germinating pollen. Analysis of the region -100 to -52 led to the identification of two novel co-dependent *cis*-regulatory elements within the minimal pollen-specific activator unit -71 to -52.

In addition to multiple upstream *cis*-regulatory elements which control the level and specificity of *lat*52 transcription, the 5'-untranslated region (5'-UTR) conferred a dramatically increased translational yield to heterologous transcripts in a pollen-specific and highly developmentally regulated manner during the final stages of pollen maturation in transgenic tobacco. Taken together, these data suggest that *lat*52 expression is developmentally regulated in a tri-phasic manner during tobacco pollen development.

Acknowledgements

I would like to thank my supervisor Dr. David Twell for excellent supervision and constant approachability throughout the course of this project. I would also like to thank Dr. Gary Foster for helpful discussion on all things translational. Thanks to Dr. Colin Eady, Dr. Joy Wilkinson, Caroline Spurr and Man-Kim Cheung for allowing me to present their data in this thesis. Also I would like to thank Caroline Spurr, Gemma Allnutt, Fiona Cooke and Victoria Davies for technical assistance. Gratitude extends to Dr. Hilary Rogers (Cardiff), Dr. Roisin Turner and Dr. Koen Weterings (Nijmegen) for their gifts of plasmids.

I would also like to thank members of the Botany Department and Biocentre (past and present) for making my time in Leicester such an enjoyable one. Special thanks to Natasha Sheldon for constant support and extreme patience while I was writing up, and finally to my parents for their continued support in everything that I do.

I acknowledge with thanks, the financial support of the B.B.S.R.C.

Contents **Page**

List of figures

Figure title **Page**

List of tables

Table Title **Page**

Abbreviations used

A	adenine
ATP	adenosine triphosphate
BAP	6-benzylaminopurine
bp	base pair
BSA	bovine serum albumin
C	cytosine
°C	degrees centigrade
C3'	*CaMV35S* 3'-untranslated region
cDNA	complementary DNA
cm	centimetre
cv	cultivar
2,4 D	2,4-dichlorophenoxyacetic acid
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
EDTA	diaminoethanetetra-acetic acid
g	gram
g	gravity
G	guanine
GEB	GUS extraction buffer
gus	ß-glucuronidase gene
GUS	ß-glucuronidase protein
Hepes	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
Kb	Kilobase pair
Kda	Kilodalton
l	litre
LEB	luciferase extraction buffer
luc	luciferase gene
LUC	luciferase protein
M	molar
MES	2-(N-Morpholino)ethane-sulphonic acid
µg	microgram
µl	microlitre
mg	milligram
ml	millilitre
min	minute
mM	millimolar
mRNA	messenger RNA

MS	Murashige and Skoog
4-MU	4-methylumbelliferone
MUG	4-methylumbelliferyl-glucuronide
NAA	α-naphthaleneacetic acid
ng	nanogram
nos	nopaline synthase
O.D.	optical density
p	plasmid
pg	picogram
PCR	polymerase chain reaction
PEG	polyethylene glycol
room temp	room temperature
RNA	ribonucleic acid
RT-PCR	reverse transcriptase-PCR
sec	second
SDS	sodium dodecyl sulphate
SH_2O	sterile distilled water
T	thymine
TAE	tris-acetate EDTA
TBE	tris-borate EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	tris(hydroxymethyl)aminomethane
UTR	untranslated region
v/v	volume per volume
WT	wild type
w/v	weight per volume
X-Gluc	5-bromo-4-chloro-3-indoyl-ß-glucuronide

Chapter 1

Introduction

1.1 The pathway of pollen development: from microsporocyte to pollen germination

In flowering plants the male gametophyte or pollen grain is a highly reduced structure which is specialized for the production and delivery of two sperm cells to the egg and central cells within the embryo sac, where fertilization takes place (Heslop-Harrison, 1987; Bedinger, 1992; Smyth, 1997). Pollen development in angiosperms has been described previously (Mascarenhas, 1989; Scott *et al.*, 1991b; Bedinger, 1992; McCormick, 1993; Twell, 1994; Owen and Makaroff, 1995). Diploid microsporocytes or pollen mother cells are contained within a fluid-filled locule enclosed by four anther wall layers (tapetum, middle layer, endothecium and the epidermis). The microsporocytes secrete a surrounding wall of ß-1,3-glucans (callose) and then undergo meiosis. Two meiotic divisions transform the diploid microsporocyte into a tetrad of haploid microspore nuclei contained within a common cytoplasm (fig. 1.1). Following meiosis II, the microspores are separated from each other by simultaneous cytokinesis and the formation of intersporal walls (callose) (Horner, 1977). Recently, the allele TETRASPORE has been identified as essential for male meiotic cytokinesis in *Arabidopsis thaliana* (Spielman *et al.*, 1997). Mutation of TETRASPORE delays intersporal wall formation which results in all four microspore nuclei remaining within the same common cytoplasm.

When the microsporocytes are undergoing meiosis, the tapetal cells differentiate into binucleate polar secretory cells, which do not contain a primary cell wall. These cells contain ribosomes, mitochondria, endoplasmic reticulum and Golgi in abundance. In addition, many secretory vesicles are located on the locular face of the plasma membrane which are connected by cytoplasmic bridges early in development.

Following meiosis and the formation of the intersporal walls, pollen wall formation begins almost immediately. The microspores contained within the tetrad synthesise a cellulosic primexine outside the plasmalemma, which provides a matrix for the deposition of sporopollenin. Initially, sporopollenin is synthesised from precursors present within the microspore cytoplasm and secreted outside the cell between the plasma membrane and the callose wall. Once the young microspore possesses a partially formed exine the callose wall of the tetrad is rapidly broken down releasing the free single microspores into the locule (fig. 1.1). Microspore release is brought about by a ß-1,3-glucanase (callase), which is synthesised by the tapetal cells and secreted into the locule. The timing of the release of this enzyme is critical for continued microspore development. Worrall *et al.*, (1992) expressed a modified basic PR ß-1,3-glucanase under the control of tapetum-specific promoters (Scott *et al.*, 1991a; Paul *et al.*, 1992) that are transcriptionally active during meiosis in transgenic tobacco. The expression of ß-1,3-glucanase at an earlier developmental stage than normal resulted in partial or complete male sterility.

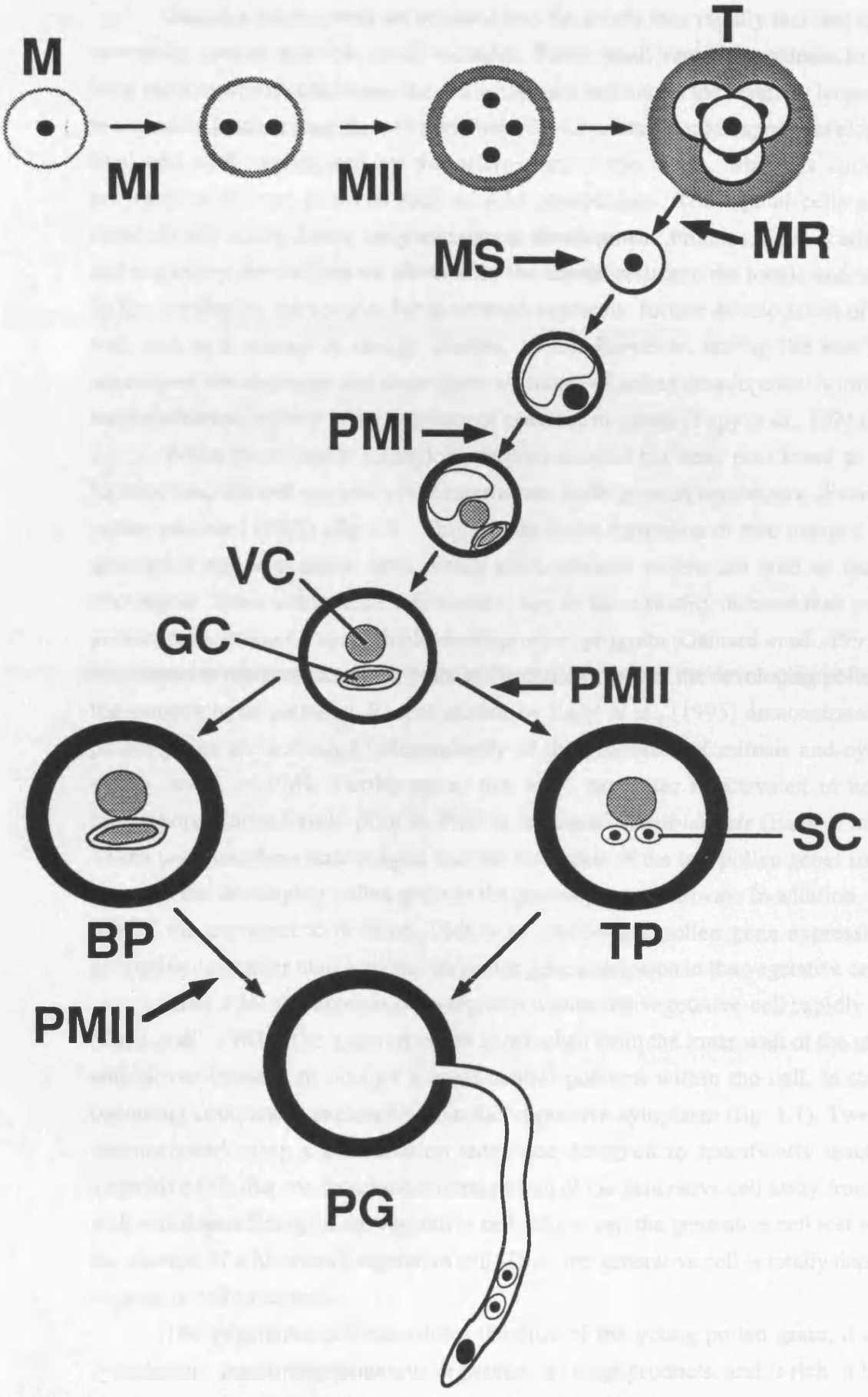


Figure 1.1. Schematic representation of the major stages during pollen development. Abbreviations: M, microsporocyte; MI, Meiosis I; MII, Meiosis II; T, Tetrad; MR, Microspore Release; MS, Microspore; PMI, Pollen Mitosis I; VC, Vegetative Cell; GC, Generative Cell; PMII, Pollen Mitosis II; SC, Sperm Cell; BP, Bicellular Pollen; TP, Tricellular Pollen; PG, Pollen Germination.

Once the microspores are released into the locule they rapidly increase in size and eventually contain multiple small vacuoles. These small vacuoles coalesce to form one large vacuole which compresses the cell cytoplasm and moves the centrally located nucleus to a specific location near the cell periphery (fig 1.1). During microspore development the final cell wall synthesised by the microspore is the intine, which is composed of polysaccharides and proteins such as acid phosphatase. The tapetal cells are highly metabolically active during early microspore development. Proteins, lipids, carbohydrates and secondary metabolites are secreted by the tapetal cells into the locule and are utilised by the developing microspore for membrane synthesis, further development of the exine wall and as a source of energy (Pacini, 1990). However, during the late stages of microspore development and throughout all stages of pollen development normal growth can be achieved *in vitro* in the presence of essential nutrients (Tupy *et al.*, 1991).

When the centrally located microspore nucleus has been positioned at a specific location near the cell periphery, the microspore undergoes an asymmetric division called pollen mitosis I (PMI) (fig 1.1). This results in the formation of two unequal cells, the generative and vegetative cells which are contained within the wall of the original microspore. Since uninucleate microspores can be more readily induced than postmitotic pollen grains to enter a sporophytic developmental program (Gaillard *et al.*, 1991), PMI is considered to represent a critical point in the commitment of the developing pollen grain to the gametophytic pathway. Recent studies by Eady *et al.*, (1995) demonstrated that late pollen genes are activated independently of the processes of mitosis and cytokinesis, which occur at PMI. Furthermore, the *lat*52 promoter is activated in uninucleate microspores immediately prior to PMI in transgenic *Arabidopsis* (Eady *et al.*, 1994). Taken together, these data suggest that the activation of the late pollen genes irreversibly commits the developing pollen grain to the gametophytic pathway. In addition, a possible role of the asymmetric division, PMI is to silence late pollen gene expression in the generative cell rather than activate late pollen gene expression in the vegetative cell.

After PMI the amount of cytoplasm within the vegetative cell rapidly increases (Tupy *et al.*, 1983). The generative cell is detached from the inner wall of the microspore and moves inwards to occupy a more central position within the cell, in the process becoming completely enclosed within the vegetative cytoplasm (fig. 1.1). Twell, (1995) demonstrated using a cell ablation technique designed to specifically inactivate the vegetative cell, that the detachment or migration of the generative cell away from the inner wall was dependent upon the vegetative cell. Moreover, the generative cell lost viability in the absence of a functional vegetative cell. Thus, the generative cell is totally dependent on vegetative cell functions.

The vegetative cell constitutes the bulk of the young pollen grain, it is densely cytoplasmic, containing numerous organelles, storage products, and is rich in both RNA

and protein. In contrast, the generative cell inherits a very small amount of the microspore cytoplasm, organelles or storage products. Whereas the chromatin in the nucleus of the generative cell is highly condensed, the vegetative nucleus is larger, has more nuclear pores and has decondensed chromatin. This would suggest that the vegetative nucleus is more transcriptionally active than the generative cell (Wagner *et al.*, 1990). *In situ* hybridizations to pollen with several pollen-expressed mRNAs (Hanson *et al.*, 1989; Ursin *et al.*, 1989) indicated that these genes are expressed specifically within the vegetative cell, but for only the *NTP303* gene (Weterings, 1994) was the resolution sufficient to exclude expression from within the generative cell. Analysis of promoter-*gus*-nuclear targeting signal fusions in transgenic plants demonstrated vegetative cell-specific expression for *lat*52 (Twell, 1992) and *NTP303* (Weterings, 1994). Furthermore, Blomstedt *et al.*, (1996) recently showed that the generative cell contains translatable mRNA and functional protein synthesis machinery.

Post-PMI leading up to anthesis, the young pollen grain undergoes a rapid increase in metabolic activity. Schrauwen *et al.*, (1990) determined the rate of mRNA synthesis during pollen development. This analysis showed that mRNA content of a single tobacco pollen grain increased by 13-fold from the late-uninucleate to the mid-binucleate stage of pollen development, with a further increase of 65 % up to anthesis. Also seen is a rapid increase in protein synthesis (Tupy *et al.*, 1983; Eady *et al.*, 1994) and the deposition of starch within the vegetative cytoplasm (Tupy *et al.*, 1983).

At the final stages of maturation the pollen grain dehydrates to a water content ranging from less than 6 % for species of *Populus* to 35 % and higher for grass species, of dry weight at anthesis (Heslop-Harrison, 1987). The level of dehydration within a pollen grain may reflect the time span for which the pollen grain remains viable. The pollen of *Cucurbita pepo* does not dehydrate prior to dispersal which results in a very short time period of pollen viability (Pacini *et al.*, 1997).

During the period of pollen development from PMI up to anthesis in *Brassica oleracea* tapetal cells undergo ultrastructural changes which eventually result in the breakdown of the cell membrane (Murgia *et al.*, 1991). Prior to anthesis the tapetal cell contents are released into the locule and the mature pollen grains are surrounded by disorganized lipid bodies and plastoglobuli (pollenkitt). These substances may function to protect the pollen grain from dehydration or to allow it to adhere to insect pollinators.

At maturity, the male gametophyte consists of either two or three cells (fig. 1.1), the vegetative cell and a generative cell (bicellular species), or a vegetative cell and two sperm cells (tricellular species), which lie within the cytoplasm of the vegetative cell (Mascarenhas, 1989). In bicellular species the second mitotic division (PMII) (fig. 1.1), in which the generative cell divides to form two sperm cells occurs during pollen germination. When the pollen grain lands on the stigma surface it rehydrates and a pollen

tube emerges from a pore on the exine surface. As the pollen tube elongates through the transmitting tissue to the ovule, the pollen cytoplasm, vegetative nucleus and sperm cells are transported within the tip (Heslop-Harrison, 1987).

1.2 The complexity of gene expression during pollen development

In order to understand the complexity of gene expression during pollen development initial studies concentrated on determining the overlap of gene expression between mature pollen and sporophytic tissues. Estimates of the number of genes expressed in the mature pollen grain were derived from the kinetics of hybridisation of ^{3}H-cDNA with poly (A) RNA in excess in *Tradescantia paludosa* (Willing and Mascarenhas 1984) and in maize (Willing *et al.*, 1988). This analysis has shown that in mature pollen of *Tradescantia paludosa* there are approximately 20,000 different sequences, which were divided into three classes based on their relative abundance. The first class of mRNAs represented around 40 different sequences which were extremely abundant, each present at about 26,000 copies per pollen grain. The second class constituted the major fraction of mRNA and contained 1,400 different sequences present at about 3,400 copies per pollen grain. The third class represented the least abundant fraction and consists of 18,000 different sequences at about 100 copies per pollen grain. In comparison, there were approximately 30,000 different sequences within the *Tradescantia* shoot, with the least abundant class expressed at levels 10- to 20-fold lower than the corresponding class in mature pollen. The higher complexity of mRNAs present in shoots than in pollen is probably due to the shoot containing many different cell types. Analysis of the abundance and complexity of mRNA sequences present within maize pollen and shoots closely reflects the pattern observed for *Tradescantia*.

Heterologous hybridizations between pollen cDNA and shoot poly (A) RNA and between shoot cDNA and pollen poly (A) RNA suggest that in *Tradescantia* and maize a minimum of 64 (Willing and Mascarenhas 1984) and 65 % (Willing *et al.*, 1988) respectively, of the pollen mRNA is present in shoot RNA. Further data obtained from colony hybridizations with cDNA libraries constructed from pollen poly (A) RNA and hybridized with cDNAs from pollen and vegetative tissues, indicated that of the different mRNA sequences present within mature pollen of maize and *Tradescantia* about 10 and 20 % respectively, could be pollen-specific (Stinson *et al.*, 1987; Mascarenhas, 1990). These analyses suggest that a large proportion of the genes expressed in mature pollen are also expressed in sporophytic tissues. However, examples exist where individual members of multi-gene families are expressed in a tissue-specific manner. Perhaps one of the clearest examples of this phenomenon is the eukaryotic initiation factor eIF-4A multi-gene family.

Members of this gene family are expressed in photosynthetic tissues (Owttrim *et al.*, 1991) and one member is expressed specifically in pollen (Brander and Kuhlemeier, 1995). Such is the sequence similarity between members of this gene family, that transcripts from these genes can only be distinguished from each other by using gene-specific probes designed to non-coding regions (Brander and Kuhlemeier, 1995). Other examples include poly (A) binding protein (Belostotsky and Meagher, 1993; 1996) and actin depolymerizing factor-like proteins (Lopez *et al.*, 1996). Thus, the number of genes specifically expressed in mature pollen may be greater than originally thought.

The above analyses by Willing and Mascarenhas (1984) and Willing *et al.*, (1988) provided only an incomplete picture regarding the complexity of gene expression during pollen development since analyses were performed on mature pollen. It has been reported that tobacco anthers containing microspores at an early developmental stage, contain approximately 26,000 different mRNA sequences of which about 11,000 are anther-specific (Kamalay and Goldberg, 1980; 1984). The proportion of the 11,000 anther-specific sequences which are likely to be microspore-specific remains unclear since during the stage of development assayed, the tapetum is still transcriptionally active. Schrauwen *et al.*, (1990) examined the changes in the mRNA population during tobacco and lily pollen development by extraction and *in vitro* translation of mRNA from distinct developmental stages. From this analysis transcripts could be divided into three broad groups based upon their pattern of accumulation during development. The first group consisted of mRNA species expressed at a constant level from late uninucleate microspores up to anthesis. The second group contained mRNAs that are expressed after PMI and accumulated up to anthesis. The third group of mRNAs are strongly expressed prior to PMI with a subsequent decrease in expression levels towards anthesis.

Taken together, the above analyses strongly indicates a complex pattern of gene expression during pollen development. In addition to acquiring pollen-specific isoforms of sporophytic genes the developing pollen grain utilises genes that are also expressed in sporophytic tissues. Coupled with this, the developing pollen grain expresses a unique set of developmentally regulated genes.

1.3 Isolation and characterization of genes expressed during microspore development

Microspore development begins with the formation of the tetrad and ends with an asymmetric nuclear division, PMI (section 1.1). The differential screening of cDNA libraries constructed from poly A+ isolated from whole anthers at different developmental stages has led to the isolation of a number of genes which are preferentially expressed during microspore development (reviewed by Twell, 1994; Weterings, 1994). *In situ*

localization and the analysis of promoter reporter gene fusions in transgenic plants showed that genes expressed in the developing microspore were also expressed in sporophytic tissues. For example *APG* (Roberts *et al.*, 1993a) is expressed in the tapetum, stomium and anther wall; *BA42* (Shen and Hsu, 1992) is expressed in the peripheral cells of the vascular bundle and tapetum; and *E2* (Foster *et al.*, 1992) is expressed in the tapetum. Other genes have been isolated that are expressed specifically in anthers containing developing microspores (Smith *et al.*, 1990; Tsuchiya et al., 1992), but it remains to be established whether these genes are microspore-specific. Smith *et al.*, (1990) showed that at least 3 out of the 10 stamen-specific genes isolated were expressed specifically in the tapetum. Moreover, the three tapetum-specific genes represented 0.63 % of the cDNA library prepared from stamen poly A+. Therefore, to overcome the transcriptional dominance of the tapetum and to reduce the likelihood of isolating genes which are expressed in both microspores and sporophytic tissues it would be desirable to construct cDNA libraries using poly A+ isolated specifically from microspores.

The differential screening of cDNA libraries constructed from poly A+ isolated from microspores has led to the isolation and characterization of the genes *Bp4* (Albani *et al.*, 1990), *Bp19* (Albani *et al.*, 1991) and *NTM19* (Oldenhof *et al.*, 1996). *Bp4* and *Bp19* are expressed at a maximum level during early microspore development with transcript levels decreasing considerably after PMI. In addition, transcripts homologous to *Bp4* and *Bp19* have been detected at low levels in sporophytic tissues, but since both genes belong to large gene families it is possible that sporophytic expression could originate from other family members. *NTM19* is specifically expressed in the unicellular microspore and to date, remains the only example of a microspore-specific gene.

The function of genes expressed during microspore development are barely known, but comparison with previously described genes or proteins has provided putative roles. Flavanols are required for pollen germination and tube growth in petunia and maize (Mo *et al.*, 1992). The genes *BA42* (Shen *et al.*, 1992) and *F3H* (Deboo *et al.*, 1995) share significant homology to the enzymes required for flavonoid biosynthesis, chalcone synthase and flavanone 3-hydroxylase respectively. *Bp19* shares identity to the pectin esterases of tomato and *Erwinia chrysanthemi* (Albani *et al.*, 1991). The genes *I3* (Roberts *et al.*, 1991; Roberts *et al.*, 1993b) and *E2* (Foster *et al.*, 1992) resemble oleosins and phospholipid transfer proteins respectively and may be involved in the lipid storage pathway.

1.4 Isolation and characterization of genes expressed during pollen maturation and germination

Pollen maturation and germination begins with the unequal asymmetric division

(PMI) of the microspore and ends with the delivery of two sperm cells to the egg and central cells within the embryo sac (section 1.1). Post-PMI leading up to anthesis, the young pollen grain undergoes a rapid increase in mRNA synthesis. The genes responsible for this increase in mRNA were defined as late pollen genes by Stinson *et al.*, (1987). To date, the late pollen genes represent the largest group of isolated genes which are expressed during pollen development (Twell, 1994; Weterings, 1994). This most likely results from a combination of the accessibility of mature pollen, or when whole anthers were used as a source of mRNA for cDNA library construction, the transcriptional dominance of the pollen genes over the sporophytic tissue.

The differential screening of cDNA libraries constructed from poly A+ isolated from mature pollen of *Tradescantia paludosa* and *Zea mays* facilitated the isolation of the pollen-specific cDNA clones *pTpc44*, *pTpc70* (*Tradescantia*) and *pZmc13*, *pZmc30* (*Zea mays*) (Stinson *et al.*, 1987). The technique of differential screening of cDNA libraries constructed from poly A+ isolated from mature pollen, mature anthers and floral tissues has been widely applied to many plant species. cDNA clones preferentially expressed in mature pollen have been isolated from *Lycopersicon esculentum* (McCormick *et al.* 1987), *Nicotiana tabacum* (Weterings *et al.*, 1992; Rogers *et al.*, 1992; Bucher *et al.*, 1995), *Oenothera organensis* (Brown and Crouch, 1990), *Brassica campestris* (Theerakulpisut *et al.*, 1991), *Sorghum bicolor* (Pe *et al.*, 1994), *Brassica napus* (Treacy *et al.*, 1997), *Helianthus annuus* (Baltz *et al.*, 1992), *Medicago sativa* (Wu *et al.*, 1996), *Oryza sativa* (Zou *et al.*, 1994), *Zea mays* (Allen and Lonsdale, 1993; Lopez *et al.*, 1996) and *Petunia inflata* (Mu *et al.*, 1994).

Putative functions for the late pollen genes have been designated based upon their homology with previously characterised genes and proteins. The tomato *lat*52 gene (McCormick *et al.*, 1987), shows significant homology to cDNAs isolated from maize (Hanson *et al.*, 1989), rice (Zou *et al.*, 1994), sorghum (Pe *et al.*, 1994), lilac (Batanero *et al.*, 1994) and olive (Lombardero *et al.*, 1994). The *lat*52 gene encodes an abundant cysteine-rich protein present in the mature pollen grain which is distantly related to kunitz trypsin inhibitor family (McCormick *et al.*, 1991). The *lat*52 homolog in olive encodes the major pollen-allergen, OLE EI, which is localised to the endoplasmic reticulum in mature and germinating pollen (Rodriguez-Garcia *et al.*, 1995) and accounts for approximately 1% of the total dry mass of mature pollen (Villalba *et al.*, 1993). Other examples include catabolic enzymes like ascorbate oxidase (McCormick *et al.*, 1987; Albani *et al.*, 1992; Weterings *et al.*, 1992), pectate esterase (Mu *et al.*, 1994), pectate lyase (Stinson *et al.*, 1987; Wing *et al.*, 1989; Rafnar *et al.*, 1991; Rogers *et al.*, 1992; Wu *et al.*, 1996) and polygalacturonase (Brown and Crouch, 1990; Niogret *et al.*, 1991; Allen and Lonsdale, 1993; Tebbutt *et al.*, 1994). Cytoskeletal proteins such as actin (Thangavelu *et al.*, 1993), α-tubulin (Carpenter *et al.*, 1992), ß-tubulin (Villemur *et al.*, 1994) and profilin (Valenta *et*

al., 1991). It remains to be established if the above genes actually encode for such functional proteins. Moreover, essential functional roles in pollen maturation and germination have only been demonstrated for *lat*52 (Muschietti *et al.*, 1994) and *Bcp1* (Xu *et al.*, 1995).

The spatial expression patterns of the late pollen genes have been investigated using *in situ* localisation, northern blot and RNA dot blot analyses to determine the presence of endogenous transcript. From these analyses the following genes have been shown to be specifically expressed in pollen: *NTP303* (Weterings *et al.*, 1992); *ZmABP1* and *ZmABP2* (Lopez *et al.*, 1996); *TobPDC2* (Bucher *et al.*, 1995); *Bmn1* (Treacy *et al.*, 1997); *G10* (Rogers *et al.*, 1992); *SF3* (Baltz *et al.*, 1992); *pMSb8* and *pMSb2.1* (Pe *et al.*, 1994); *Npg1* (Tebbutt *et al.*, 1994); *W2247* (Allen and Lonsdale, 1993) *Bp10* (Albani *et al.*, 1992); *Zmc13* (Hanson *et al.*, 1989); *pTpc44*, *pTpc70* and *pZmc30* (Stinson *et al.*, 1987). Northern blot analyses demonstrated that the endogenous transcripts of the late anther tomato (*lat*) clones *lat*51, *lat*52, *lat*56, *lat*58 and *lat*59, are expressed within the mature anther (Twell *et al.*, 1989b; Ursin *et al.*, 1989; Wing *et al.*, 1989). Although *in situ* localisation of mRNA showed that transcript was detectable in all cell layers within the anther wall for all *lat* genes (Ursin *et al.*, 1989), the majority of mRNA in mature anthers was attributed to expression in mature pollen (Twell *et al.*, 1989b; Ursin *et al.*, 1989). In addition, northern blot analysis of endogenous *lat*52 transcript showed low but detectable levels of *lat*52 mRNA in petals (Twell *et al.*, 1989b). Other genes which are expressed in pollen and to a lesser extent in sporophytic tissues include *ADH* (Bucher *et al.*, 1995) and *PPE1* (Mu *et al.*, 1994).

Investigation of the accumulation of steady state *lat*52 mRNA during anther development showed that transcript was first detectable in anthers containing spores undergoing PMI, leading to a substantial increase in transcript levels in anthers containing mature pollen (Twell *et al.*, 1989b). Similar developmental accumulation profiles were observed for the other *lat* genes (Ursin *et al.*, 1989; Wing *et al.*, 1989) and *pTpc44* and *pTpc70* (Stinson *et al.*, 1987). In addition, Stinson *et al.*, (1987) showed that *pTpc44* and *pTpc70* transcript levels decreased during pollen germination. Analysis of the accumulation of steady state *Bp10* (Albani *et al.*, 1992) and *ADH* (Bucher *et al.*, 1995) mRNA demonstrated that transcripts were detectable at low levels during microspore development, reached maximal levels at PMI and mid-pollen development respectively with considerably lower levels seen for both transcripts during the later stages of pollen development. In contrast to the aforementioned genes, transcripts from *NTP303* (Weterings *et al.*, 1992), *pZmc30* (Stinson *et al.*, 1987), *Bnm1* (Treacy *et al.*, 1997) and *P6* (Brown and Crouch, 1990) are first detectable during mid-pollen development or later, leading to a substantial increase in transcript levels in mature pollen. Furthermore, Weterings *et al.*, (1992) clearly demonstrated that *NTP303* was actively transcribed at the

early stages of pollen germination. Taken together, these analyses indicate differential developmental regulation of the late pollen genes.

In conjunction with *in situ* localisation, northern blot and RNA dot blot analyses of endogenous transcript, the analyses of promoter reporter gene fusions in transgenic plants has been utilised to provide further information on the spatial and temporal expression patterns of the late pollen genes. The 5' flanking region of the tomato *lat*52 (-492 to +110) and tomato *lat*59 (-1305 to +91) genes when fused to the reporter gene *gus* were sufficient to direct GUS activity specifically in mature pollen within anthers of transgenic tomato (Twell *et al.*, 1990). Furthermore, in transgenic tomato the developmental accumulation of GUS activity in pollen closely reflected the accumulation of the native *lat*52 and *lat*59 transcripts (Twell *et al.*, 1990). Histochemical localisation of GUS activity in sporophytic tissues showed that the *lat*52 and *lat*59 promoters were expressed within the endosperm of immature and mature seeds (Twell *et al.*, 1991). In addition, the *lat*59 promoter also directed GUS activity in the testa and root cap of primary and lateral roots (Twell *et al.*, 1991). In contrast to previous northern blot analysis and *in situ* localisation of endogenous *lat*52 and *lat*59 transcripts (Twell *et al.*, 1989b: Ursin *et al.*, 1989), no significant GUS activity above background was detectable within the mature anther wall for the *lat*52 and *lat*59 promoters (Twell *et al.*, 1990) and in petals for the *lat*52 promoter (Twell *et al.*, 1991). Analysis of *TUA1* promoter *gus* fusions in transgenic *Arabidopsis* also identified previously undetected promoter activity within sporophytic tissues (Carpenter *et al.*, 1992). However, a strict correlation between the spatial and temporal accumulation of endogenous transcript with the expression of promoter-*gus* fusions in transgenic *Brassica napus* and tobacco has been demonstrated for the *Brassica napus Bmn1* (Treacy *et al.*, 1997) and tobacco *Npg1* (Tebbutt *et al.*, 1994) genes respectively.

The *lat*52 promoter fused to the reporter gene *gus* directed GUS activity specifically in mature pollen within anthers of transgenic tobacco and *Arabidopsis* (Twell *et al.*, 1990; Eady *et al.*, 1994). Furthermore, Nishihara *et al.*, (1993) demonstrated *lat*52 promoter activity in lily pollen by particle bombardment. In addition, the promoters of the monocot genes *PS1* (Zou *et al.*, 1994), *Zm13* (Guerrero *et al.*, 1990) and *W2247* (Allen and Lonsdale, 1993) fused to *gus* all conferred their correct spatial and temporal expression patterns in transgenic tobacco plants. These analyses suggest that regulatory mechanisms for late pollen promoter activation are conserved among monocotyledonous and dicotyledonous species independent of whether pollen is shed in a bicellular or tricellular condition.

The analyses of promoter reporter gene fusions in transgenic plants has led to the identification of numerous plant genes which are expressed in sporophytic and gametophytic tissues. Through histochemical localization and quantitative fluorometric determination of GUS activity the promoters of the following genes were shown to be

active in mature pollen as well as sporophytic tissues: *Arabidopsis* acyl carrier protein (Baerson and Lamppa, 1993; Baerson *et al.*, 1994); *CDeT27-45* (Michel *et al.*, 1993); potato *Sus3* sucrose synthase (Fu *et al.*, 1995); soybean glutamine synthetase (Marsolier *et al.*, 1993); *Arabidopsis HMG2* encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase (Enjuto *et al.*, 1995); potato *hmg1* encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase (Bhattacharyya *et al.*, 1995); *Arabidopsis Rab2* (Moore *et al.*, 1997); rice chitinase *RCH10* (Zhu *et al.*, 1993); Arabidopsis gibberellin biosynthetic gene *GA1* (Silverstone *et al.*, 1997). With the exception of *Arabidopsis Rab2*, the presence of endogenous transcript in mature pollen remains to be established for the other genes.

1.5 Transcriptional regulation of gene expression in plants

1.5.1 TATA box: RNA polymerase II and general transcription factors

The transcription of genes encoding protein molecules is mediated by RNA polymerase II. The minimal promoter region of the majority of RNA polymerase II transcribed genes contain an initiator *cis*-element that overlaps the transcription start site and a TATA box located ~ 30 bp upstream. The TATA box is an AT rich region, with the eukaryotic consensus sequence TATAAA, which closely corresponds to TATA box sequences present within many plant genes (Joshi, 1987). Accurate basal transcription of the rice *PAL* minimal promoter *in vitro* requires both the TATA box for transcriptional activity and the initiator element for accurate placement of the transcription start site (Zhu *et al.*, 1995). In addition, the functional interaction between these two *cis*-elements is critically dependent on their spacing within the minimal *PAL* promoter. Furthermore, Kloeckener-Gruissem *et al.*, (1992) showed that duplication and deletion of the TATA box region of maize *Adh1* affected its organ-specific expression.

Before a gene can be transcribed by RNA polymerase II, a complex consisting of more than 20 proteins must be assembled on a promoter (reviewed by Buratowski, 1994; Eloranta and Goodbourn, 1996). The first step involves TFIID binding to the TATA element of a promoter. TFIID consists of a single TATA-binding protein (TBP) molecule and multiple associated factors (TAFs). Genes encoding TBPs have been cloned from *Arabidopsis* (Gasch *et al.*, 1990), maize (Haass and Feix, 1992), potato (Holdsworth *et al.*, 1992), tobacco (Iwataki *et al.*, 1997) and wheat (Kawata *et al.*, 1992), and all share significant homology with the carboxy-terminal region of TBPs from human, yeast and *Drosophila*. Indeed, the maize TBPs *Tbp1* and *Tbp2* are functionally interchangeable with yeast TBP for conferring yeast cell viability (Vogel *et al.*, 1993). Furthermore, tobacco TBP can be substituted by recombinant human TBP in heat-inactivated extracts prepared from tobacco nuclei (Iwataki *et al.*, 1997).

Genes encoding TAFs have been isolated from human, yeast and *Drosophila* (reviewed by Tijan and Maniatis, 1994; Tansey and Herr, 1997). *Drosophila* TFIID has been entirely reconstituted from recombinant proteins: TBP and eight TAFS, of which 7 are conserved in humans (reviewed by Tansey and Herr, 1997). No literature is currently available for plant TAFs, but since maize TBP is functional in yeast (Vogel *et al.*, 1993) and homologous TAFs have been isolated from human, yeast and *Drosophila* one would expect homologs to yeast TAFs to be present in plants.

The TBP component of TFIID interacts with the minor groove of the TATA box which distorts the DNA. Holdsworth *et al.*, (1992) demonstrated *in vitro* that recombinant potato TBP specifically recognised the TATA box within the class-1 potato patatin gene. In addition, recombinant maize TBP has been shown to form *in vitro* complexes with the proximal zein promoter *P2* and the *CaMV35S* promoter (Haass *et al.*, 1994). In contrast, for recombinant maize TBP to form *in vitro* complexes with the distal zein promoter *P1*, the shrunken promoter and a maize rRNA gene promoter the assistance of a yeast protein fraction with TFIIA activity is required (Haass *et al.*, 1994). TFIID possesses two known catalytic activities histone acetylation (reviewed by Wade and Wolffe, 1997) and basal-factor phosphorylation. RNA polymerase II is recruited as part of a preassembled holoenzyme complex containing many of the general transcription factors such as TFIIB, TFIIF and TFIIH, which recognises the TFIID-promoter complex (reviewed by Greenblatt, 1997).

1.5.2 *Cis*-regulatory elements and transcription factors

The differentiated state of higher plants is achieved via a complex pattern of gene expression. Analysis of nuclear RNAs in tobacco tissues showed that a large proportion was represented in the nuclear RNA of the entire plant, with a significant fraction (10 to 40 %) proposed to be organ-specific (Kamalay and Goldberg, 1984). Spatial and temporal gene expression is primarily regulated at the level of transcription. This is mediated by sequence-specific DNA binding proteins (transcription factors) which interact directly or via intermediary proteins, with the basal transcription machinery (section 1.5.1).

Initial studies designed to identify transcription factor binding sites (*cis*-regulatory elements) within plant promoters were performed on the cauliflower mosaic virus 35S RNA promoter (*CaMV35S*). Analysis of 5' and 3' *CaMV35S* promoter deletion mutants in transgenic tobacco demonstrated that the majority of promoter activity was provided by the region -343 to -46 (Fang *et al.*, 1989). Furthermore, the region -343 to -46 was subdivided into three functional regions: -343 to -208, -208 to -90 and 90 to -46. Transient expression analysis of 5' *CaMV35S* promoter deletion mutants in tobacco protoplasts further defined the functional regions to -134 to -108 and -89 to -73 (Ow *et al.*,

1987). The spatial and temporal expression pattern of the *CaMV35S* promoter has been determined in transgenic tobacco seeds and seedlings (Benfey *et al.*, 1989). This analysis showed that domain A (-90 to +8) directed GUS activity preferentially in the embryo radicle and the seedling root. Domain B (-343 to -90) directed GUS activity preferentially in the seed cotyledon and vascular and leaf tissues of the developing seedling. In addition, deletion from -90 to -72 virtually abolished GUS activity in seedlings.

DNAse I footprinting and gel mobility retardation assays demonstrated that a cellular factor from tobacco leaf tissue (ASF-1) recognises two tandem TGACG motifs within the region -82 to -62 of domain A (Lam *et al.*, 1989; Prat *et al.*, 1989). Moreover, substitution mutation of both motifs within the region -343 to +2 dramatically reduced promoter activity in stem and roots of transgenic tobacco plants. A tobacco cDNA clone has been isolated which encodes the putative transcription factor TGA1a, which has ASF-1 like activity (Katagiri *et al.*, 1989). Northern blot analysis showed that *TGA1a* was preferentially expressed in roots and to a lesser extent in leaves. A functional role has been demonstrated for TGA1a. Addition of recombinant TGA1a to an *in vitro* transcription system stimulated transcription by increasing the number of active preinitiation complexes (Yamazaki *et al.*, 1990). Furthermore, microinjection of recombinant TGA1a into transgenic tobacco cotyledon cells activated transgenes containing *as-1*-linked promoters (Neuhaus *et al.*, 1994). In addition to TGA1a, the putative leaf-specific transcription factors ASF-2 (Lam and Chua, 1989) and MNF1 (Yanagisawa and Izui, 1992) recognise the *CaMV35S* promoter regions -106 to -85 and -279 to -236 respectively.

One of the most well characterised native plant gene promoters in terms of defined *cis*-regulatory elements and the transcription factors which recognise these elements is pea *rbcS-3A*. *rbcS-3A* is regulated by light, restricted to chloroplast-containing cells and encodes the small subunit of ribulose-1,5-bisphosphate carboxylase which is the key enzyme in photosynthetic carbon assimilation. Substitution mutational analysis of the *rbcS-3A* promoter in transgenic tobacco identified two light-regulatory motifs, box II and box III, required for maximal expression of the region -175 to +22 (Kuhlemeier *et al.*, 1988). Moreover, a 2 bp substitution mutation within box II abolished activity of the promoter region -175 to +22. In addition, 4 copies of box II inserted upstream of the -90 *CaMV35S* promoter, is sufficient to confer light responsive expression in transgenic tobacco and *Arabidopsis* (Lam and Chua, 1990; Puente *et al.*, 1996). Gel retardation analysis of 2 bp substitution mutations throughout box II defined the core sequence GGTTAA as critical for the binding of GT-1 nuclear protein (Green *et al.*, 1988). Furthermore, the nuclear factor GT-1 was shown to recognise box III, and four other upstream boxes II*,III*, II** and III**.

A cDNA which encodes a putative transcription factor GT-1a, that interacts with box II has been isolated (Gilmartin *et al.*, 1992; Perisic and Lam, 1992). Northern blot

analysis showed that *GT-1a* transcript was detectable in roots as well as photosynthetic tissues (Perisic and Lam, 1992). This would imply that GT-1a does not only mediate light responsive transcription of *rbcS-3A*. Indeed, nuclear GT-1-like factors have been shown to interact with non light-responsive genes such as the tobacco *PR-1a* gene (Buchel *et al.*, 1996) and with the pollen-specific *NTP303* gene (Hochstenbach *et al.*, 1996). The box VI motif from the pea *rbcS-3A* gene, has been shown to bind the transcription factor 3AF1 (Lam *et al.*, 1990). Northern blot analysis demonstrated that *3AF1* transcript was detectable in all tissues examined including anthers and non-photosynthetic tissues such as roots (Lam *et al.*, 1990). Therefore, the transcription factor (s) which regulate the light responsive transcription of *rbcS-3A* remain to be identified. cDNAs have been isolated which encode putative transcription factors which recognise light responsive gene promoters. These include *GT-2* (Dehesh *et al.*, 1990), *CPRF-1*, *CPRF-2* and *CPRF-3* (Weisshaar *et al.*, 1991) of which only the transcription of *CPRF-1* is regulated by light. Moreover, only GT-2 has been shown to activate gene expression in a binding site-dependent manner *in vivo* (Ni *et al.*, 1996).

Opaque-2 (*O2*) represents one of the most well characterised plant transcription factors. The *O2* gene was isolated via transposon tagging (Schmidt *et al.*, 1987; Motto *et al.*, 1988) and the corresponding cDNA encodes a protein which possesses the characteristic features of the basic domain/leucine zipper class of transcription factors (Hartings *et al.*, 1989). *O2* is specifically expressed in the endosperm of maize and its transcription is precisely controlled during seed development (Gallusci *et al.*, 1994). Analysis of steady state zein transcript accumulation within the endosperm of maize harbouring an *opaque-2* mutation demonstrated that α–zein transcription was severely reduced (Kodrzycki *et al.*, 1989). Moreover, the expression of genes encoding 22 KDa zein proteins was almost abolished. *In vitro* DNAse I footprinting and gel mobility shift assays with recombinant O2 protein localised multiple O2 binding sites in the promoter of 22 KDa zein genes (Schmidt *et al.*, 1992; Muth *et al.*, 1996). Furthermore, *O2* has been shown to transactivate 22 KDa zein promoter-reporter gene fusions in tobacco protoplasts and yeast. *O2* has also been shown to transactivate *in vivo* the promoters of the seed-specific genes *b-32* (Lohmer *et al.*, 1991), 27 KDa zein (Ueda *et al.*, 1992), 22-KDa-like *α-prolamin* (Yunes *et al.*, 1994) and *2S* storage protein (Vincentz *et al.*, 1997) in addition to its own promoter (Lohmer *et al.*, 1991). Interestingly, these analyses have also shown that O2 protein recognises different *cis*-regulatory elements within seed-specific promoters. Recently, a maize zinc-finger protein has been isolated which binds to the prolamin box in zein gene promoters and interacts *in vitro* with O2 protein (Vicente-Carbajosa *et al.*, 1997).

1.6 Transcriptional regulation of gene expression during pollen development

The differential expression patterns of late pollen genes during pollen maturation and pollen germination (section 1.4) would suggest specific mechanisms of regulation for groups of late pollen genes. In order to understand the pathway of late pollen gene regulation it is important to define the *cis*-regulatory elements and isolate the *trans*-acting factors responsible for transcriptional activation and their specific expression in the developing pollen grain.

To date only a small number of late pollen gene promoters have been functionally characterised. Initial studies were first performed on the promoter regions of the *lat* genes *lat*52, *lat*56 and *lat*59 (Twell *et al.*, 1991). From sequence analysis of the *lat*52 promoter, the core motif TGTGGTT, which is closely related to the *rbcS-3A* Box II *cis*-regulatory element (Green *et al.*, 1988; Kuhlemeier *et al.*, 1988) was identified at positions -438 (PBIII), -178 (PBII) and -96 (PBI). To test if the PB core motifs contributed to *lat*52 transcription 5' promoter deletion mutant-*gus* fusions were stably introduced into tomato. This analysis led to the identification of two upstream positive *cis*-regulatory regions which contribute to the activation of the *lat*52 promoter in anthers (fig. 1.2): region 52-A -492 to -125 (containing PBIII and PBII) and region 52-B -124 to -72 (containing PBI). This analysis also defined the minimal proximal promoter region sufficient to direct preferential reporter gene activity in anthers as the region 52-C (-71 to +110) (fig. 1.2).

A more detailed study of the region -492 to +110 was performed using a transient expression assay system in tobacco pollen (Twell *et al.*, 1991). The region 52-A was sub-divided into a redundant domain A1 containing PBIII, and a functional domain A2 containing PBII (fig 1.2). This analysis further localised the *cis*-regulatory element PBI to the region -100 to -87 and redefined the minimal pollen-specific proximal promoter region as -41 to +110 (fig 1.2). Gain of function analyses showed that both PBII (-194 to -176) (Twell *et al.*, 1991) and the 30 bp proximal promoter region (-84 to -55) (Eyal *et al.*, 1995) conferred pollen-specific expression when linked to the heterologous *CaMV35S* core promoter.

Analysis of *lat*59 5' promoter deletion mutant-*gus* fusions in transgenic tomato led to the identification of three positive and one negative *cis*-regulatory regions which regulate the expression of the *lat*59 promoter in mature anthers (fig. 1.2) (Twell *et al.*, 1991). Positive *cis*-regulatory regions are located at -1305 to -805 (region 59-A), -141 to 116 and -115 to -46 (fig 1.2). The negative *cis*-regulatory region is located between -804 to -142. Since progressive 5' promoter deletions from -804 to -142 gradually restored promoter activity to the level observed for the -1305 5' promoter deletion mutant in anthers, it would imply that multiple negative *cis*-regulatory elements reside within the

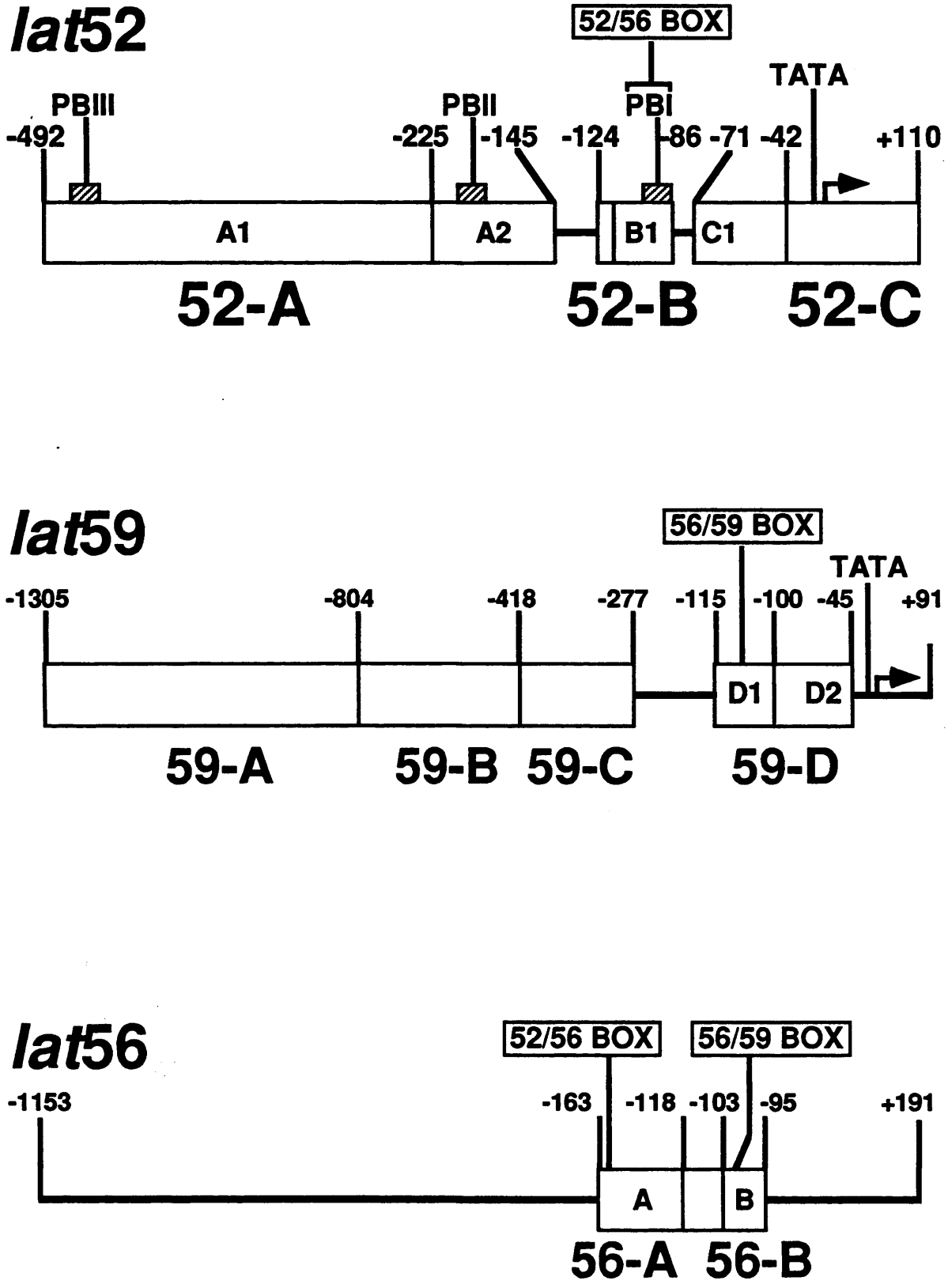


Figure 1.2. Functional maps of the *lat*52, *lat*56 and *lat*59 promoters as defined by Twell *et al.*, (1991).

region -804 to -142. This analysis also defined the minimal proximal promoter region sufficient to direct preferential reporter gene activity in anthers as the region -115 to +91 (fig. 1.2).

A more detailed study of the region -1305 to +91 was performed using a transient expression assay system in tobacco pollen (Twell *et al.*, 1991). This analysis subdivided the *lat*59 promoter into three major activating domains: -1305 to -278, -115 to -101 and -100 to -46 (fig. 1.2). Furthermore, this analysis redefined the minimal pollen-specific proximal promoter region as -45 to +91 (fig. 1.2). In contrast with the results observed in transgenic tomato anthers, deletion of the region -1305 to -805 did not decrease promoter activity. In addition, the negative influence seen for the region -804 to -142 in transgenic tomato anthers was not detected in the transient assay.

Detailed 5' promoter deletion analysis of the *lat*56 promoter using a transient expression assay system in tobacco pollen led to the identification of four positive cis-regulatory regions: -1153 to -176, -163 to -119, -118 to -104 and -103 to -95 (fig. 1.2) (Twell *et al.*, 1991). In addition, this analysis defined the minimal pollen-specific proximal promoter region as -103 to +191 (fig. 1.2).

Sequence analysis of the positive *cis*-regulatory regions defined using the transient expression assay system in tobacco pollen, identified shared motifs between the *lat* promoters (Twell *et al.*, 1991). The 52/56 box TGTGGTTATATA, is located at positions -96 to -85 and -166 to -155 within the *lat*52 and *lat*56 promoters respectively (fig. 1.2). The 56/59 box GAA$^{T}/_{A}$TTGTGA is located at positions -103 to -94 and -114 to -105 within the *lat*56 and *lat*59 promoters respectively (fig. 1.2). Targeted 5' deletion and substitution mutational analysis of these shared motifs within the *lat* promoters identified the core sequences GTGG and GTGA as essential for maximal activity of the 52/56 and 56/59 boxes respectively (Twell *et al.*, 1991).

Gain of function analyses showed that PBII (-194 to -176) (Twell *et al.*, 1991) and the *lat*52 (-84 to -55) and *lat*59 (-98 to -69) proximal promoter regions (Eyal *et al.*, 1995) conferred pollen-specific expression when linked to the heterologous *CaMV35S* core promoter. In addition, Eyal *et al.*, (1995) demonstrated using the transient expression assay system that multimers of the *lat*52 promoter region -84 to -55 could effectively compete for transcription factors with the *lat*52 and *lat*59 promoter regions -100 to +110 and -115 to +91 respectively, but not with the *lat*56 promoter region -164 to +191. These data suggest that the *lat*52 and *lat*59 proximal promoter regions are the target for a common transcription factor.

Although no extensive sequence homology to either the 52/56 or 56/59 boxes are present within other late pollen promoters, copies of the core motifs GTGG and GTGA are located within functional regions. Analysis of the *Zm13* promoter using the transient expression assay in *Tradescantia* pollen demonstrated that the promoter region -260 to -

101 was required for high level activity (Hamilton *et al.*, 1992). Contained within this region are the core motifs GTGG and a reverse copy of the GTGA motif. Additional sequences essential for pollen-specificity were retained within the region -101 to -54 which contains GTGG in reverse orientation (Hamilton *et al.*, 1992). Recent analysis of *Npg1* 5' promoter deletion mutants in tobacco pollen using the transient expression assay functionally defined the minimal promoter region required for pollen-specificity as -182 to +85 (Tebbutt and Lonsdale, 1995). Further 5' deletion analysis of the region -182 to +85 demonstrated that deletion from -182 to -86 was sufficient to abolish promoter activity in pollen. These analyses strongly indicated that sequences required for pollen-specific expression of *Npg1* reside within the region -182 to -87. Contained within this region are two copies of the GTGG motif. Also present is the PG box which was identified from sequence comparisons between the tobacco *Npg1* promoter and the promoter of its maize homolog (Allen and Lonsdale, 1993). Deletion analysis of the *Arabidopsis TUA1* promoter in transgenic *Arabidopsis* identified two upstream pollen-specific activator regions -1500 to -534 and -271 to -218 (Carpenter *et al.*, 1992). This analysis also showed that the minimal core promoter -97 to +56 was sufficient to direct gene expression specifically in pollen. Further deletion from -97 to -39 abolished promoter activity in pollen which indicated the presence of essential sequences within the region -97 to -40. The regions -271 to -218 and -97 to -40 both contain a single GTGA motif.

Detailed 5' promoter deletion analysis of the *NTP303* promoter using a transient expression assay system in tobacco pollen led to the identification of two positive *cis*-regulatory regions: -103 to -87 and -86 to -59 (Weterings *et al.*, 1995). In addition, this analysis defined the minimal pollen-specific proximal promoter region as -86 to +175. Furthermore, this analysis identified the novel *cis*-regulatory element AAATGA, of which the TGA triplet was shown to constitute an active part. Moreover, the promoter fragment -103 to -51 which contains two copies of the AAATGA motif, one forward, one reverse, activated a minimal *CaMV35S* promoter specifically in pollen. The sequence AAATGA and its position relative to the TATA box is completely conserved within the promoters of *NTP303* and its *Brassica napus* homolog *Bp10* (Albani *et al.*, 1992).

Even though the above analyses have identified the precise sequences of *cis*-regulatory elements which regulate pollen-specific activity of the late pollen promoters, the identity or properties of the transcription factors which recognise these elements remain unknown. Genes which encode putative transcription factors/nucleic acid binding proteins, which are expressed in pollen have been isolated. Heterologous screening of cDNA libraries using probes with sequence identity to known transcription factors has facilitated the isolation of a *myb*-related gene *Nt.myb* (Sweetman, 1996), two rice MADS box genes *OsMADS2* and *OsMADS4* (Chung *et al.*, 1995) and an *Antirrhinum* MADS-box gene *DEFH125* (Zachgo *et al.*, 1997). In addition, *SF3*, which encodes a putative

nucleic acid binding protein has been isolated by differential screening of a sunflower inflorescence cDNA library (Baltz *et al.*, 1992). Spatial and temporal analysis of endogenous transcript and or endogenous protein demonstrated that *Nt.myb* (Sweetman, 1996), *DEFH125* (Zachgo *et al.*, 1997) and *SF3* (Baltz *et al.*, 1992) were expressed predominantly within the developing pollen grain. *In situ* hybridization analyses showed that *OsMADS2* and *OsMADS4* were expressed mainly in pollen, tapetum and stigma (Chung *et al.*, 1995). Analysis of a *TGA6* promoter-*gus* fusion in transgenic *Arabidopsis* suggests that the DNA-binding protein TGA6 is expressed in mature pollen (Xiang *et al.*, 1997). Previous analysis using RT-PCR showed that GT-1-related transcript products were detectable in tobacco pollen (Eyal *et al.*, 1995). Sequence analysis of the RT-PCR products identified two different but homologous genes, one of which was proposed to encode the transcription factor GT-1a (Gilmartin *et al.*, 1992). Perhaps, the heterologous screening of cDNA libraries with sequences of known transcription factors represents the easiest method for the isolation of genes encoding transcription factors expressed in pollen.

1.7 Translational regulation of gene expression

The importance of transcriptional controls in regulating both the level and specificity of expression in plants is well established (section 1.5.2), commonly being determined by the distribution and combinatorial activities of sequence-specific DNA binding proteins. In contrast, the contribution of post-transcriptional mechanisms in controlling these processes is far less well understood. However, a few well documented examples exist in the literature of gene regulation at the level of translation, which demonstrate that the 5' untranslated regions (5'-UTRs) of certain viral and native cellular RNAs are used to modulate the rate of translational initiation. Specific regions of the 5'-UTRs of several animal viruses (polio, encephalomyocarditis and foot and mouth) from the picornavirus group, are recognised as internal ribosome entry sites (reviewed by Belsham and Sonenberg, 1996; Jackson *et al.*, 1996). Furthermore, specific sequences which mediate internal ribosome entry have been identified in the native animal cellular RNAs *Antennapedia* (Oh *et al.*, 1992) and immunoglobulin heavy-chain binding protein (Yang and Sarnow, 1997). In addition to internal ribosome entry, translational repression represents an alternative mechanism of translational control of cellular mRNAs. Translation of testis-specific superoxide dismutase mRNA and ferritin mRNA are regulated through the binding of a repressor protein to their respective 5'-UTRs (Gu and Hecht, 1996; Klausner *et al.*, 1993).

The major emphasis on post-transcriptional control in plants has been focused upon plant viruses. The ability of plant virus 5'-UTRs to enhance gene expression at the

level of translation is well documented (reviewed by Gallie, 1993; Turner and Foster, 1995). How this is achieved is not clear, although they appear to act at the level of translational initiation rather than elongation, RNA processing or mRNA stability. Moreover, cap-independent translational enhancement of gene expression mediated by certain plant viruses has been shown to be dependent on the 3'-UTR (Timmer *et al.*, 1993; Wang and Miller, 1995). Whether such sequences enhance translation differentially in different cell types and/or are modulated during development has not been thoroughly investigated. Alternative mechanisms of translational initiation are utilised by plant viruses, these include internal ribosomes entry (Basso *et al.*, 1994) and ribosome shunting (Schmidt-Puchta *et al.*, 1997).

Translational enhancers have only recently been identified within native plant genes. These include sequences within the 5'-UTR of the photosystem I gene *psaDb* (Yamamoto *et al.*, 1995) and the plasma membrane protein-ATPase *pma1* (Michelet *et al.*, 1994). Both have been shown to enhance reporter gene activity at the level of translation in either transgenic plants (*psaDb*) or from RNA transfection studies in protoplasts (*pma1*). But, enhancement was not shown to be independent of the linked coding sequence or the 3'-UTR. Moreover, no evidence for developmental or tissue-specific translational control was presented for either 5'-UTR. One example of a translational enhancer which does show some degree of tissue specificity is located within a barley α-amylase gene (Gallie and Young, 1994). From RNA transfection studies the α-amylase 5'- and 3'-UTR regions when fused to the reporter gene luciferase enhanced activity in maize aleurone and endosperm protoplasts but not in suspension cells. Other examples of 5'-UTR mediated enhancement of chimeric gene fusions in transgenic plants have been described (Marcotte *et al.*, 1989; Wong *et al.*, 1992; Kuhlemeier, 1992; Casper and Quail, 1993; Sullivan and Green, 1993). These remain to be rigorously investigated because in at least one case this has been shown to result from the presence of transcriptional regulatory elements within the 5'-UTR (Bolle *et al.*, 1994). A further example of translational control is one modulated by an environmental signal in which preferential translation of *adh1* (Fennoy and Bailey-Serres, 1995) occurs during hypoxia. Preferential translation of *adh1* under low-oxygen conditions was shown to be dependent on the presence of both 5' and 3' UTR mRNA sequences (Bailey-Serres and Dawe , 1996).

1.8 An overview of thesis aims

The overall aim of this thesis was to investigate the mechanisms which regulate late pollen gene expression during pollen development. At the onset of this thesis the tomato *lat52* gene represented one of the most well characterised examples of genes expressed during late pollen development (Twell *et al.*, 1989b; Ursin *et al.*, 1989; Twell *et al.*, 1990;

Twell *et al.*, 1991).

The first part of this analysis was to determine how the *lat*52 promoter was regulated during pollen development and to further investigate the mechanisms controlling pollen-specific expression. This involved the detailed analysis of *lat*52 5' promoter deletion mutant-*gus* fusions in transgenic tobacco plants (chapter 3).

The second part of this project was to delimit promoter sequences essential for pollen-specific expression of *lat*52 and in addition, determine the functional organisation of identified *cis*-regulatory elements. This involved gain of function and targeted substitution mutational analyses (chapter 4).

The final part of this thesis was to establish the role of the *lat*52 5'-UTR in controlling pollen-specific expression of *lat*52 by substituting it for a synthetic polylinker sequence. (chapter 5).

The information derived from these analyses are discussed in relation to the cell-specific control of the tomato *lat*52 gene and other known late pollen genes.

Chapter 2

Materials and methods

2.1 Sources of molecular biology reagents, enzymes and plant tissue/bacterial culture chemicals

Chemicals and reagents were purchased from Sigma Chemical Company Ltd. and BDH Ltd. Enzymes were obtained from Gibco BRL, Promega, Pharmacia and Boehringer Mannheim. Tissue culture salts and hormones were purchased from Flow Laboratories. Agar was obtained from Difco Laboratories and bacterial media from Unipath Ltd.

2.2 Plant tissues

2.2.1 Collection of dehisced *N.tabacum* cv Samsun pollen

Mature pollen used for particle bombardment (section 2.9.5) was collected by hoovering the open flowers from approximately forty mature greenhouse grown tobacco plants using a vacuum cleaner. Pollen was retained on a 20 micron mesh contained within the vacuum cleaner nozzle. When all the open flowers had been hoovered the pollen was scraped off the 20 micron mesh into a 1.5 ml microfuge tube and stored at -70 °C until required.

2.2.2 Surface sterilization of *N.tabacum* cv Samsun tissues

2.2.2.1 Surface Sterilization of Leaves

Mature fully expanded leaves were excised at the petiole from six-week old plants grown under normal greenhouse conditions. In a laminar flow hood the leaves were placed in a large sterile casserole dish. Following submersion in 500 ml of domestos diluted to 10 % v/v with tap water the leaves were incubated for 15 min with occasional agitation. The leaves were then transferred to a new sterile casserole dish and washed four times with 400 ml aliquots of sterile water to remove all traces of domestos.

2.2.2.2 Surface sterilization of seeds

Approximately 200 seeds were placed into a 5 ml plastic bijou bottle, 1 ml of 70 % v/v ethanol was added and the seeds shaken for 1 min. The 70 % v/v ethanol was removed using a p 1000 gilson pipette and 2 ml of domestos diluted to 10 % v/v with tap water was added. The seeds were incubated in the presence of domestos with occasional shaking for 15 min. The domestos was removed using a p 1000 gilson pipette and the

seeds subsequently washed four times with 2 ml aliquots of sterile water.

2.2.3 Generation and propagation of *in vitro N.tabacum* cv Samsun shoot cultures

MS30 media: per litre

Muarashige and Skoog (MS) salts	4.71 g
Sucrose	30 g

pH adjusted to 5.6 - 5.8 with 1 M KOH
Autoclaved 15 min at 120 °C

MS30 media was solidified by the addition of 0.8 % w/v agar and autoclaved as above.

Approximately 30 seeds, surface sterilised as in section 2.2.2.2, were plated out onto a 9 cm petri dish containing solid MS30 media. The petri dish was sealed with whatman tape and incubated at 25 °C in continuous white light. After 3 weeks single seedlings were transferred to magenta pots containing solid MS30 media and grown for a further 6 weeks at 25 °C in continuous white light. *In vitro* shoot cultures were propagated by excising axillary nodes from the 6 week old plants. Each axillary node was transferred to a new magenta pot containing solid MS30 media and grown at 25 °C in continuous white light .

2.2.4 Generation and maintenance of *N.tabacum* cv Samsun cell suspension cultures

Callus inducing media: per litre

MS30 media (section 2.2.4)	
NAA	2 mg
BAP	2 mg

pH adjusted to 5.6 - 5.8 with 1 M KOH

Media was solidified by the addition of 0.8 % w/v agar and autoclaved for 15 min at 120 °C

UM media: per litre

MS salts	4.71 g
sucrose	30 g
caesin hydrolysate	2 g
thiamine HCl	9.9 mg
pyridoxine HCl	9.5 mg
nicotinic acid	4.5 mg
2,4 D	2 mg
kinetin	0.25 mg

pH adjusted to 5.6 - 5.8 with 1 M KOH
Autoclaved 15 min at 120 °C

N.tabacum cv. Samsun cell suspension cultures were generated by placing 2 cm by 2 cm sterile leaf squares isolated from *in vitro* grown axial shoot cultures (section 2.2.3) onto petri dishes containing solid callus inducing media. The petri dishes were sealed with whatman tape and incubated at 25 °C in continuous white light. After 3 to 4 weeks the callus which had formed at the cut edges of the leaf squares was removed and transferred to a 250 ml sterile flask containing 50 ml of UM media. Multiple flasks were set up and placed onto an orbital shaker at 140 rpm, 25 °C in continuous white light. Two weeks later one fifth of the cell suspension culture was transferred to a sterile 250 ml flask containing 50 ml of UM media and grown as before. The latter step was repeated at weekly intervals to maintain the cell suspension culture.

2.2.5 Transformation and regeneration of *N.tabacum* cv. Samsun

shoot inducing media: per litre

solid MS30 (section 2.2.4)	
Cefotexime (dissolved in H_2O and filter sterilised)	200 mg
Kanamycin (dissolved in H_2O and filter sterilised)	100 mg
BAP (dissolved in 0.1 M KOH and filter sterilised)	1 mg

Rooting media: per litre

solid MS30 (section 2.2.4)

Cefotexime (dissolved in H_2O and filter sterilised)	200 mg
Kanamycin (dissolved in H_2O and filter sterilised)	50 mg

Agrobacterium-mediated leaf disc transformation was performed essentially as described by Twell, (1992). Liquid cultures of *Agrobacterium tumefaciens* harbouring the desired gene fusion were grown by inoculating 50 ml of LB (section 2.3.1) containing Kanamycin (section 2.3.3) with a single colony and shaking at 200 rpm on an orbital shaker at 28 °C for two days. In a laminar flow hood, leaves from six week old *in vitro* shoot cultures (section 2.2.3) or surface sterilised leaves (section 2.2.2.1) were sliced up into 2 cm by 2 cm squares using a sterile scalpel and forceps and placed into 80 ml of liquid MS30 media (section 2.2.3) to prevent desiccation. When sufficient leaf squares had been prepared 20 ml of the *Agrobacterium* culture was added and mixed evenly, after an incubation of 5 min the leaf squares were removed blotted on sterile whatman number 1 filter paper, to remove excess *Agrobacterium* , and placed abaxial side up on solid MS30 plates (section 2.2.3). The plates were sealed using whatman sealing film and placed in continuous white light at 25 °C to facilitate *Agrobacterium* infection. After 2 days the leaf squares were transferred onto shoot inducing media and placed at 25 °C in continuous white light. Resistant shoots which appeared after 2-3 weeks were excised from the leaf squares and placed individually into sterile pots containing rooting media and grown in continuous white light at 25 °C. Healthy rooted plants were transferred to sterile soil and grown in a propagator for 2 weeks under greenhouse conditions. Established plants were transferred to 4 inch square pots and grown under greenhouse conditions until maturity.

2.2.6 Determination of loci number

Mature seed was collected from transgenic plants and surface sterilised as in section 2.2.2.2. Approximately 50 seeds were evenly plated out onto 9 cm petri dishes containing solid MS30 media (section 2.2.3) and Kanamycin at a concentration of 50 mg/l. Petri dishes were sealed with whatman sealing film and incubated at 25 °C in continuous white light for three weeks. After three weeks seedlings were scored as either resistant (well established roots) or sensitive (small stunted roots). From the ratio of resistant: sensitive seedlings the loci number of each individual transgenic line was estimated.

2.3 Bacterial culture and storage

2.3.1 Media for the growth of bacteria

Luria Bertani (LB) media: per litre

bacto-tryptone	10 g
bacto-yeast extract	5 g
NaCl	10 g

pH adjusted to 7.0 with 1 M NaOH
Autoclaved 15 min at 120 °C

LB media was solidified by the addition of 1.5 % w/v agar and autoclaved as above.

2YT media: per litre

bacto-tryptone	16 g
bacto-yeast extract	16 g
NaCl	5 g

pH adjusted to 7.0 with 1 M NaOH
Autoclaved 15 min at 120 °C

2XL media: per litre

bacto-tryptone	16 g
bacto-yeast extract	10 g
NaCl	5 g
18 % w/v glucose	

pH adjusted to 7.0 with 1 M NaOH
Autoclaved 15 min at 120 °C

2.3.2 Strains and genotypes

2.3.2.1 *Eschericia coli* (*E.coli*)

XL1-Blue: *recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, lac,* {*F', proAB, lacI*^{q}, *ZDM15, Tn10,* (*tet*^{R})}. (Bullock *et al.*, 1987)

2.3.2.2 *Agrobacterium tumefaciens (A.tumefaciens)*

LBA4404 (pAL4404) (Ooms *et al.*, 1982)

2.3.3 Antibiotics for bacterial selection

Name	Concentration (mg/l) for:-	
	E.coli	*A.tumefaciens*
Ampicillin	100	-
Tetracycline	10	-
Kanamycin	50	50
Rifampicin	-	50
Streptomycin	-	200

2.3.4 Long term storage of bacterial cultures

Viable bacterial cultures were maintained for long periods of time (up to 2 years) by transferring 0.5 ml of a dense bacterial culture to a cryogenic storage tube. An equal volume of sterile 50 % v/v glycerol was added to the culture, the contents vortexed and stored at - 70 °C.

2.4 Isolation of plasmid DNA

2.4.1 Large scale isolation of plasmid DNA

Solution I

glucose	50 mM
Tris-HCl pH 8.0	25 mM
EDTA pH 8.0	10 mM

Autoclaved 15 min at 120 °C

Solution II

0.2 M NaOH
1 % w/v SDS

Solution III

Potassium acetate	117.77 g
glacial acetic acid	46 ml
X ml H_2O to total volume of 400 ml	

Autoclaved 15 min at 120 °C

The following method was adapted from Sambrook *et al.*, (1989) and routinely yielded approximately 0.5 mg of plasmid DNA for high copy number plasmids such as pUC19, and 50 µg for medium to low copy number plasmids such as pBIN19. From a 5 ml overnight starter culture 0.4 ml was taken and added to 400 ml of LB (section 2.3.1) with the appropriate antibiotic selection and grown overnight on an orbital shaker at 200 rpm, 37 °C. The overnight culture was decanted into 250 ml polypropylene bottles and the bacteria pelleted by centrifugation at 10,000 *g*, room temp for 10 min. The supernatant was decanted and the bacterial pellet resuspended in 8 ml of solution I. To lyse the bacteria 12 ml of solution II was added and the resulting viscous solution gently inverted 3-4 times. Following the addition of 10 ml of solution III, the contents were thoroughly mixed to allow the precipitation of chromosomal DNA and protein. The white precipitate was removed by centrifugation at 10,000 *g*, room temp for 5 min. To remove all traces of the contaminating white precipitate the supernatant was filtered through myra cloth into a clean 250 ml polypropylene bottle.

An equal volume of isopropanol was added and the contents of the bottle mixed by inversion. The remaining nucleic acid was isolated by centrifugation at 10,000 *g*, 4 °C for 10 min. The supernatant was discarded and the nucleic acid pellet resuspended in 3 ml of SH_2O and transferred to a 50 ml polypropylene bottle, 3 ml of 8 M LiCl was added, the contents vortexed and centrifuged at 10,000 *g*, 4 °C for 10 min to selectively precipitate large RNA species. The supernatant was transferred to a clean 50 ml tube, an equal volume of isopropanol was added and the contents vortexed. The plasmid and remaining RNA species were pelleted by centrifugation at 10,000 *g*, 4 °C for 10 min. The supernatant was removed and discarded, the pellet was resuspended in 500 µl of SH_2O and transferred to a 1.5 ml microfuge tube. Following the addition of 20 µl of RNAse A (20 mg/ml) the contents were mixed by vortex and incubated at 37 °C for 1 hour. 500 µl of 1.6 M NaCl containing 13 % (w/v) PEG 8000 was added to the plasmid DNA solution, mixed and placed on ice for 45 to 60 min to aid the selective precipitation of the plasmid.

The DNA was collected by centrifugation at 14,000 *g*, room temp for 10 min. The supernatant was discarded and the pellet resuspended in 500 µl of SH_2O. 500 µl of phenol/chloroform was added and the contents vortexed and centrifuged at 14,000 *g*, room temp for 5 min. The top aqueous layer was transferred to a 1.5 ml microfuge tube.

The phenol/chloroform step was repeated to remove the final traces of PEG 8000. The top aqueous layer was transferred to a microfuge tube and 100 µl of 10 M ammonium acetate and two volumes of absolute ethanol were added to precipitate the plasmid. The plasmid was recovered by centrifugation at 14,000 *g*, room temp for 5 min. The supernatant was discarded and the pellet washed with 200 µl 70 % v/v ethanol to remove remaining salt and centrifuged at 14,000 *g*, room temp for 1 min. The supernatant was discarded and the pellet dried under vacuum to remove residual ethanol. The pellet was resuspended in an appropriate volume of SH_2O. The quality and quantity of the plasmid was estimated by running 1 µl on a 1% TAE agarose gel with appropriate DNA markers (section 2.6.2).

2.4.2 Midi scale isolation of plasmid DNA

This method was used to obtain ~ 50 µg of clean plasmid DNA from high copy number plasmids such as pUC19, which was a sufficient amount to carry out both detailed restriction enzyme analysis (section 2.6.1) and dideoxy chain termination sequencing (section 2.8.1). A bacterial culture was set up by inoculating 50 ml of LB (section 2.3.1) containing the appropriate antibiotic selection, with either 50 µl of bacterial culture from a positive clone identified by small scale isolation of plasmid DNA (section 2.7.5.1), or a loop full of bacteria from a positive clone identified using colony PCR (section 2.7.5.2). The bacterial culture was grown overnight on an orbital shaker at 200 rpm, 37 °C. The overnight culture was decanted into a 50 ml polypropylene bottle and the bacteria pelleted by centrifugation at 10,000 *g*, room temp for 10 min. The supernatant was decanted and the bacterial pellet resuspended in 2 ml of solution I (section 2.4.1). To lyse the bacteria 3 ml of solution II (section 2.4.1) was added and the resulting viscous solution was gently inverted 3-4 times. Following the addition of 2.5 ml of solution III (section 2.4.1), the contents were thoroughly mixed to allow the precipitation of chromosomal DNA and protein. The white precipitate was removed by centrifugation at 10,000 *g*, room temp for 5 min. To remove all traces of the contaminating white precipitate the supernatant was filtered through myra cloth into a clean 50 ml polypropylene bottles. An equal volume of isopropanol was added and the contents mixed by inversion.

The remaining nucleic acid was isolated by centrifugation at 10,000 *g*, 4 °C for 10 min. The supernatant was discarded and the nucleic acid pellet resuspended in 300 µl of SH_2O and transferred to a 1.5 ml microfuge tube, 300 µl of 8 M LiCl was added, the contents vortexed and centrifuged at 14,000 *g*, room temp for 5 min. The supernatant was transferred to a clean microfuge tube and the remaining nucleic acid precipitated by the addition of 600 µl of isopropanol. The contents were mixed by vortex followed by

centrifugation at 14,000 *g*, room temp for 5 min. The supernatant was discarded and the nucleic acid pellet resuspended in 500 µl of SH_2O. Following the addition of 10 µl of RNAse A (20 mg/ml) the contents were mixed by vortex and incubated at 37 °C. After 1 hour 500 µl of phenol/chloroform was added, vortexed and centrifuged at 14,000 *g*, room temp for 5 min. The top aqueous layer was transferred to a 1.5 ml microfuge tube to which 100 µl of 10 M ammonium acetate and two volumes of absolute ethanol were added to precipitate the plasmid. The plasmid was recovered by centrifugation at 14,000 *g*, room temp for 5 min. The supernatant was discarded and the pellet washed with 200 µl 70 % v/v ethanol to remove remaining salt and centrifuged at 14,000 *g*, room temp for 1 min. The supernatant was discarded and the pellet dried under vacuum to remove residual ethanol. The pellet was resuspended in an appropriate volume of SH_2O. The quality and quantity of the plasmid was estimated by running 1 µl on a 1% TAE agarose gel with appropriate DNA markers (section 2.6.2).

2.5 Amplification of DNA by the polymerase chain reaction

2.5.1 Purification of oligonucleotides

Oligonucleotide primers were synthesised using either an ABI 3808 or ABI 394 DNA synthesizer by PNACL sevices Leicester University. Due to the method of synthesis the oligonucleotide primers were stored in a solution containing ammonia. Oligonucleotide purification was carried out according to Sawadogo and Van Dyke, (1991). A 100 µl aliquot containing impure oligonucleotide was transferred to a 1.5 ml microfuge tube, 1 ml of butan-1-ol was added, the contents vortexed for 15 sec and centrifuged at 14,000 *g*, room temp for 1 min. The supernatant was discarded and the oligonucleotide pellet resuspended in 100 µl of SH_2O, a further 1 ml of butan-1-ol was added, the contents vortexed for 15 sec and centrifuged at 14,000 *g*, room temp for 1 min. The supernatant was discarded and the precipitated oligonucleotide was vacuum dried for approximately 30 min to remove the remaining butan-1-ol. After resuspension in 100 µl of SH_2O the oligonucleotide concentration was calculated by diluting 1 µl in 200 µl of SH_2O. The O.D. of the dilution was measured over the range of 200-300 nm using a Hewlett Packard 8452A diode array spectrophotometer. The oligonucleotide concentration was calculated using the following formula:

$$O.D._{260} \times 20 \times \text{dilution} = \text{mg/ml}$$

2.5.2 Amplification of plasmid DNA

The polymerase chain reaction (PCR) was utilised to amplify specific fragments of the *lat*52 promoter using tagged oligonucleotide primers. Primers were designed to introduce specific restriction enzyme sites at the 5' and 3' ends of the amplified DNA fragments to facilitate cloning.

A typical PCR reaction mixture was prepared as follows:

10 x Taq DNA polymerase buffer	10 µl
2 mM dNTP's	10 µl
plasmid DNA template	5 ng
5' primer	100 ng
3' primer	100 ng
Taq DNA polymerase	0.4 u
SH_2O to 100 µl	

The above PCR reaction mixture was assembled in a 0.5 ml microfuge tube and over-layed with 2 drops of paraffin oil to prevent evaporation. The tube was placed in a Perkin-Elmer DNA thermal cycler. The DNA fragment was amplified for 30 cycles.

Each cycle consisted of the following steps:

1 min at 95 °C DNA template denatured
1 min at 50 °C Primer annealing
1 min at 72 °C Synthesis of DNA fragment

For the synthesis of fragments greater than 0.5 kb in length, the synthesis time at 72 °C was increased by 1 min increments for each additional 0.5 kb to be amplified. When primers containing mismatches were used, 2 cycles containing a lower primer annealing temperature were carried out prior to using the above PCR conditions.

2.5.3 Purification of amplified DNA fragments

The quantity of amplified DNA was verified by subjecting 10 µl of the PCR reaction to agarose gel electrophoresis (section 2.6.2). An equal volume of phenol/chloroform was added to the remainder of the PCR reaction, the contents mixed by vortex and centrifuged at 14,000 *g*, room temp for 5 min. The top aqueous layer was transferred to a 1.5 ml microfuge tube and the amplified DNA fragment precipitated by the addition of 10 µl of 3 M sodium acetate pH 5.2 and 220 µl of absolute ethanol. The contents were

mixed by vortex and centrifuged at 14,000 *g*, room temp for 20 min. The supernatant was discarded and the DNA pellet washed with 100 µl of 70 % v/v ethanol to remove remaining salt and centrifuged at 14,000 *g*, room temp for 1 min. The supernatant was discarded and the DNA pellet air dried at room temp, followed by resuspension in 30 µl of SH_2O.

2.6 Manipulation of DNA for cloning of fragments into plasmid vectors

2.6.1 Digestion of DNA with restriction enzymes

DNA was digested using restriction enzymes to generate fragments with compatible ends for cloning, or to verify newly constructed plasmids. The reaction condition for each enzyme(s) was followed as per manufacturers instructions. Digests were normally set up in a total volume of 30 µl which consisted of:

DNA	x µl
Restriction enzyme * (10 units/µl)	1 µl
10 x React buffer	3 µl
X µl SH_2O to 30 µl	

Digests were normally carried out at 37 °C for 3-4 hours unless otherwise stated.

*When more than two enzymes were used together the total reaction volume was increased proportionally to prevent the glycerol concentration in the reaction exceeding 5 %, which with certain restriction enzymes would lead to random cleavage of the DNA.

2.6.2 Agarose gel electrophoresis

Gel loading buffer (10 x)

0.5 % w/v orange G
50% v/v glycerol

10 x TAE buffer: per litre

Trisma base	48.4 g
sodium acetate trihydrate	37.2 g

EDTA 3.72 g

pH adjusted to 8.2 with glacial acetic acid

Neutral agarose gels were utilised to quantify and asses the quality of newly isolated plasmid DNA, DNA fragments generated using PCR and also to verify the products from restriction enzyme analysis of the aforementioned DNA. Agarose was melted in 1 x TAE buffer using a microwave, cooled to ~ 60 °C and ethidium bromide was added to a final concentration of 125 ng/ml. The liquid agarose was poured into a perspex tray with an appropriate comb and allowed to set at room temp. When set the agarose gel including the tray and comb was submerged in an electrophoresis tank containing 1 x TAE buffer. The comb was removed and the wells washed out by agitation.

Gel loading buffer (10 x) was added to DNA samples to a final concentration of 1 x. DNA samples were loaded into individual wells and electrophoresed at 8 volts/cm. Quantification of the amount (ng) and size (Kb) of DNA fragments was determined by comparison of their mobility with the reference DNA molecular weight standard 1 Kb ladder. Generally plasmid DNA and DNA fragments > 0.5 Kb were electrophoresed in 1 % w/v volume agarose gels. The concentration of agarose was increased to 2 % w/v volume for DNA fragments in the size range of 0.2-0.5 Kb and to 4 % w/v volume for DNA fragments < 0.2 Kb.

Following electrophoresis gels were photographed on a transilluminator using a camera.

2.6.3 Purification of DNA fragments from agarose gel slices

DNA fragments in the size range of 0.4-15 Kb were purified from agarose gel slices using a Geneclean DNA purification kit (Bio 101 Ltd.). The kit was used as per manufacturers instructions. Two elution steps were performed and combined. The quantity of isolated DNA was verified by subjecting 2 µl of the eluted DNA to agarose gel electrophoresis (section 2.6.2).

DNA fragments in the size range of 0.05-0.4 Kb were purified from agarose gel slices using a Mermaid DNA purification kit (Bio 101 Ltd.). The kit was used as per manufacturers instructions. Two elution steps were performed and combined. The quantity of isolated DNA was verified by subjecting 2 µl of the eluted DNA to agarose gel electrophoresis (section 2.6.2).

2.6.4 Removal of 5' terminal phosphate groups from cleaved plasmid DNA

10 x Phosphatase buffer

$ZnCl_2$	10 mM
$MgCl_2$	10 mM
Tris-HCl (pH 8.3)	100 mM

In order to reduce the efficiency with which plasmid DNA cleaved by a single restriction enzyme re-ligated minus the required insert, the 5' terminal phosphate groups were removed using calf intestinal alkaline phosphatase (CIP). This was achieved using the following method which is suitable for restriction enzymes which cleave DNA to leave a 5' overhang. About 5-6 µg of plasmid DNA was digested at 37 °C overnight with the required restriction enzyme e.g.

Plasmid DNA (6 µg)	6 µl
10 x react buffer 2	6 µl
XhoI (20 units)	2 µl
SH_2O	46 µl

Complete digestion of the plasmid was confirmed by subjecting 10 µl of the digest to agarose gel electrophoresis (section 2.6.2).

To the remaining 50 µl digest the following were added:

CIP (0.5 unit)	1 µl
10 x Phosphatase buffer	5 µl
SH_2O	44 µl

The reaction was incubated at 37 °C for 30 min in a water bath. To inactivate CIP and the restriction enzyme an equal volume of phenol/chloroform was added to the reaction, the contents were mixed by vortex and centrifuged at 14,000 *g*, room temp for 5 min. The top aqueous layer was transferred to a 1.5 ml microfuge tube and the DNA was precipitated by the addition of 10 µl of 3 M sodium acetate pH 5.2 and 220 µl of absolute ethanol. The contents were mixed by vortex and centrifuged at 14,000 *g*, room temp for 20 min. The supernatant was discarded and the DNA pellet washed with 100 µl of 70 % v/v ethanol to remove remaining salt and centrifuged at 14,000 *g*, room temp for 1 min.

The supernatant was discarded and the DNA pellet air dried at room temp, followed by resuspension in 30 µl of SH_2O.

2.6.5 Addition of 5' terminal phosphate groups to oligonucleotides

Oligonucleotides are synthesised without a 5' terminal phosphate groups. To prepare oligonucleotides for cloning a 5' terminal phosphate group was added using the following method. Oligonucleotides were purified and their concentration determined as in section 2.5.1.

The following reaction was set up in a total volume of 30 µl and incubated at 37 °C for 60 min.

Tris-HCl (pH 7.5)	50 mM
$MgCl_2$	10 mM
DTT	5 mM
oligonucleotide	1-10 µg
ATP	1 mM
T4 polynucleotide Kinase	20 u
BSA	50 ng/µl

The T4 polynucleotide kinase was inactivated by the addition of 70 µl of SH_2O and 100 µl of phenol/chloroform, mixed by vortex and centrifuged at 14,000 *g*, room temp for 5 min. The top aqueous layer was transferred to a 1.5 ml microfuge tube and the oligonucleotide precipitated by the addition of 10 µl of 3 M sodium acetate pH 5.2 and 220 µl of absolute ethanol. The contents were mixed by vortex and centrifuged at 14,000 *g*, room temp for 20 min. The supernatant was discarded and the oligonucleotide pellet washed with 100 µl of 70 % v/v ethanol to remove remaining salt and centrifuged at 14,000 *g*, room temp for 1 min. The supernatant was discarded and the oligonucleotide pellet air dried at room temp.

2.6.6 Ligation of DNA fragments

Recombinant plasmids were created by annealing cut fragments together using T4 DNA ligase. For ligations involving vectors which contained an origin of replication for high copy number such as pBluescript KS(+), pUC19 or pSL301 25 ng of cut vector plasmid was used, for ligations using low copy number plasmids such as pBIN19 100 ng of cut vector plasmid was used. The quantity of insert for each ligation was worked out using a vector:insert ratio of 1:3 with respect to ends and the following equation:

$$\frac{\text{INSERT SIZE}}{\text{VECTOR SIZE}} \times \frac{3}{1} \times \text{ng OF CUT VECTOR} = \text{ng OF INSERT}$$

Standard ligations were set up as follows and incubated at 12 °C overnight:

Vector	x µl
Insert	x µl
5 x T4 DNA ligase buffer	2 µl
T4 DNA ligase (5 units)	1 µl
SH_2O to 10 µl	

Vector only control

Vector	x µl
5 x T4 DNA ligase buffer	2 µl
T4 DNA ligase (5 units)	1 µl
SH_2O to 10 µl	

2.7 Bacterial transformation

2.7.1 Preparation of competent *E.coli*

$Ca^{2+}Mn^{2+}$ solution:

sodium acetate	40 mM
$CaCl_2$	100 mM
$MnCl_2$	70 mM

pH adjusted to 5.5 with 1 M HCl
Filter sterilised

Competent *E.coli* for the cloning of recombinant plasmids were prepared using the following method (Hanahan, 1983). A single colony of XL1-blue (section 2.3.2.1) was inoculated into 25 ml of 2XL broth (section 2.3.1) containing teytracycline (section 2.3.3) and grown at 37 °C, 200 rpm overnight. 1 ml of the overnight bacterial culture was added to 100 ml of 2XL broth pre-warmed to 37 °C, and incubated at 37 °C, 200 rpm on an orbital shaker. When the $O.D._{600}$ of the bacterial culture had reached 0.2,

sterile 1 M $MgCl_2$ was added to a final concentration of 20 mM. The bacterial culture was grown as before. When the bacterial culture had obtained an $O.D._{600}$ of 0.45-0.55, it was poured into 50 ml sterile tubes and placed on ice for 2 hours. Bacteria were pelleted by centrifugation at 3,000 *g*, 4 °C for 5 min. The supernatant was discarded and each bacterial pellet gently resuspended in 25 ml of $Ca^{2+}Mn^{2+}$ solution (prechilled to 4 °C) using a sterile disposable plastic pasteur pipette. The bacteria were incubated on ice for 45 min to allow the chemicals to weaken the bacterial cell walls. The now competent bacteria were harvested as before, the supernatant discarded and the two pellets combined by gentle resuspension in a total volume of 5 ml of $Ca^{2+}Mn^{2+}$ solution containing 15 % v/v glycerol. The cells were quickly aliquoted (0.2 ml) into pre-chilled 1.5 ml microfuge tubes, flash frozen in liquid nitrogen and stored at -70 °C where they remained viable for 1-2 months.

2.7.2 Transformation of *E.coli* with ligated plasmid DNA

To ascertain if the competent *E.coli* cells prepared in section 2.7.1, were of a sufficient quality for the cloning of recombinant plasmids the transformation efficiency of these cells was calculated by transformation of a known amount of supercoiled plasmid DNA. To confirm that the cells had not been contaminated with plasmid DNA during their preparation a no DNA control was utilised. The competent cells were transfected with plasmid DNA using the following method (Hanahan, 1983).

Two 200 µl aliquots of competent cells were thawed on ice. To one aliquot (DNA control) 10 pg of supercoiled plasmid (e.g pUC19) in a volume of 25 µl SH_2O was added and mixed with the cells by gently flicking the tube. To the second aliquot (no DNA control) 25 µl of SH_2O was added and mixed with the cells by gently flicking the tube. The cells were incubated on ice for 30 min then heat shocked at 37 °C for 5 min. The cells were transferred into a 10 ml tube containing 925 µl of LB (section 2.3.1) and grown at 37 °C, 200 rpm in an orbital shaker for 1 hour to allow the cells to acquire resistance to antibiotic selection. Aliquots were plated out onto LB agar (section 2.3.1) plates containing antibiotic selection (section 2.3.3), the plates were dried and placed upside down in a dry incubator at 37 °C overnight.

The transformation efficiency (TE) was calculated in the following way:

$$\text{Number colonies} \times \text{dilution} \times \frac{1\ \mu g}{\text{amount DNA used}}$$

If 100 µl of the above transformation yielded 10 colonies then the TE would equal:

$10 \times 10 \times 10^5 = 1 \times 10^7$

If the competent cells had a TE of 1×10^6 to 1×10^7, and if no bacterial colonies were detected with the no DNA control they were deemed suitable for use.

Plasmid ligations prepared in section 2.6.7 were transfected in to competent cells essentially as described above. Except 2 µl of each ligation reaction was mixed with 23 µl of SH_2O, chilled on ice, then 50 µl of thawed competent cells were added.

2.7.3 Preparation and transformation of competent *A.tumefaciens*

$T_{10}E_1$ (pH 7.5)

Tris-HCl pH 7.5	10 mM
EDTA pH 8.0	1 mM

Recombinant plasmids based on the binary vector pBIN19 were introduced into *A.tumefaciens* by the method described by Hofgen and Willmitzer, (1988). A sterile universal containing 10 ml of 2YT (section 2.3.1) was inoculated with a single colony of *A.tumefaciens* (section 2.3.2.2) from a freshly streaked plate and grown overnight at 28 °C with shaking (200 rpm). 200 ml of 2YT was inoculated with the 10 ml overnight bacterial culture and grown for a further 4 hours as before. Bacteria were harvested by centrifugation at 3000 *g*, 4 °C for 10 min. The supernatant was discarded and the bacterial pellet washed with 20 ml of ice cold $T_{10}E_1$ (pH 7.5) and centrifuged as before. The supernatant was discarded and the pellet resuspended in 20 ml of ice cold 2YT, 0.5 ml aliquots in 1.5 ml microfuge tubes were taken, flash frozen in liquid nitrogen and stored at -70 °C.

Competent *A.tumefaciens* were transfected with plasmid DNA as follows. An aliquot of cells was thawed on ice, 1 µg of plasmid DNA in 10 µl of SH_2O was added and mixed by flicking the tube. This was incubated on ice for 5 min, in liquid nitrogen for a further 5 min and then thawed at 37 °C. The cells were transferred to a 10 ml tube containing 1 ml of 2YT and grown at 28 °C with shaking (200 rpm) for 2-4 hours. 300 µl aliquots were plated out on LB agar (section 2.3.1) plates containing the appropriate antibiotic selection (2.3.3), air dried, inverted and placed at 28 °C in a dry incubator. Single bacterial colonies could be picked for further analysis after 2 days.

2.7.4 Identification of positive recombinant plasmids

2.7.4.1 Restriction enzyme analysis of plasmid DNA

Positive recombinant plasmids generated from section 2.7.2 were identified by restriction enzyme analysis of small quantities of crude plasmid isolated using the following method. 5 ml of LB (section 2.3.1) plus antibiotic selection (section 2.3.3) in a sterile glass universal was inoculated with a single bacterial colony and grown overnight on an orbital shaker at 200 rpm, 37 °C. 1.5 ml of the bacterial culture was pelleted in a 1.5 ml microfuge tube by centrifugation at 14,000 *g*, room temp for 5 min. The supernatant was discarded and the bacterial pellet resuspended in 100 µl of solution I (section 2.4.1), 150 µl of solution II (section 2.4.1) was added to lyse the bacteria with the subsequent addition of 125 µl of solution III (section 2.4.1) to precipitate chromosomal DNA and protein. To remove the white precipitate the tube was centrifuged at 14,000 *g*, room temp for 5 min. The supernatant was transferred to a clean microfuge tube leaving all of the white precipitate behind. For *endA1* bacterial strains such as XL1-blue (section 2.3.2.1), treatment of the supernatant with phenol/chloroform was unnecessary. The crude plasmid was precipitated by the addition of 2.5 volumes of absolute ethanol, mixed by vortex, and centrifugation at 14,000 *g*, room temp for 5 min. The supernatant was discarded and the pellet washed with 100 µl of 70% v/v ethanol and centrifuged at 14,000 *g*, room temp for 1 min. The supernatant was discarded and the pellet air dried for about 30 min to remove residual traces of ethanol. Once dry the pellet was resuspended in 25 µl of SH_2O. A typical digest was then performed on the crude plasmid in the following way:

plasmid	5 µl
RNAse A (20 mg/ml)	1 µl
10 x react buffer	3 µl
restriction enzyme	1 µl
SH_2O to 30 µl	
37 °C for 2-3 hours.	

The products of the restriction enzyme digest were visualised on a TAE agarose gel with appropriate DNA markers (section 2.6.3).

2.7.4.2 Colony PCR

Positive recombinant clones were normally identified by restriction enzyme analysis of small quantities of plasmid DNA isolated using the method described in section 2.7.4.1. This method proved inappropriate when large numbers of putative recombinant

clones had to be screened in order to identify a low frequency cloning event. Also using this method insufficient plasmid could be isolated to verify the introduction of small DNA fragments into receptor plasmids and for the verification of the introduction of recombinant clones into *Agrobacterium tumefaciens* (section 2.7.3). To overcome these problems Colony PCR was utilised. A 50 µl aliquot of a PCR mixture (section 2.5.2) in a 0.5 ml microfuge tube was inoculated with a single bacterial colony using a sterile toothpick. Bacteria still present on the toothpick were streaked out onto LB agar (section 2.3.1) plates with antibiotic selection (section 2.3.3). The PCR reaction was then performed as before (section 2.5.2) but with an initial step of 10 min at 95 °C to burst the bacteria. PCR products were visualised on a TAE agarose gel with appropriate DNA markers (section 2.6.3).

2.8 DNA sequencing

2.8.1 Dideoxy-mediated chain termination using T7 DNA polymerase

2.8.1.1 Denaturation of double stranded plasmid DNA for sequencing

Plasmid DNA suitable for sequencing was isolated using the method described in section 2.4.2. In a sterile 1.5 ml microfuge tube approximately 10 µg of plasmid DNA in a volume of 16 µl was denatured by the addition of 4 µl of 1 M NaOH. The contents were mixed by vortex and incubated at room temp for 5 min. The NaOH was neutralised by the addition of 6.7 µl of 3 M ammonium acetate pH 4.8 and mixed by vortex. The denatured DNA was precipitated by the addition of 66.7 µl of absolute ethanol, mixed by vortex and incubated at -70 °C for 10 min, followed by centrifugation at 14,000 *g*, room temp for 20 min. The supernatant was discarded and the pellet washed with 100 µl of 70 % v/v ethanol and centrifuged at 14,000 *g*, room temp for 1 min. The supernatant was discarded and the pellet air dried for about 30 min, room temp to remove residual traces of ethanol.

2.8.1.2 DNA sequencing reactions

DNA sequencing reactions were prepared using the reagents and methods supplied with the pharmacia T7 sequencing kit. The denatured DNA pellet prepared in section 2.8.1.1 was resuspended in 20 µl of SH_2O. Primer was annealed to the denatured DNA template by the addition of 2 µl (20 ng) of primer oligonucleotide and 2 µl of annealing buffer to 10 µl of the resuspended denatured DNA in a clean 1.5 ml microfuge tube. The contents were mixed by vortex, centrifuged briefly and incubated at 65 °C for 3 min, 37

°C for 20 min and room temp for 10 min.

Before the sequencing reaction was carried out an enzyme premix was prepared on ice which consisted of the following:

SH_2O	1.5 µl
Labelling mix A	3.0 µl
T7 DNA polymerase (1.5 units/µl)	2.0 µl
Labelled dNTP (^{35}S α-dATP)	0.5 µl

To the primer annealed template 6 µl of the enzyme premix was added, mixed gently and incubated at room temp for 5 min. 4.5 µl of the reaction mix was transferred onto the side of 4 microfuge tubes labled A, C, G or T, each containing 2.5 µl of the appropriate dideoxynucleotide solution. The contents were mixed by brief centrifugation and incubated at 37 °C for 5 min. The reactions were terminated by the addition of 5 µl of stop solution. Sequencing reactions were stored at -20 °C until required.

2.8.1.3 Sequencing gels and autoradiography

5 x TBE buffer: per litre:

H_3BO_3	27.5 g
Trisma base	54 g
0.5 M EDTA pH 8.0	20 ml
pH 8.3	

Sequencing reactions prepared in section 2.8.1.2 were analyzed by polyacrylamide gel electrophoresis using a Sequi-Gen nucleic acid sequencing cell (BIO-RAD). Prior to assembly of the gel apparatus all components were cleaned with detergent and warm water. All traces of detergent were removed and the two glass plates wiped with absolute ethanol. To prevent the acrylamide gel from sticking to both glass plates 2-3 ml of sigma coat was evenly applied to the front glass plate and allowed to dry in a fume hood. The apparatus was assembled according to the manufactures instructions.

A 6 % polyacrylamide gel mix was prepared as follows:

46 % w/v Urea	37.50 ml
40 % Acrylamide/Bis-Acrylamide	11.25 ml

5 x TBE	15.00 ml
SH_2O	11.25 ml

The bottom space between the 2 glass plates was sealed by taking 15 ml of the above gel mix and initiating polymerisation by the addition of 50 µl of TEMED and 50 µl of ammonium persulphate (250 mg/ml). This was poured onto a piece of 3 MM whatman filter paper in a casting tray, the sequencing apparatus was placed onto it and left in a upright position for 15 min to allow the acrylamide to polymerise.

The remaining gel mix was filtered through a 0.45 µm filter. Polymerisation was initiated by the addition of 50 µl of TEMED and 50 µl of ammonium persulphate (250 mg/ml) to the filtered gel mix. A 50 ml syringe was filled with the gel mix and a wide bore needle attached. The mix was poured at the corner between the two glass plates, while holding the apparatus at an angle of about 45°. When the gel mix had reached the top of the back plate the apparatus was lowered to an angle of 30°. The sharks tooth comb was inserted upside down between the two plates so that the bottom of the holes in the comb were in line with the top of the back plate. The comb was clamped in place using a bulldog clip and the gel left to polymerise for 2 hours.

Once the sequencing gel had polymerised the apparatus was assembled. Approximately 400 ml of 1 x TBE buffer was placed in the bottom reservoir and 600 ml of 1 x TBE buffer was poured into the back plate reservoir. The comb was removed and the wells washed out with 1 x TBE buffer using a syringe. The gel was run at 65 watts, 1,500 volts until the temperature reached 55 °C, the comb was re-inserted so that the shark teeth penetrated the gel by about 1-2 mm. Each sequencing reaction was boiled at 95 °C for 2 min and then placed back on ice. The wells were washed out as before and 2.5 µl of each sample was loaded between the teeth in the order A, C, G and T. The gel was run at a constant temperature of 50 °C (~ 50 watts, 1,500 volts). For 'short gels' the bromophenol blue dye front was allowed to run off, for 'long gels' both dye fronts were run off before loading the gel for a short run.

Following electrophoresis the apparatus was disassemble and the gel transferred onto a piece of 3 MM whatman filter paper. A layer of 'clingfilm' was placed on top of the gel/3 MM whatman filter paper and a further piece of 3 MM whatman filter paper placed underneath. The gel was dried under vacuum at a temperature of 80 °C for 2 hours. The dried gel was placed in contact with X-ray film and left at room temp. After a suitable exposure time the X-ray film was developed.

2.8.2 Dideoxy-mediated chain termination using the automated ABI Prism™ procedure

2.8.2.1 Isolation of plasmid DNA for sequencing

Plasmid DNA used in automated sequencing reactions was isolated using the Wizard® *plus SV* miniprep DNA purification system (Promega) according to manufacturers instructions. The quantity of isolated plasmid DNA was verified by subjecting 5 µl of the eluted DNA to agarose gel electrophoresis (section 2.6.2).

2.8.2.2 Automated sequencing reactions

Automated sequencing reactions were prepared in a 0.5 ml microfuge tube as follows:

Terminator ready reaction mix	8.0 µl
double stranded DNA (0.2 µg/µl)	2.0 µl
primer (3.2 pmol)	1.0 µl
SH_2O	9.0 µl

Each sequencing reaction was over-layed with a single drop of paraffin oil and placed into a Perkin-Elmer DNA thermal cycler model 480. Extension products were amplified over 25 cycles using the following conditions:

Rapid thermal ramp to 96 °C
96 °C for 30 sec
Rapid thermal ramp to 50 °C
50 °C for 15 sec
Rapid thermal ramp to 60 °C
60 °C for 4 min

Extension products were purified by transferring the 20 µl sequencing reaction to a 1.5 ml microfuge tube containing 2.0 µl of 3 M sodium acetate (pH 5.2) and 50 µl of 95 % v/v ethanol. The contents were mixed by vortex and placed on ice for 10 min. The extension products were precipitated by centrifugation at 14,000 *g*, room temp for 20 min. The supernatant was discarded and the pellet washed with 250 µl of 70 % v/v ethanol. The 70 % v/v ethanol was removed and the pellet air dried. Purified extension products were analyzed on a 24 cm well-to-read gel (ABI 373) by PNACL services (Leicester University).

2.8.3 Storage and analysis of DNA sequence data

Sequence information was stored and manipulated using the computer program Gene Jockey designed for DNA sequence analysis on the apple macintosh computer.

2.9 Microprojectile bombardment

2.9.1 Equalisation of test plasmid DNA

Plasmid DNA for microprojectile bombardment was isolated as described in section 2.4.1. For the accurate comparison of test plasmid activities, test plasmids had to be the same concentration. Since no two plasmid isolations yield exactly the same quantity of DNA and plasmid isolated using this method was contaminated with trace amounts of chromosomal DNA, RNA and other nucleotides. A method other than calculating plasmid DNA concentration from spectrophotometric analysis was used. The predominant structural form of plasmid isolated was supercoiled, this allowed test plasmids to be equalised to each other using agarose gel electrophoresis (section 2.6.3). 5 µl of test plasmid (stock) was diluted in 495 µl of SH_2O. A 10 µl aliquot was electrophoresed against 1 µg of 1 Kb ladder on a 1 % agarose gel. When 1 µg of 1 Kb ladder is used the 1.6 Kb band represents 100 ng of DNA. The relative intensity of each diluted test plasmid was compared to that of the 1.6 Kb band, if the intensity of the diluted test plasmid equals that of the 1.6 Kb band then the stock test plasmid is at a concentration of 1 µg/µl. If the intensity of the diluted test plasmid was greater than the 1.6 Kb band then the stock test plasmid was diluted accordingly, and a new plasmid dilution prepared and electrophoresed against 1 µg of 1 Kb ladder on a 1 % agarose gel. This was repeated until the concentration of each test plasmid stock was 1 µg/µl.

2.9.2 Preparation of M10 tungsten microprojectiles

M10 tungsten microprojectiles were resuspended in absolute ethanol at a concentration of 50 mg/ml. 0.5 ml aliquots were transferred to 1.5 ml microfuge tubes and centrifuged at 14,000 *g*, room temp for 5 min. The supernatant was discarded, the microprojectiles were washed with 0.5 ml of SH_2O, the contents mixed by vortex and centrifuged at 14,000 *g*, room temp for 5 min. The wash step was repeated a further 3 times. The M10 tungsten microprojectiles were resuspended in 0.5 ml of SH_2O and stored at -20 °C until required.

2.9.3 Precipitation of plasmid DNA onto M10 tungsten microprojectiles

Plasmid DNA was precipitated onto M10 tungsten microprojectiles by mixing 7 µl of

test plasmid (1 µg/µl) with 3 µl of reference plasmid (1 µg/µl) in a 1.5 ml microfuge tube, to this 25 µl of M10 tungsten microprojectiles (section 2.9.2) was added and mixed by rapid pipetting. 25 µl of 1 M $CaCl_2$ was mixed in followed by 10 µl of 0.1 M spermidine free base. This was left for 10 min after which 25 µl of the supernatant was removed and discarded.

2.9.4 Microprojectile bombardment

The DNA coated tungsten particles prepared in section 2.9.3 were dispersed by briefly touching the bottom of the 1.5 ml microfuge tube to a horn type sonicator, 2 µl was quickly pipetted onto the surface of a macroprojectile. This was inverted and rammed down the barrel of a biolistic gun, a blank .22 charge was placed directly above the macroprojectile and the firing apparatus assembled. A stopping plate was placed in the chamber approximately 15 cm above the plant tissue and the chamber evacuated. At a vacuum of 25 inches of mercury the firing mechanism was activated. The vacuum was gently released and the spent stopping plate and cartridge discarded.

2.9.5 Microprojectile bombardment of mature pollen

Pollen germination media (PGM) (Tupy *et al.*, 1991)

Sucrose	0.3 M
H_3BO_3	1.6 mM
$Ca(NO_3)_2$	3 mM
$MgSO_4$	0.8 mM
KNO_3	1 mM
MES	25 mM

pH adjusted to 5.9 with 1 M KOH
Autoclaved 15 min at 120 °C

PGM media was solidified by the addition of 0.8 % w/v agar and autoclaved as above.

On a 9 cm petri dish containing solid PGM a 7 cm whatman number 1 filter paper was placed. A 3 x 3 cm nylon membrane was over-layed onto the filter paper and allowed to wet. Dehisced pollen isolated in section 2.2.1 was resuspended in liquid PGM at a concentration of 50 mg/ml and a 400 µl aliquot was pipetted onto the nylon membrane. After 5 min excess liquid PGM had soaked into the solid PGM to leave a wet layer of

pollen on the nylon membrane which was ready for bombardment. Pollen was bombarded as in section 2.9.4 with DNA coated tungsten particles prepared as in section 2.9.3. The bombarded pollen was incubated at 25 °C in continuous white light for 16 hours prior to protein extraction.

2.9.6 Microprojectile bombardment of *in vitro* grown *N.tabacum* leaves

In vitro grown *N.tabacum* leaves (section 2.2.3)were placed axial side up on a 9 cm petri dish containing solid MS30 media (section 2.2.3). Leaves were bombarded as in section 2.9.4 with DNA coated tungsten particles prepared as in section 2.9.3. The bombarded leaves were incubated at 25 °C in continuous white light for 24 hours prior to protein extraction.

2.9.7 Microprojectile bombardment of *N.tabacum* cell suspension culture

A 1.5 ml aliquot of a 7 day old *N.tabacum* cell suspension culture (section 2.2.4) was partially dried down onto a 5 cm nitrocellulose membrane using a bucner flask. The nitrocellulose membrane with the cell suspension culture was placed onto a 9 cm petri dish containing solid MS30 media (section 2.2.3). The cell suspension culture was bombarded as in section 2.9.4 with DNA coated tungsten particles prepared as in section 2.9.3. The bombarded cell suspension culture was incubated at 25 °C in continuous white light for 24 hours prior to protein extraction.

2.10 Enzyme assays of plant extracts

2.10.1 Protein extraction of bombarded plant material

Luciferase extraction buffer (LEB):

0.1 M KPO_4 pH 7.5	9.9 ml
0.1 M DTT	0.1 ml

Bombarded pollen (section 2.9.5) was scraped off the nylon membrane using a clean microscope slide and transferred to a clean mortar and pestle. 500 µl of LEB was added and the pollen ground for 1 min. The ground pollen was poured into a 1.5 ml microfuge tube and placed on ice.

Bombarded leaf material (section 2.9.6) was transferred using forceps to a clean mortar and pestle. 500 µl of LEB was added and the leaf ground for 1 min. The ground

leaf was poured into a 1.5 ml microfuge tube and placed on ice.

Bombarded cell suspension culture (section 2.9.5) was scraped off the nitrocellulose membrane using a clean microscope slide and transferred to a clean mortar and pestle. 1 ml of LEB was added and the cell suspension culture ground for 1 min. The ground cell suspension culture was poured into a 1.5 ml microfuge tube and placed on ice.

When all tissue samples from one bombardment experiment had been ground, samples were centrifuged at 14,000 *g*, room temp for 5 min to pellet cell debris prior to enzyme assays.

2.10.2 Luciferase (LUC) assay

ATP buffer:

20 mM ATP (10 mM KPO_4 pH 7.5)	5 ml
0.5 M Hepes pH 7.5	1 ml
1 M $MgCl_2$	0.2 ml
SH_2O	3.8 ml

LUC activity in extracts of bombarded plant material (section 2.10.1) was measured using a Berthold Clinilumat Luminometer. 25 µl of the supernatant was aliquoted into a clean tube, background LUC activity was determined prior to the injection of 200 µl of ATP buffer followed by 200 µl of 0.05 mM d-luciferin, total light units were counted for 10 sec.

2.10.3 ß-Glucuronidase (GUS) assay

2.10.3.1 Fluorometric determination of GUS activity

Gus Extraction Buffer (GEB):

0.1 M $NaPO_4$ pH 7.0	250 ml
0.5 M EDTA pH 8.0	10 ml
10 % v/v Triton	5 ml
SH_2O	235 ml
ß-Mercaptoethanol	35 µl/50 ml

GUS activity in bombarded tissue extracts (section 2.10.1) was measured essentially as described by Jefferson *et al.*, (1987) with some modifications. A 100 µl aliquot of the

supernatant was added to 400 µl of 1 mM MUG (4-methylumbelliferyl-glucuronide) in GEB pre-warmed at 37 °C for 5 min. The reaction was incubated at 37 °C. After 10 min 100 µl was removed from the reaction and directly mixed with 100 µl of 0.2 M Na_2CO_3 in a microtitre plate to stop the reaction. Three consecutive 100 µl aliquots at set time points were removed from the reaction and added directly to 100 µl aliquots of 0.2 M Na_2CO_3 in a microtitre plate. Fluorescence of each aliquot was measured using a Perkin-Elmer luminescence spectrometer LS50. The excitation and emission wavelengths were set at 365 nm and 455 nm respectively each with a slit width of 10 nm. GUS activity was calculated as the change in fluorometric units per hour.

2.10.3.2 Histochemical localization of GUS activity

Histochemical localization of GUS activity was performed according to Twell, (1992) by incubating plant tissues in GEB (section 10.3.1) containing 1 mM X-Gluc and 0.5 mM potassium ferricyanide at room temp for 16 hours.

2.10.4 Calculation of test plasmid activity

The relative activity of a test plasmid containing the reporter gene luciferase assayed using microprojectile bombardment (section 2.9) was calculated from a ratio of LUC:GUS i.e.

$$\frac{\text{light units/ 10 sec}}{\text{fluorometric units/hour}}$$

normalised to a control test plasmid.

The relative activity of a test plasmid containing the reporter gene ß-Glucuronidase assayed using microprojectile bombardment (section 2.9) was calculated from a ratio of GUS:LUC i.e.

$$\frac{\text{fluorometric units/hour}}{\text{light units/ 10 sec}}$$

normalised to a control test plasmid.

2.10.5 Calculation of reporter gene activity in transgenic plants

Transgenic plant material assayed for GUS activity was ground in the presence of

GEB (section 2.10.3.1). GUS activity was determined as in section 2.10.3.1. A calibration curve was prepared using 4-methyl umbelliferone (the reaction product 4-MU) dissolved in 0.2 M Na_2CO_3 as a standard, such that the amount in pmol of 4-MU formed per hour could be determined from the change in fluorometric units per hour. Protein concentration was determined using Bradford reagent and a calibration curve prepared using known quantities of BSA dissolved in GEB. The specific activity of GUS was calculated as pmol 4-MU formed per min per mg of total protein.

Transgenic plant material assayed for LUC activity was ground in the presence of LEB (section 2.10.2). LUC activity was determined as in section 2.10.2. Protein concentration was determined using Bradford reagent and a calibration curve prepared using known quantities of BSA dissolved in LEB. The specific activity of LUC was calculated as light units per 10 sec per mg of total protein.

Chapter 3

Pollen-specific transcription is developmentally regulated by multiple activator regions

3.1 Introduction

Previous analysis of the accumulation of steady state *lat*52 mRNA during anther development showed that transcript was first detectable in anthers containing spores undergoing pollen mitosis I, leading to a substantial increase in transcript levels in anthers containing mature pollen (Twell *et al.*, 1989b). The 5' flanking region of the *lat*52 gene (-492 to +110) when fused to the reporter gene *gus* was sufficient to direct GUS activity specifically in mature pollen within anthers of transgenic tomato. In addition, the developmental accumulation of GUS activity in transgenic tomato pollen closely reflected the accumulation of the native *lat*52 transcript (Twell *et al.*, 1989b). Analysis of 5' promoter deletion mutant-*gus-nos3'* fusions in tomato led to the identification of two upstream positive regulatory regions which contribute to the activation of the *lat*52 promoter specifically in pollen: region 52-A -492 to -125 and region 52-B -124 to -72 (fig. 1.2). This analysis also defined the minimal proximal promoter region sufficient to direct preferential reporter gene activity in pollen as the region 52-C (-71 to +110) (fig. 1.2). Using the biolistic transient expression system in tobacco the minimal functional region sufficient for pollen-specific expression was further delimited to the region -41 to +110 (Twell *et al.*, 1991). In addition, the functional activity of the region 52-B (-124 to -72) in pollen was shown to be due to sequences present within the region -100 to -72.

The main aim of this chapter was to identify the regulatory regions of the *lat*52 promoter involved in developmental and cell-specific control. A series of *lat*52 5' promoter deletion mutants fused to the reporter gene *gus* were stably introduced into tobacco. Analysis of these promoter 5' deletion mutants would confirm in transgenic plants the functional activities of the *lat*52 promoter regions defined using the transient expression assay system. Moreover, these plants would facilitate the investigation of the developmental and tissue-specific regulation of the *lat*52 promoter regions.

3.2 Functional characterisation of the *lat*52 promoter in transgenic tobacco

3.2.1 Construction and introduction of *lat52-gus-nos3'* gene fusions into *N. tabacum* cv Samsun via *Agrobacterium* mediated gene transfer

*lat*52 promoter 5' deletion mutants (-100 to +110, -71 to +110 and -41 to +110) were transcriptionaly fused at the ATG of the *E. coli* ß-glucuronidase gene *uidA* (*gus*) and the nopaline synthase 3' region (*nos3'*) (Twell *et al.*, 1991). The steps involved in the construction of *lat52-gus-nos3'* binary plasmid fusions are shown schematically in figure 3.1. *lat52-gus-nos3'* fusions were excised from the plasmids pLAT52-25 (-100 to

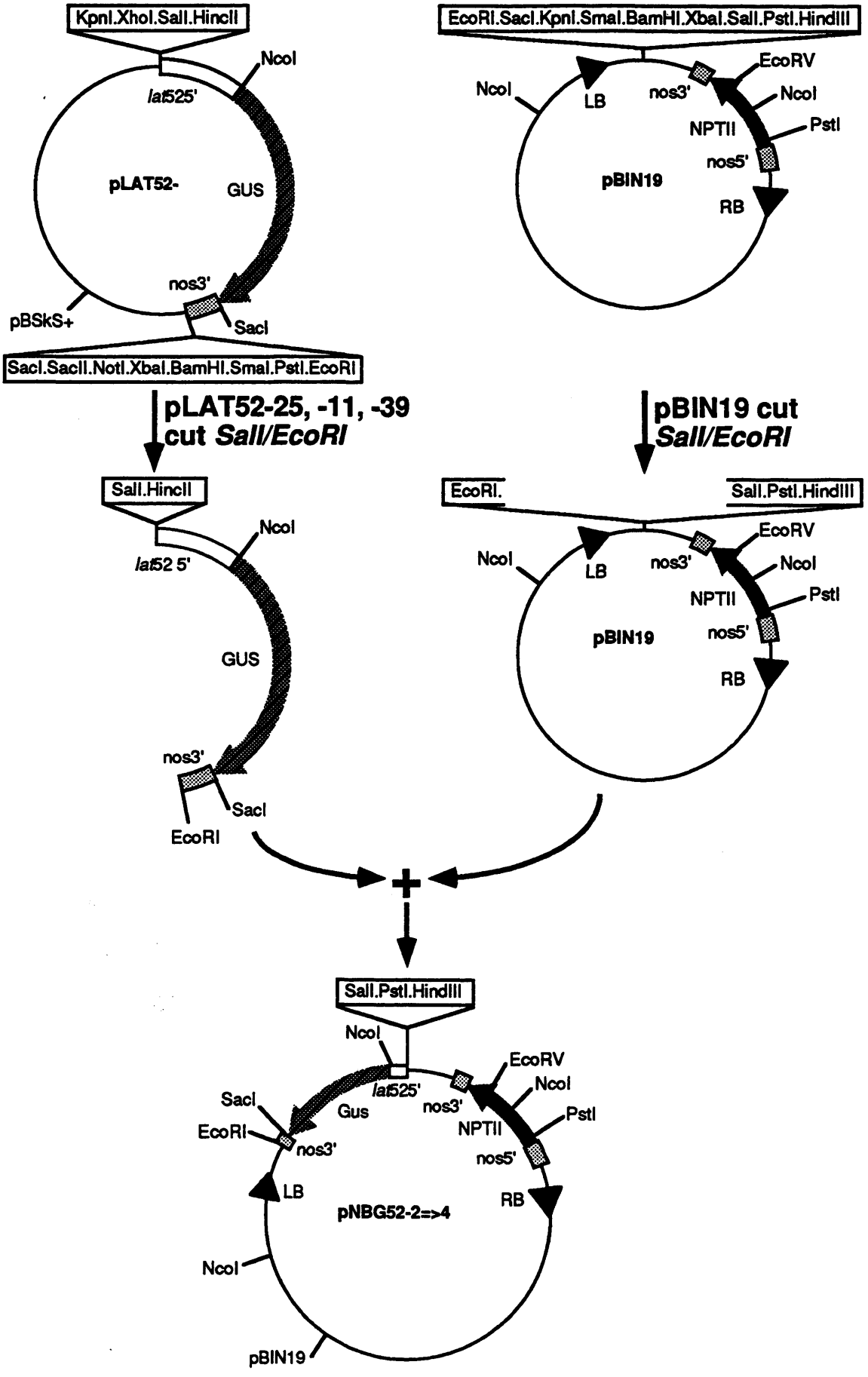


Figure 3.1. Construction of *lat52-gus-nos3'* binary plasmid fusions. *lat*52 promoter 5' deletion mutants -100 to +110, -71 to +110 and -41 to +110 fused to the *gus* reporter gene and nopaline synthase 3'-UTR (*nos3'*) were excised as *SalI/EcoRI* fragments and cloned into the binary vector pBIN19.

+110), pLAT52-11 (-71 to +110) and pLAT52-39 (-41 to +110) (Twell *et al.*, 1991) as *SalI/EcoRI* fragments and subcloned into the binary vector pBIN19 (Bevan, 1984) to give the plasmids pNBLG52-2, 3 and 4 respectively.

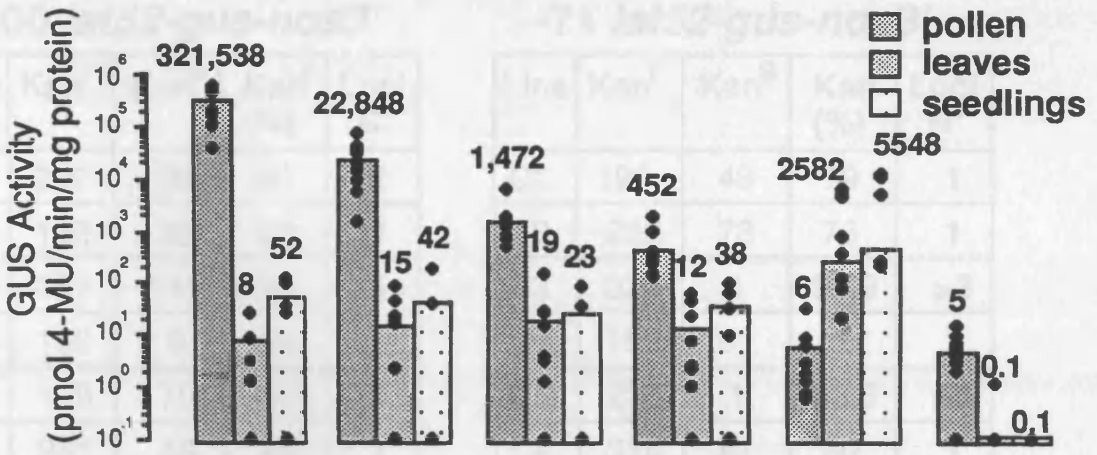
The *lat*52 promoter 5' deletion mutant -492 to +110-*gus-nos3'* fusion was cloned into pBIN19 as a *SalI/EcoRI* fragment and transformed into *A.tumefaciens* by D. Twell (Twell *et al.*, 1990). Transformation of *N. tabacum* cv Samsun was carried out by C. Eady (Eady *et al.*, 1994). The *CaMV35S-gus-nos3'* fusion was constructed and introduced into *N. tabacum* cv Samsun by D. Twell.

3.2.2 Analysis of *lat*52 5' promoter deletion mutants in transgenic tobacco

Twelve to fourteen independent transgenic lines were generated for each *lat52-gus-nos3'* construct. GUS activity was determined (section 2.10.5) in extracts prepared from mature pollen and leaves for all primary transformants (-492 *lat52-gus-nos3'* pollen data from Eady *et al.*, 1994) . GUS activity was determined in extracts prepared from ten independent *CaMV35S-gus-nos3'* transformants (pollen data provided by J. Wilkinson) and a similar number of untransformed plants. GUS activity was also determined in extracts prepared from T_2 seedlings from five independent transformants for each construct and a similar number of untransformed plants. The results of this analysis are presented in figure 3.2.

Deletion of the *lat*52 promoter from -492 to -100 decreased GUS activity by 14-fold to ~ 7 % of the activity of the full-length (-492) promoter, whilst deletion from -100 to -71 decreased GUS activity a further 15-fold to 0.46 %. Removal of 30 bp from -71 to -41 led to a further 3-fold drop in GUS activity to 0.14 %, which was still ~ 75-fold above the activity of the *CaMV35S-gus-nos3'* and untransformed controls. No significant variation in loci number (determined in section 2.2.6) for each group of transformants harbouring each construct (table 3.1: Eady *et al.*, 1994) was seen. Therefore, differences observed in the mean relative activity between *lat*52 promoter 5' deletion mutants are most likely accounted for by the intrinsic transcriptional activity of each 5' promoter deletion and not by an overall increase in transgene loci number for a group of plants harbouring a particular construct.

The above quantitative differences seen between GUS activities for plants harbouring -100 *lat52-gus-nos3'* or -41 *lat52-gus-nos3'* constructs and a untransformed control was visibly highlighted by collecting mature pollen from one line containing each construct and subjecting it to histochemical analysis (section 2.10.3.2). The results of this analysis are presented in figure 3.3. The -100 *lat52-gus-nos3'* line showed a stronger blue-staining than the -41 *lat52-gus-nos3'* line, which in turn was clearly



*lat*52 Deletion end point (bp)	-492	-100	-71	-41	CaMV 35S	WT
Relative activity (%) (pollen)	100	7.11	0.46	0.14	0.002	0.002

Figure 3.2. Analysis of *lat*52 5' promoter deletion mutants in transgenic tobacco plants. The scatter histograms show GUS activities determined in pollen, leaves and T2 seedlings of transgenic plants harbouring either *lat*52 5' promoter deletion mutant-*gus-nos3'* or *CaMV35S gus-nos3'* fusions. Also shown are scatter histograms for GUS activities determined in pollen, leaves and T2 seedlings of non-transformed plants. Each point represents the value derived from the analysis of a single independent transgenic plant, with the mean values indicated by a shaded box. The data summarized below represents the relative activity (%) in pollen determined from the mean GUS activities for each *lat*52 5' promoter deletion mutant-*gus-nos3'* or *CaMV35S-gus-nos3'* fusion and for non-transformed plants.

-100 *lat52-gus-nos3'*

Line	Kan^r	Kan^s	Kan^r (%)	Loci N°
B1	297	28	91	2
B2	168	35	83	1
B3	367	117	76	1
B4	160	9	95	2
B5	109	70	61	1
B6	251	46	85	1
B7	172	75	70	1
B8	137	61	69	1
B9	98	40	71	1
B10	149	31	83	1
B11	250	56	82	1
B12	179	11	94	2
B13	99	0	100	>3
B14	146	88	62	1

-71 *lat52-gus-nos3'*

Line	Kan^r	Kan^s	Kan^r (%)	Loci N°
C1	190	49	79	1
C2	232	73	76	1
C3	320	1	99.9	>3
C4	193	51	79	1
C5	230	1	99.6	>3
C6	334	81	80	1
C7	222	25	90	2
C8	235	66	78	1
C9	129	30	81	1
C10	160	15	91	2
C11	164	52	76	1
C12	326	27	92	2

-41 *lat52-gus-nos3'*

Line	Kan^r	Kan^s	Kan^r (%)	Loci N°
D1	208	5	98	3
D2	311	24	93	2
D3	298	4	99	>3
D4	142	43	77	1
D5	181	97	65	1
D6	179	60	75	1
D7	194	4	98	3
D8	207	15	93	2
D9	182	59	76	1
D10	223	96	70	1
D11	175	49	78	1
D12	250	3	99	>3

Table 3.1. Determination of transgene loci number. Segregation analysis for kanamycin resistance of plants harbouring the -100, -71 and -41 *lat52-gus-nos3'* gene fusions was performed as in section 2.2.6. Loci number was estimated from the percentage of Kanamycin resistance seedlings (Kan^r %).

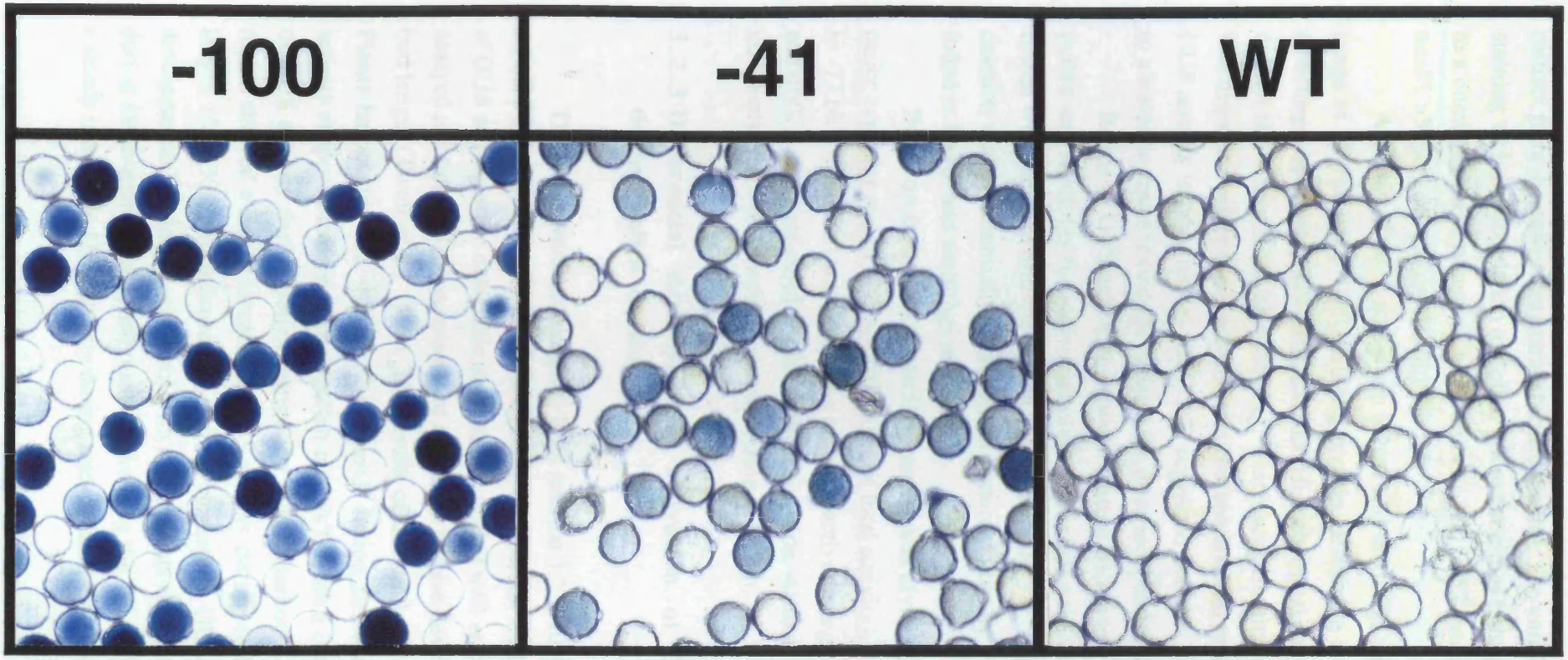


Figure 3.3. Histochemical localisation of GUS activity in pollen. Mature pollen from one line harbouring the -100 *lat52-gus-nos3'* fusion (-100), one line harbouring the -41 *lat52-gus-nos3'* fusion (-41) and a wild type control (WT) were incubated at room temperature in the presence of 1 mM X-gluc and 0.5 mM potassium ferricyanide for 16 hours. Magnification x 200.

distinct from the colourless untransformed control. The variability in the strength of blue-staining within a population of pollen grains for either transgenic lines is most likely due to a combination of transgene loci number (-100 *lat52-gus-nos3'*, 2 loci: -41 *lat52-gus-nos3'*, >3 loci) and pollen viability.

Analysis of GUS activity in leaves and seedlings showed GUS activities in the range of 10-180 pmol 4-MU/min/mg protein, detectable in ~ 30 % (leaves) and 55 % (seedlings) of independent transformants. Histochemical analysis of leaves and seedlings did not show detectable blue-staining, which suggests that this low level of expression was dispersed rather than localised to a few specific cell types. Fluorometric analysis of GUS activity in other tissues (table 3.2) suggested that this low level activity was present to a lesser or greater extent throughout the plant independent of promoter length.

Ratios of GUS activities between mature pollen and leaves and between mature pollen and seedlings for individual plants harbouring the -41 *lat*52 promoter-*gus-nos3'* fusion were derived and the results shown in table 3.3. This analysis showed that this deletion was preferentially expressed in mature pollen by 75- and 33-fold above levels found in leaves and seedlings respectively.

Based on these data the *lat*52 promoter was divided into three functional units: a major activator region -492 to -101 (93 % of total activity), a minor activator region -100 to -72 (6.7 % of total activity) and a minimal activator region -71 to -42 (0.3 % of total activity). It also confirmed *in planta*, that the region -41 to +110 is the minimal functional unit required for preferential gene expression in pollen.

3.2.3 Differential developmental activation of *lat*52 activator regions during pollen development

Three independent transgenic lines (section 3.2.3) harbouring each *lat*52 promoter deletion mutant were selected for detailed developmental analyses. Since negligible levels of GUS activity were detectable in the anther wall (Twell *et al.*, 1990), GUS activity was assayed (section 2.10.5) in extracts of anthers at precise developmental stages defined by bud length (Twell *et al.*, 1993). The results of this analysis are presented in figure 3.4. Plants harbouring the -492 *lat*52-*gus-nos3'* fusion showed a sharp increase in GUS activity at approximately 16 mm (pollen-mitosis I), with a continuous rapid accumulation of GUS activity during pollen maturation. Plants harbouring the -100 *lat*52-*gus-nos3'* fusion showed a steady increase in GUS activity during early-pollen development (bud length 16 to 30 mm), followed by a sharp increase in GUS activity during mid-pollen development (bud length ~ 30 mm) with a continuous rapid accumulation of GUS activity during maturation. In contrast, plants harbouring the -71 *lat*52-*gus-nos3'* fusion showed a steady increase in GUS activity during mid- to late-pollen maturation (bud length ~ 30

-492 *lat52-gus-nos3'*

Line	Pollen	Sepal	Petal	Pistil	Immature seed
A1	91,657	94	20	22	29
A3	442,788	ND	56	3	14
A13	678,451	67	73	3	16

-100 *lat52-gus-nos3'*

Line	Pollen	Sepal	Petal	Pistil	Immature seed
B1	28,292	ND	30	ND	297
B2	39,101	ND	ND	ND	54
B4	69,693	37	10	ND	6

-71 *lat52-gus-nos3'*

Line	Pollen	Sepal	Petal	Pistil	Immature seed
C1	338	73	10	22	5
C2	2,628	8	ND	ND	ND
C12	1,538	ND	ND	ND	18

-41 *lat52-gus-nos3'*

Line	Pollen	Sepal	Petal	Pistil	Immature seed
D7	474	ND	ND	ND	24
D10	101	ND	ND	ND	158
D12	593	200	ND	5	18

CaMV35S5'-gus-nos3'

Line	Pollen	Sepal	Petal	Pistil	Immature seed
ST24	47	12,905	2,291	1,173	106

WT

Line	Pollen	Sepal	Petal	Pistil	Immature seed
	6	ND	ND	ND	2

Table 3.2. GUS enzyme activity in plant tissue extracts. GUS activity, pmol 4-MU per min per mg of total protein, was determined in extracts prepared from pollen, sepal, petal, pistil and immature seed for three individual lines harbouring the -492, -100, -71 and -41 *lat52-gus nos3'* transgenes, for one line harbouring the *CaMV35S-gus-nos3'* gene fusion and also for a wild type (WT) control. ND represents non detectable.

Line	Pollen	Leaf	Seedling	$\frac{\text{Pollen}}{\text{Leaf}}$	$\frac{\text{Pollen}}{\text{Seedling}}$
D1	381	7	7	54	54
D2	152	1	ND	152	-
D3	226	2	59	113	4
D4	313	2	ND	157	-
D5	278	21	7	13	40
D6	273	ND	ND	-	-
D7	996	ND	30	-	33
D8	153	ND	8	-	19
D9	391	43	6	9	65
D10	255	ND	9	-	28
D11	125	ND	ND	-	-
D12	1881	64	91	29	21

Mean $\frac{\text{Pollen}}{\text{Leaf}}$ = 75-fold Mean $\frac{\text{Pollen}}{\text{Seedling}}$ = 33-fold

Table 3.3. GUS enzyme activity in tissue extracts of plants harbouring the transgene -41 *lat52 gus-nos3'*. GUS activity, pmol 4-MU per min per mg of total protein, was determined in extracts prepared from pollen, leaf and seedlings for all lines harbouring the -41 *lat52-gus nos3'* transgene. ND represents non detectable. Ratios of GUS activity between pollen and leaf; pollen and seedling were calculated and the average fold enhancement presented.

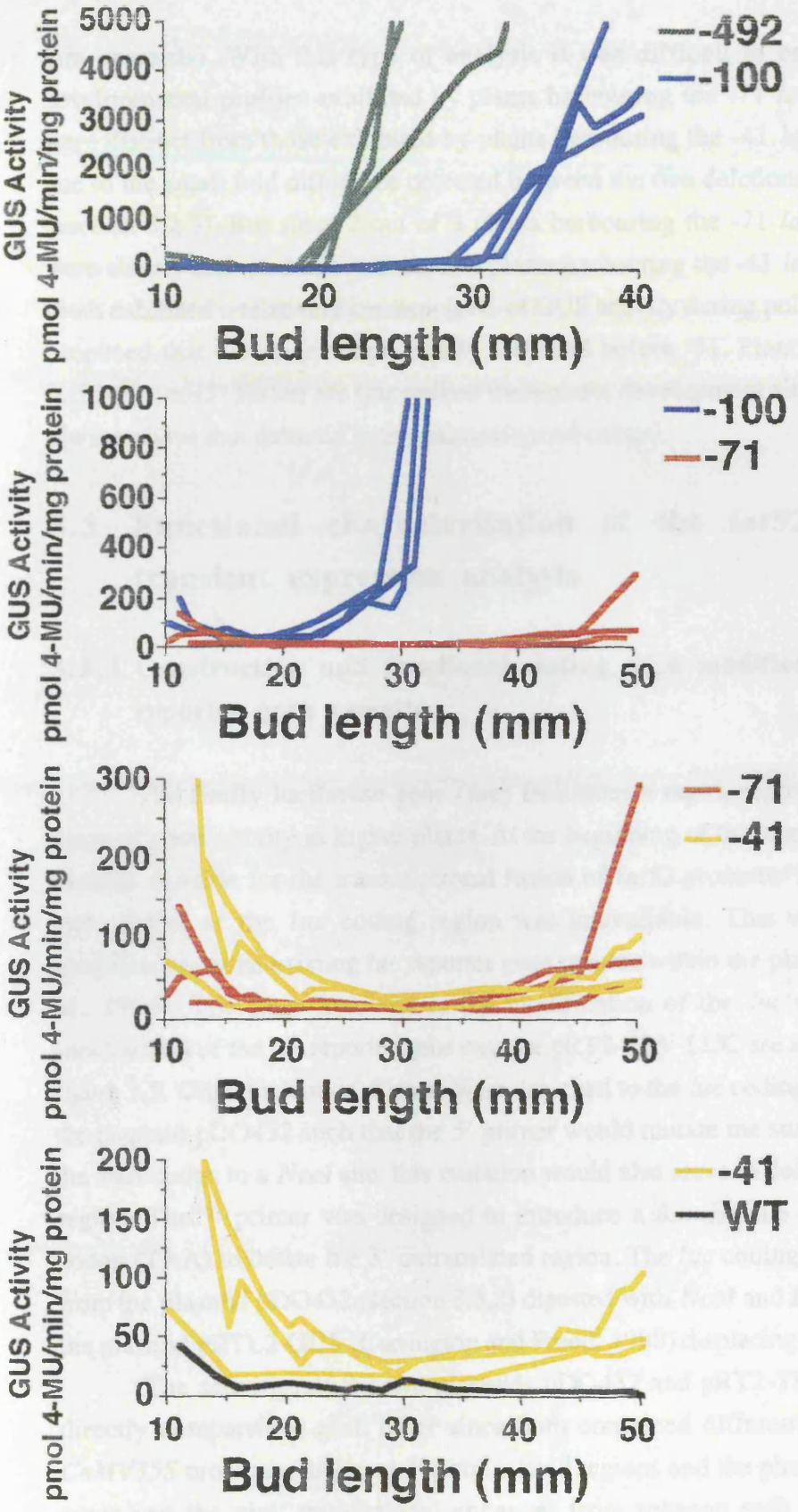


Figure 3.4. Developmental regulation of the -492, -100, -71 and -41 *lat*52 5' promoter deletion mutants in transgenic tobacco. The GUS activities in extracts prepared from anthers isolated from buds of different length are shown for: three individual lines harbouring the *lat*52 5' promoter deletion -492 to +110-*gus-nos3'* fusion (green lines), three individual lines harbouring the *lat*52 5' promoter deletion -100 to +110-*gus-nos3'* fusion (blue lines), three individual lines harbouring the *lat*52 5' promoter deletion -71 to +110-*gus-nos3'* (red lines), three individual lines harbouring the *lat*52 5' promoter deletion -41 to +110-*gus-nos3'* (yellow lines) and a wild type (WT) control (black line).

mm onwards). With this type of analysis it was difficult to establish whether the developmental profiles exhibited by plants harbouring the -71 *lat52-gus-nos3'* fusion were distinct from those exhibited by plants harbouring the -41 *lat52-gus-nos3'* fusion due to the small fold difference detected between the two deletions in transgenic plants (section 3.2.3). But since 2 out of 3 plants harbouring the -71 *lat52-gus-nos3'* fusion were clearly activated above 2 out of 3 plants harbouring the -41 *lat52-gus-nos3'* fusion (both exhibited a relatively constant level of GUS activity during pollen development) it is proposed that -71 is developmentally activated before -41. Plants harbouring the -41 *lat52-gus-nos3'* fusion are transcribed throughout development since GUS activity was always above that detected in the untransformed control.

3.3 Functional characterisation of the *lat52* promoter by transient expression analysis

3.3.1 Construction and functional testing of a modified firefly luciferase reporter gene cassette

The firefly luciferase gene (*luc*) facilitates a rapid, reproducible and sensitive assay of gene activity in higher plants. At the beginning of this thesis a *luc* reporter gene cassette suitable for the transcriptional fusion of *lat52* promoter fragments at the ATG start codon of the *luc* coding region was unavailable. This was overcome by the modification of an existing *luc* reporter gene present within the plasmid pDO432 (Ow *et al.*, 1986). The steps involved in the modification of the *luc* reporter gene and the construction of the *luc* reporter gene cassette pRT2-TEV-LUC are shown schematically in figure 3.5. Oligonucleotide primers were designed to the *luc* coding region present within the plasmid pDO432 such that the 5' primer would mutate the surrounding sequence of the start codon to a *NcoI* site, this mutation would also serve to delete the 5' untranslated region. The 3' primer was designed to introduce a *BamHI* site directly after the stop codon (TAA) to delete the 3' untranslated region. The *luc* coding region was amplified from the plasmid pDO432 (section 3.3.2) digested with *NcoI* and *BamHI* and cloned into the plasmid pRTL2-GUS (Carrington and Freed. 1990) displacing *gus*.

The activities of the two plasmids pDO432 and pRT2-TEV-LUC could not be directly compared to each other since both contained different combinations of the *CaMV35S* promoter, different 3' untranslated regions and the plasmid pRT2-TEV-LUC contained the viral translational enhancer from tobacco etch virus (TEV 5'-UTR) (Carrington and Freed, 1990). Therefore a promoterless version (pNBL-0) of pRT2-TEV-LUC was constructed. The steps involved in the construction of the plasmid pNBL-0 are shown schematically in figure 3.6. The plasmid pRT2-TEV-LUC was digested with

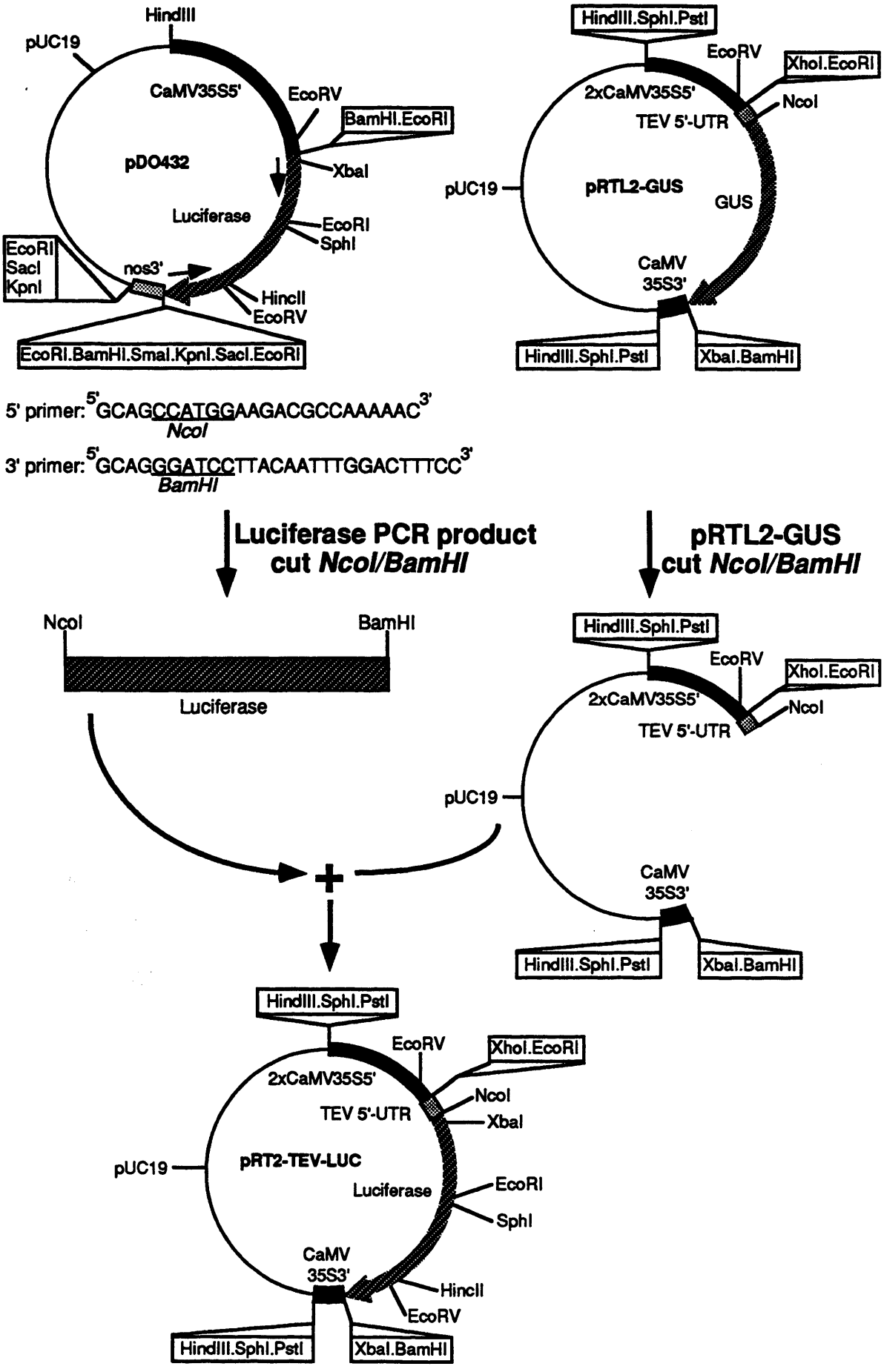


Figure 3.5. Construction of the modified *luc* reporter gene cassette pRT2-TEV-LUC. The *luc* reporter gene was amplified by PCR, cut *NcoI/BamHI* and cloned into the plasmid pRTL2GUS.

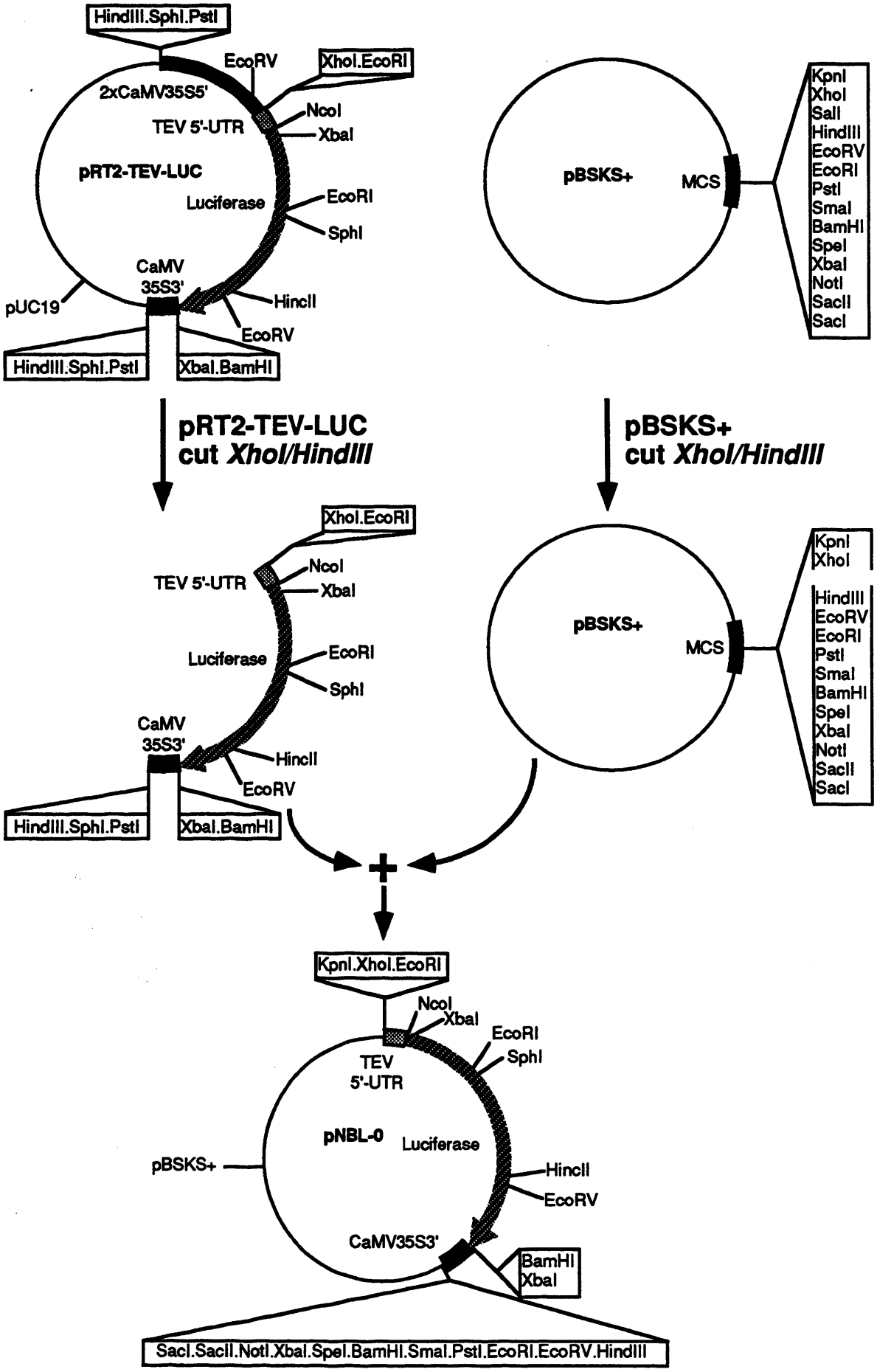


Figure 3.6. Construction of the plasmid pNBL-0. *TEV 5'-UTR-luc-CaMV35S3'* was excised as an *XhoI/HindIII* fragment from pRT2-TEV-LUC and inserted into the multiple cloning site (MCS) of pBSKS+.

XhoI and *HindIII* (section 2.6.1) and the 2.0 kb *TEV 5'-UTR-luc-CaMV35S3'* fragment subcloned into pBSKS+.

The test plasmids pDO432, pRT2-TEV-LUC and pNBL-0 were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in leaves (section 2.9.6) in conjunction with a no DNA control. The results of this analysis are presented in figure 3.7. The activity of the plasmid pRT2-TEV-LUC was ~ 800- and 6,000-fold above the promoterless control (pNBL-0) and the no DNA control respectively. These results showed that the modified *luc* gene was functional and also highlighted the extremely low level of background luciferase activity in leaf tissue.

The activity of pRT2-TEV-LUC was 19-fold above that seen for pDO432. This difference in activity is most likely the result of a stronger promoter-5'-UTR combination in pRT2-TEV-LUC (Ow *et al.*, 1987: Carrington and Freed, 1990). Therefore, it can be assumed that no major negative mutations were introduced into the modified *luc* gene in the plasmid pRT2-TEV-LUC.

3.3.2 Construction and transient expression analysis of *lat52* promoter 5' deletion mutant-*luc-CaMV35S3'* gene fusions in pollen

*lat*52 promoter 5' deletion mutants (-492 to +110, -100 to +110, -71 to +110, -41 to +110 and -17 to +110) were transcriptionaly fused at the ATG of the *E. coli* ß-glucuronidase gene *uidA* (*gus*) and the nopaline synthase 3' region (*nos3'*) (Twell *et al.*, 1991). The steps involved in the construction of *lat52-luc-CaMV35S3'* plasmid fusions are shown schematically in figure 3.8. *lat*52 promoter regions were excised from the plasmids pLAT52-7 (-492 to +110), pLAT52-25 (-100 to +110), pLAT52-11 (-71 to +110), pLAT52-39 (-41 to +110) and pLAT52-19 (-17 to +110) (Twell *et al.*, 1991) as *KpnI/NcoI* fragments and subcloned into the plasmid pNBL-0 displacing the TEV 5'-UTR. Plasmids constructed were termed pNBL52-1 (-71), -2 (-100), -5 (-492), -6 (-17) and -7 (-41).

The *lat*52 promoter regions -67 to +110 and -51 to +110 were amplified from the template pNBL52-5 (section 2.5.2) in conjunction with specific 5' oligonucleotides containing a *SalI* site at the relevant 5' border sequence of the *lat*52 promoter and a 3' oligonucleotide complementary to the nonsense strand of *luc*. The generated PCR products were digested with *SalI and NcoI* (section 2.6.1) and subcloned into the plasmid pNBL52-5 displacing the *lat*52 promoter region (-492 to +110). Plasmids constructed were termed pNBL52-10 (-67) and -8 (-51). The plasmid pNBL52-26 (chapter 5) was digested with *XhoI* and *HindIII* and the *lat52 5'-UTR (+1 to +110) -luc-CaMV35S3'* fragment subcloned into pBSKS+ to give the plasmid pNBL52-30 (+1).

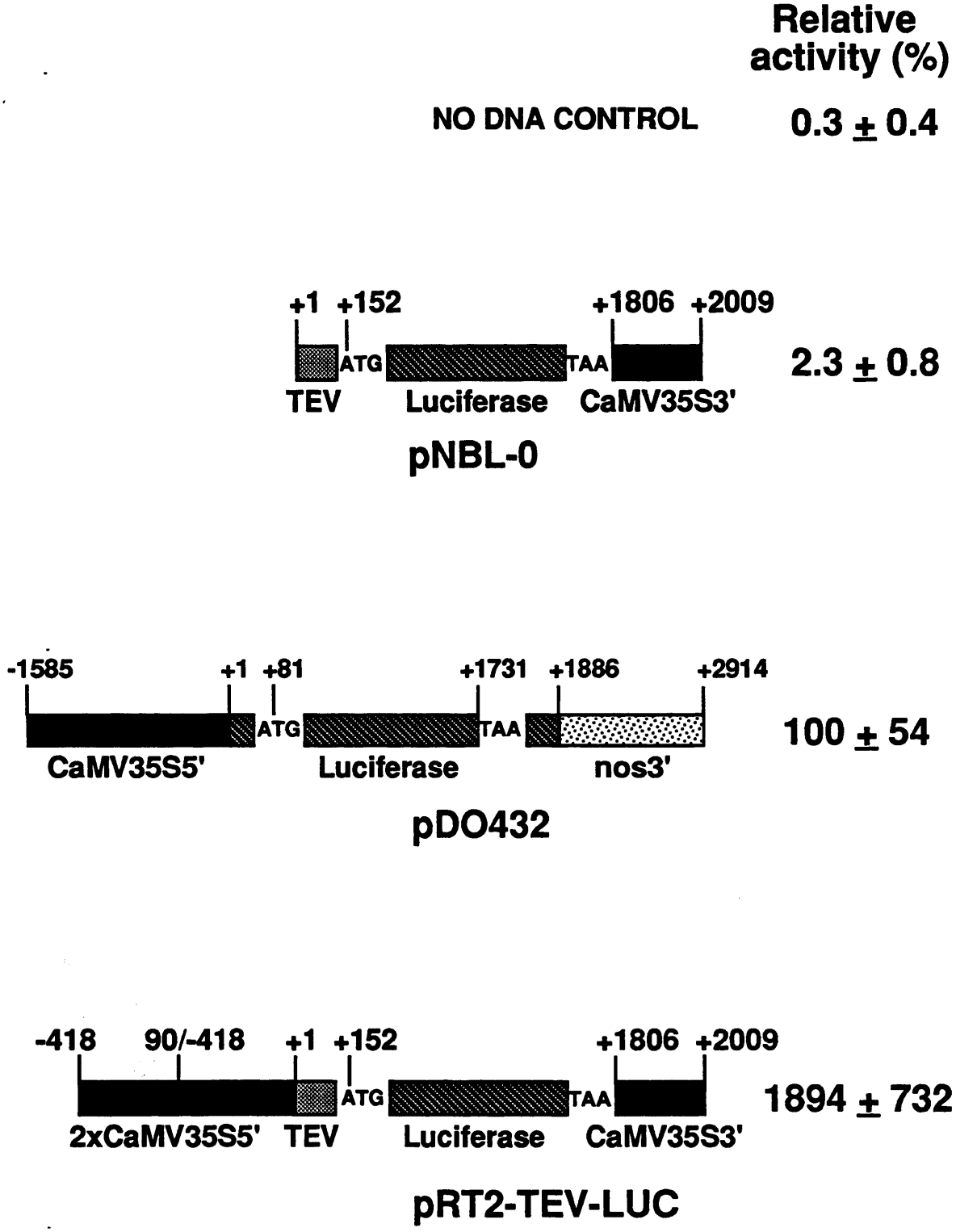


Figure 3.7. Transient expression analysis of the modified luciferase reporter gene in leaves. The test plasmids pNBL-0, pDO432 and pRT2-TEV-LUC were co-bombarded into leaves with the reference plasmid pRTL2GUS in conjunction with a no DNA control. Relative activities of test plasmids were calculated from a ratio of LUC:GUS normalised to the activity of the test plasmid pDO432 (average light units per bombardment was 125363). Data from two independent experiments are shown, representing six independent bombardments for each test plasmid. The standard error is shown on the right.

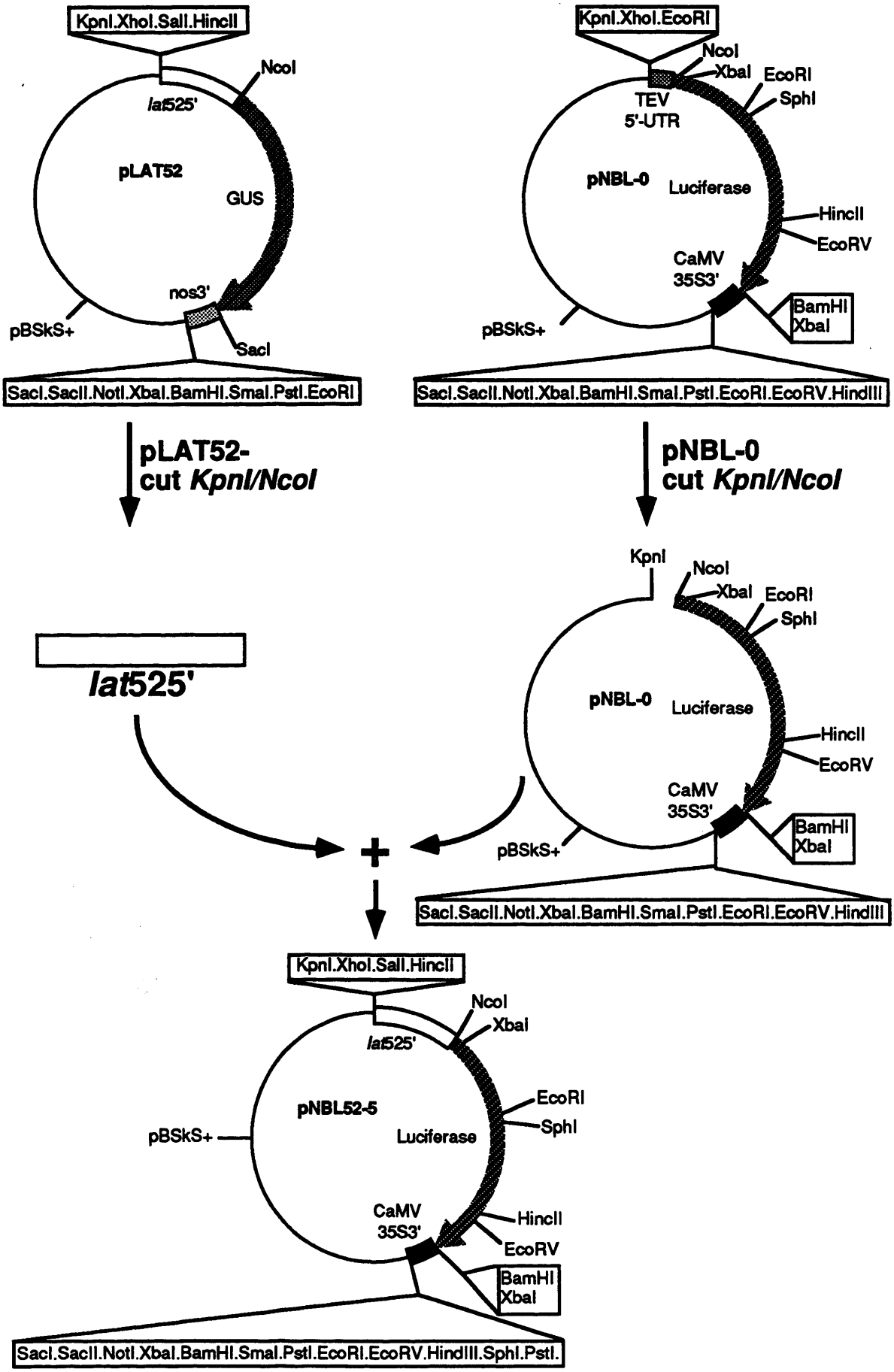


Figure 3.8. Construction of *lat52-luc-C3'* plasmid fusions. The *lat52* promoter 5' deletion mutants -492, -100, -71, -41 and -17 to +110 were excised from the relevant pLAT52 (*lat52 gus-nos3'*) plasmids (Twell *et al.*, 1991) as *SalI/NcoI* fragments and cloned into the plasmid pNBL-0.

At the beginning of this thesis a method existed for the transient expression analysis of promoter-reporter gene fusions in pollen using microprojectile bombardment (Twell *et al.*, 1989a). In this method pollen was cultured on MS30 (section 2.2.3), which severely restricted pollen germination and tube growth. This problem was overcome by using an optimised media for pollen germination and tube growth (PGM) (Tupy *et al.*, 1991). In order to establish the optimum time in culture in which maximum levels of gene expression were obtained from bombarded pollen a time course was performed. The plasmid pNBL52-5 was transiently expressed in pollen (section 2.9.6) along with a no DNA control. Luciferase (LUC) activity was determined in pollen (section 2.10.2) at 2 hour intervals post bombardment up to a maximum time of 24 hours. The results of this analysis are presented in figure 3.9. In pollen bombarded with the plasmid pNBL52-5 LUC activity was first detectable at 2 hours. LUC activity increased proportionally up to 16 hours followed by a reduction in activity, presumed to be due to a cessation in gene expression coupled with enzyme turn-over. In pollen bombarded with the no DNA control an extremely low level of LUC activity was detectable throughout the time course. From this analysis the optimum time to analyse gene expression in pollen was 16 hours post bombardment. Since the optimum time for gene expression in pollen cultured on MS30 was 6 hours (Twell *et al* ., 1989a), the extended culture time seen for pollen cultured on PGM would indicate that pollen remained viable longer, resulting in higher levels of gene expression.

The test plasmids pNBL52-1, -2, -5, -6, -7, -8, -10 and -30 were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5). Test plasmid activities were calculated (section 2.10.4) and normalised to the -492 to +110 *lat*52 5' promoter deletion mutant. The results of this analysis are presented in figure 3.10. Deletion from -492 to -100 decreased LUC activity by 2.5-fold, deletion from -100 to -71 decreased LUC activity a further 3-fold. Deletion from -71 to -67 decreased LUC activity by 4.5-fold, further deletions to -51 and -41 increased LUC activity by 1.6- and 1.7-fold respectively. Deletion of the whole region from -71 to -41 decreased LUC activity by 1.5-fold. Further deletion from -41 to -17, removing the first putative TATA box led to a 8-fold decrease in LUC activity, but deletion from -17 to +1, which removed a second putative TATA box did not lead to a further decrease in LUC activity. These data confirm the function of the regions -492 to -101, -100 to -72 and -71 to -42 as activator domains of the *lat*52 promoter in germinating pollen and indicate that the region -41 to -18 is the minimal transcriptional unit required for pollen-specific gene expression.

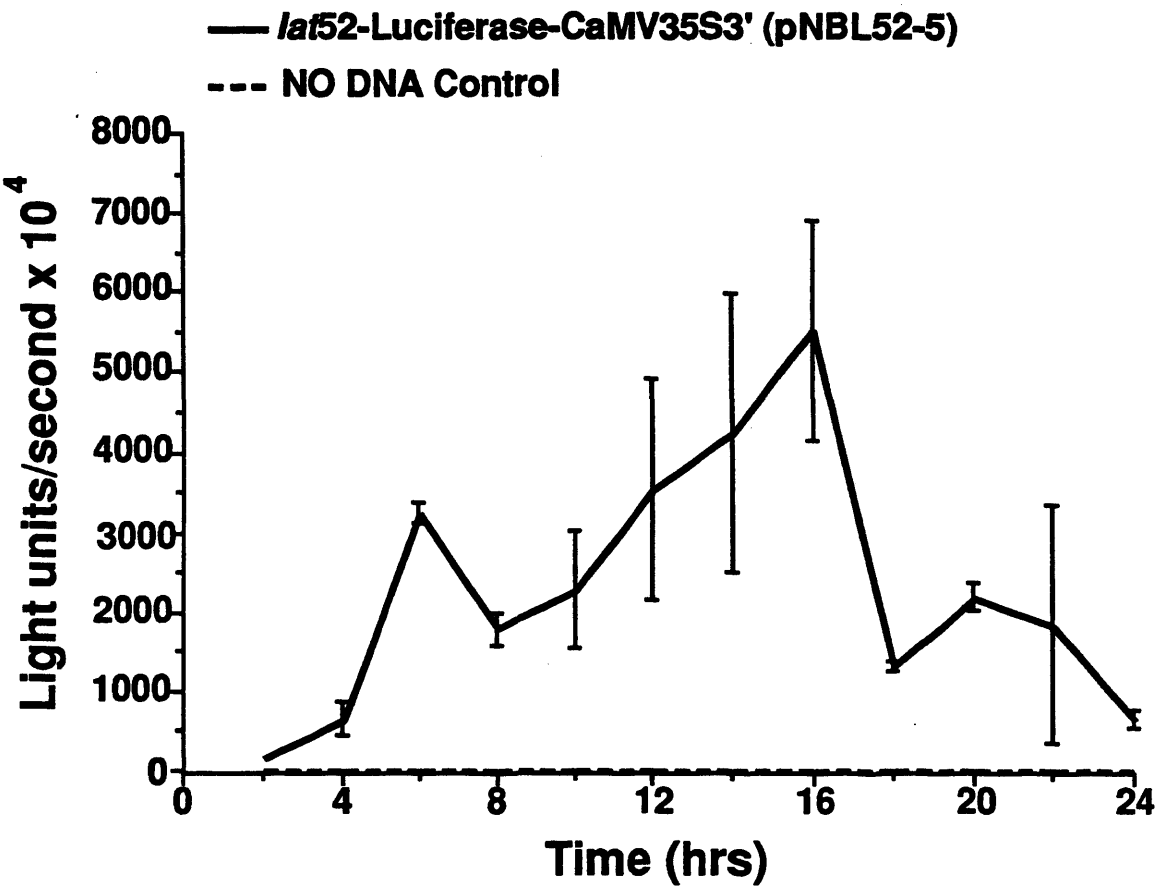


Figure 3.9. Transient gene expression during pollen germination. The plasmid pNBL52-5 was transiently expressed in pollen in conjunction with a no DNA control. Luciferase (LUC) activity was determined in pollen at 2 hour intervals post bombardment up to a maximum time of 24 hours. Data from one experiment is shown, representing the mean of two independent bombardments. Error bars represent the standard error.

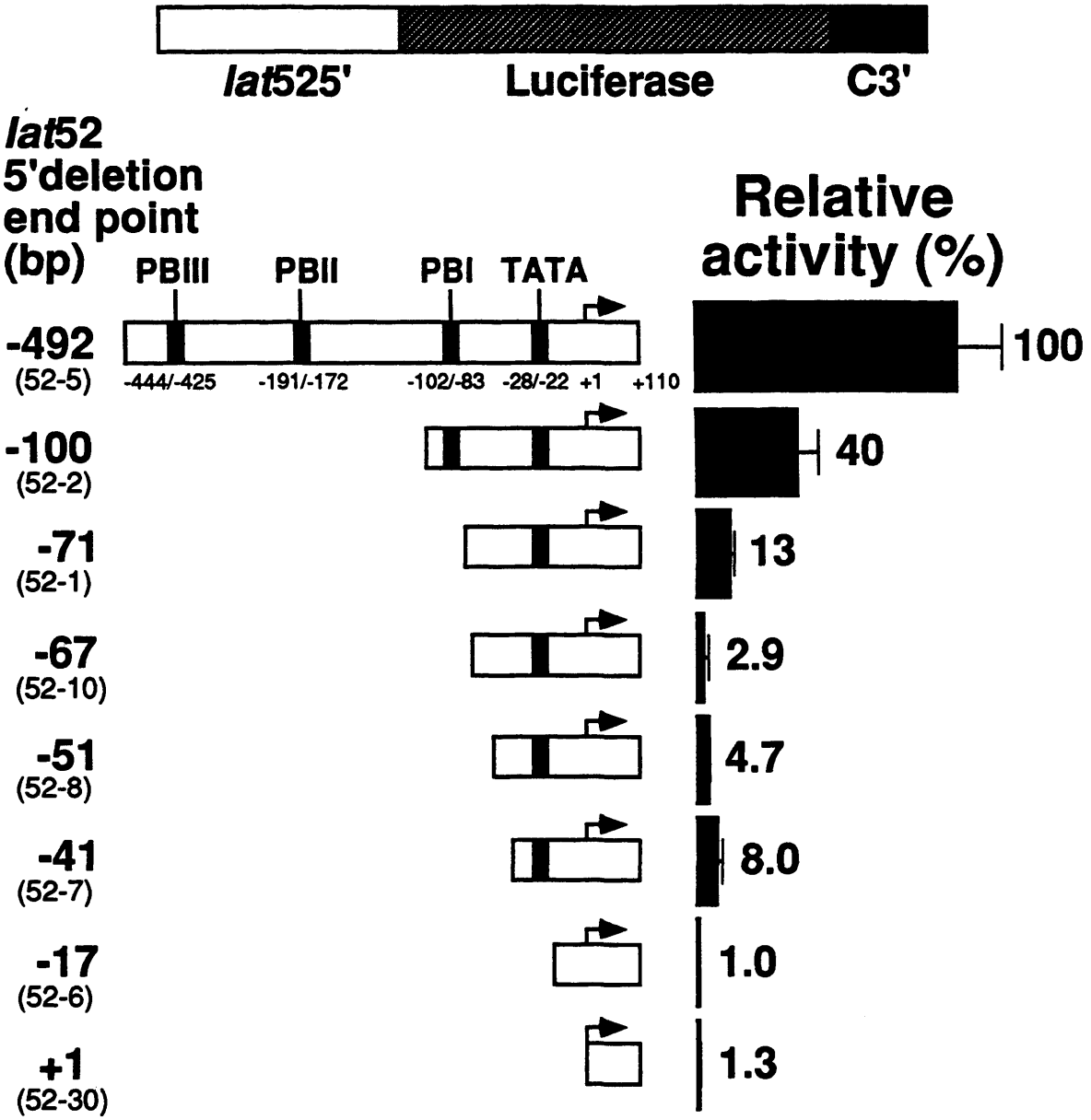


Figure 3.10. Transient expression analysis of *lat*52 5' promoter deletion mutants in pollen. Test plasmids containing *lat*52 5' promoter deletion-*luc-C3'* gene fusions were co-bombarded into pollen with the reference plasmid pLAT52-7. Relative activities of test plasmids containing *lat52-luc-C3'* fusions were calculated from a ratio of LUC:GUS normalised to the activity of the largest 5' deletion mutant (-492 to +110; average light units per bombardment was 356674380). Data from two independent experiments are shown, representing six independent bombardments for each plasmid. Error bars represent the standard error.

3.3.3 Ectopic activation of *lat*52 5' promoter deletion mutants by a heterologous enhancer

The data presented in section 3.2.3 highlighted a low but detectable level of GUS activity in both leaves and seedlings of transgenic plants harbouring *lat52-gus-nos3'* fusions. To determine whether GUS activity in sporophytic tissues was the result of genomic position effects or to a low level of constitutive *lat*52-promoter activity hybrid *CaMV35S/lat52* promoter fusions were constructed. Previous analysis showed that insertion of the *CaMV35S* enhancer region domain B upstream of the chalcone synthase promoter, transcriptionally enhanced promoter activity specifically within the cell-types that chalcone synthase is normally expressed (van der Meer *et al.*, 1992). Therefore, the hypothesis was tested that the addition of the *CaMV35S* enhancer region upstream of the *lat*52 promoter would only enhance activity in the sporophytic tissue types where *lat*52 is expressed.

The steps involved in the construction of *CaMV35S/lat52* promoter fusions are shown schematically in figure 3.11. The duplicated *CaMV35S* region (-418 to -90/-418 to -78) was amplified from the plasmid pRT2-TEV-LUC (section 3.3.1) with specific oligonucleotides designed to introduce a *KpnI* site at the 5' end and a *SalI* site at the 3' end. The generated PCR product was digested with *KpnI* and *SalI* and introduced upstream of the *lat*52 promoter regions in the plasmids pNBL52-1, -2, -6 and -7. The *CaMV35S-lat52-luc-CaMV35S3'* fusions were termed pNBL52-34 (-71), -35 (-17), -36 (-41) and -53 (-100).

The test plasmids pNBL52-1, -2, -6, -7, -34, -35, -36 and -53 were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in leaves (section 2.9.6). Test plasmid activities were calculated (section 2.10.4) and normalised to the -100 to +110 *lat*52 5' promoter deletion mutant. The results of this analysis are presented in figure 3.12. The -100, -71 and -41 5' promoter deletions enhanced LUC activity 4- to 5-fold above the TATA-less -17 control. Addition of the *CaMV35S* enhancer region to each promoter deletion enhanced LUC activity 20- to 33-fold, but the relative differences between each promoter deletion remained approximately the same. These results strongly suggest that the *lat*52 promoter is weakly activated in sporophytic tissue, and that since the activities of all deletion mutants were strongly activated by the duplicated *CaMV35S* enhancer, no strong negative sporophytic *cis*-regulatory elements reside within the region -100 to +110. These results also indicate that pollen-specific transcriptional activator sequences are absent from the region -41 to -18, since deletion of this region significantly reduced LUC activity both in leaves and pollen (section 3.3.2).

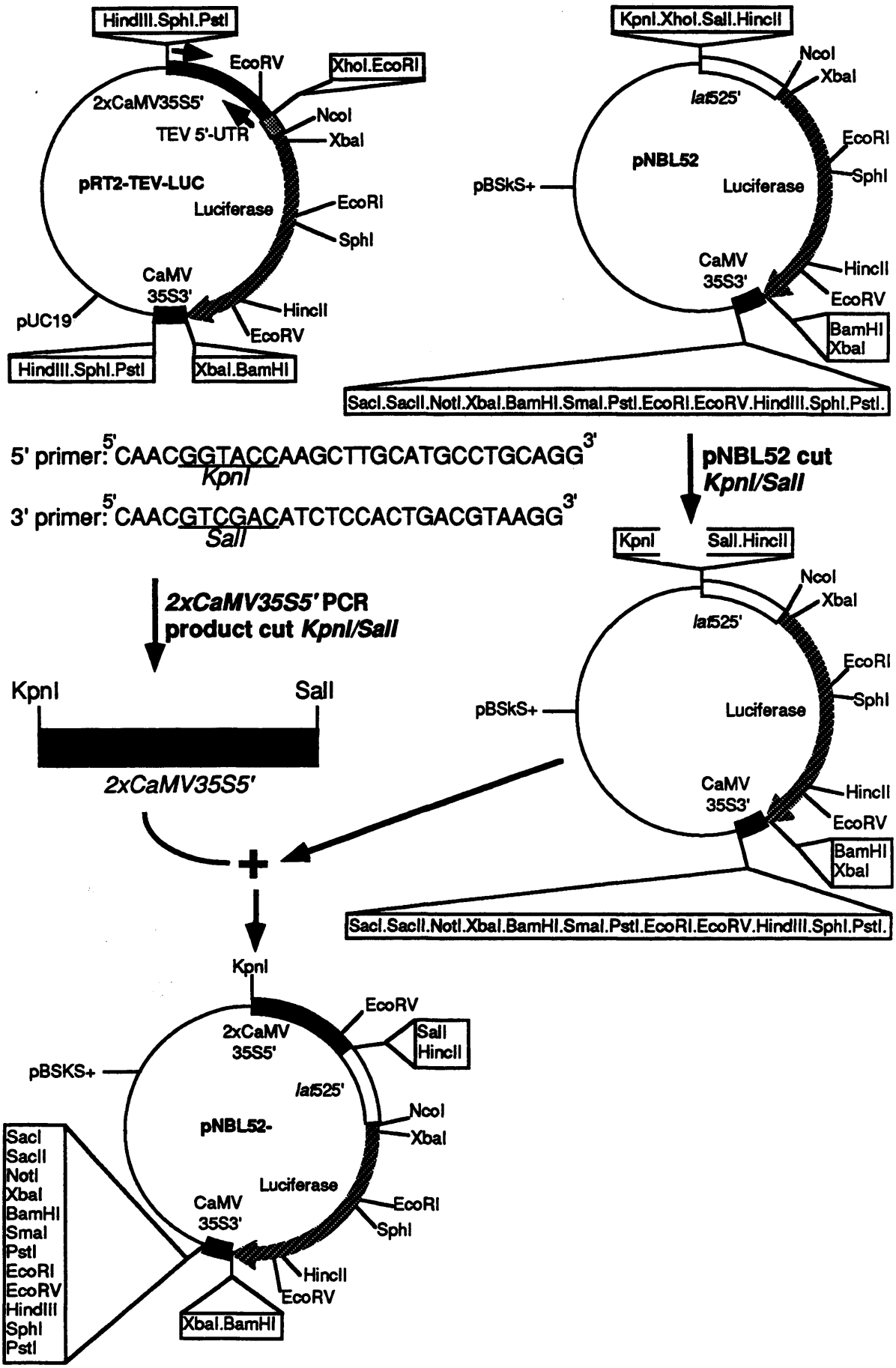


Figure 3.11. Construction of *2xCaMV35S5'/lat52* hybrid promoter fusions. The duplicated *CaMV35S* region (-418 to -90/-418 to -78) was amplified from the plasmid pRT2-TEV-LUC by PCR digested with *KpnI* and *SalI* and cloned into the plasmids pNBL52-1 (-71), -2 (-100), -6 (-17) and -7 (-41).

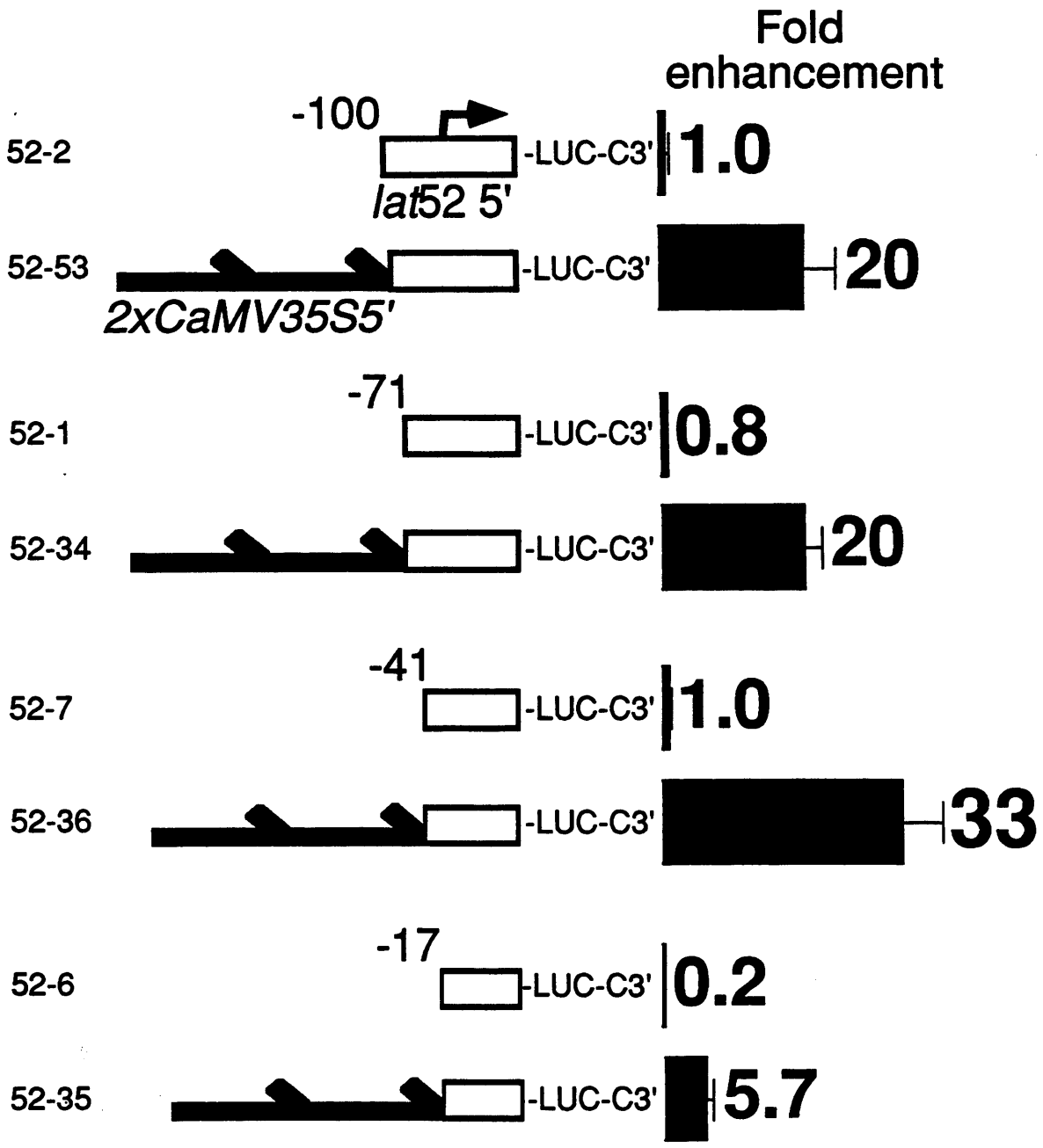


Figure 3.12. Transient expression analysis of *lat*52 5' promoter deletion mutants in leaves. Test plasmids containing either *lat*52 5' promoter deletion mutant-*luc-C3'* or *2xCaMV35S5'-lat52* 5' promoter deletion mutant-*luc-C3'* gene fusions were co-bombarded into leaves with the reference plasmid pRTL2GUS. Relative activities of test plasmids were calculated from a ratio of LUC:GUS normalised to the activity of the *lat*52 5'deletion mutant, -100 to +110 (average light units per bombardment was 14285). Data from two independent experiments are shown, representing six independent bombardments for each test plasmid. Error bars represent the standard error.

3.4 Discussion

3.4.1 Major pollen-specific transcriptional *cis*-regulatory elements reside upstream of -41

Analysis of *lat*52 5' promoter deletion mutant-*gus* fusions in transgenic tobacco showed that the region -492 to -42 accounted for 99.86 % of *lat*52 promoter activity in pollen (section 3.2.2). In addition, this analysis also demonstrated that in transgenic tobacco the minimal *lat*52 proximal promoter region -41 to +110 preferentially enhanced GUS activity in pollen to levels ~ 75- and 33-fold above those in leaves and seedlings respectively. Progressive 5' deletion analysis of the region -41 to +110 showed that deletion of the promoter to -17 was sufficient to reduce activity to the basal level of the promoterless control (+1 to +110) in pollen (section 3.3.2). These results imply that sequences controlling the preferential expression of the region -41 to +110 in mature pollen reside between -41 to -18. However, deletion of the region -41 to -18 led to a similar decrease in promoter activity in pollen (section 3.3.2) and leaves (section 3.3.3). Furthermore, substitution mutational analysis of the region -41 to -29 in context of the 5' promoter deletion -100 to +110 did not affect promoter activity in pollen (Eyal *et al.*, 1995). Since no apparent pollen-specific transcriptional activator sequences lie within the region -41 to -1 pollen-specific regulatory sequences are most likely to be found within the 5' untranslated region (5'-UTR).

A *cis*-regulatory element which influences mRNA levels specifically in pollen has been identified in the 5'-UTR of maize *Adh1* via transposon mutagenesis (Dawe *et al.*, 1993). The level at which this element regulates *Adh1* remains unclear since it could either transcriptionally up regulate *Adh1* expression or increase the stability of *Adh1* mRNA. Therefore, it would be important to determine the level at which the *lat*52 5'-UTR regulates *lat*52 expression in pollen. Calculation of the translational yield from *lat52-gus* transcripts in pollen and leaves/ or seedlings of transgenic plants harbouring the -41 *lat52-gus-nos3'* fusion would indicate if pollen-specific transcriptional *cis*-regulatory elements are present within the *lat*52 5'-UTR.

In conclusion the region -41 to +110 contains a TATA box (-28 to -25) which is sufficient to support similar basal levels of transcription in both pollen and leaves. Preferential expression of the region -41 to +110 in pollen is achieved solely as a result of the pollen-specific enhancement mediated by the *lat*52 5'-UTR. Moreover, the predominant pollen-specific transcriptional *cis*-regulatory elements reside upstream of-41.

3.4.2 Pollen-specific transcription is coordinated by multiple developmentally regulated activator regions

Results presented in figure 3.2 show low but detectable levels of GUS activity in leaves and seedlings from transgenic plants harbouring *lat*52 5' promoter deletion mutant-*gus* fusions. Deriving a ratio of mean GUS activity between pollen and leaves indicates that the -492 to +110 promoter region is preferentially expressed in pollen by ~ 40,000-fold. Previous analysis of endogenous transcript in tomato showed that the steady state level of *lat*52 mRNA was at least 200-fold higher in pollen than in leaves, where no signal was detected (Twell *et al.*, 1989b). Taken together, these analyses would indicate that *lat*52 is expressed at a very low level in sporophytic tissues but transcripts were not detected due to the sensitivity of the northern blot analyses used.

Alternatively, low level GUS activity detected in sporophytic tissues could have arisen from the integration of *lat*52 transgene fusions into or close to sporophytic enhancer/activator elements. Analysis of a promoterless *gus* reporter gene in transgenic tobacco showed that the spatial pattern and level of GUS activity was extremely variable but strictly dependent upon the sequences surrounding the point of transgene integration (Topping *et al.*, 1991). The range of GUS activities was no more variable in sporophytic tissues than that seen in pollen for transgenic plants harbouring the same *lat52-gus-nos3'* fusions (section 3.3.2). This suggests that GUS activity in sporophytic tissues originated from a low level of constitutive *lat*52-promoter activity independent of surrounding sequences in the genome. *lat*52 5' promoter deletion mutants were analysed independent of flanking plant DNA by transient expression analysis in leaves (section 3.3.3). This showed that only the region -41 to -18 (TATA box) was required to support low levels of promoter activity in leaves. Addition of a duplicated *CaMV35S* enhancer directly upstream of each 5' deletion mutant (section 3.3.3) strongly enhanced the low level activity by 20- to -33- fold. Taken together, these data strongly suggest that the *lat*52 promoter region -492 to +110 contains no dominant negative elements but does contain the region -41 to -18 (TATA box) which is sufficient to support a low level of promoter activity in sporophytic tissues.

Previous analysis of *lat*52 promoter 5' deletion mutants in transgenic tomato (Twell *et al.*, 1991) showed that deletion of the promoter from -492 to -71 reduced activity by 1000-fold in anthers. The same deletion (-492 to -71) in transgenic tobacco anthers reduced promoter activity by 641-fold (fig. 3.4). But due to transgene variability it can be assumed that the *lat*52 promoter region -494 to -72 is functionally active to a similar degree in both tomato and tobacco. Transient expression analysis of *lat*52 promoter 5' deletion mutants in tobacco pollen (fig 3.10) showed a similar pattern of promoter activity as previously presented by Twell *et al.*, (1991). Thus confirming that

no deleterious effects were introduced using the improved transient expression assay system presented in this thesis.

Transgenic and transient analysis of *lat*52 5' deletion mutants (figures 3.2 and 3.10) divided the promoter region -492 to -42 into 3 functional pollen-specific regions: a major activator region -492 to -101, a minor activator region -100 to -72 and a minimal activator region -71 to -42. Moreover, developmental analysis of the *lat*52 5' deletion mutants in transgenic tobacco showed that the aforementioned activator regions were actively transcribed at different stages during pollen maturation (fig. 3.4). The organisation of activator regions within the *lat*52 promoter appears to be a general characteristic of pollen-specific promoters. Functional analysis of the promoters of the pollen-specific genes *NTP303* (Weterings *et al.*, 1995), *Zm13* (Hamilton *et al.*, 1992) and *Npg1* (Tebbutt and Lonsdale, 1995) revealed that the promoters contained upstream activator regions which enhanced the activity of a minimal pollen-specific activator region contained within the core promoter.

In summary, the *lat*52 promoter is transcribed at a very low level in sporophytic tissues, but appears pollen-specific as a result of strong positive pollen-specific regulatory sequences contained within the region -492 to -42. In addition, the *lat*52 promoter is comprised of three activator regions which are differentially regulated during pollen development.

3.4.3 The relative activities of *lat*52 promoter activator regions defined by transgenic analysis are dramatically reduced in transient expression assays

Comparison of the fold decrease in *lat*52 promoter expression resulting from deletion of the regions -492 to -101, -100 to -72 and -71 to -42 in transient and transgenic analyses showed a reduced relative activity of each region in the transient expression assay system (fig. 3.10). This quantitative difference could arise from differences in the timing of the utilisation of each region during pollen maturation compared with that measured during pollen germination in the transient assay system.

Another possible explanation is based upon the assumption that a complex of transcription factors is required for maximum activation of the *lat*52 promoter in pollen. If the concentration of one or more transcription factor(s) was limiting in germinating pollen the introduction of an excess of the *lat*52-reporter gene fusions per cell in the transient expression assay could lead to a decrease in the number of fully active transcription factor complexes available per template.

Results presented in section 3.2.3 support the conclusion that the increased quantitative differences between the activities of the *lat*52 5' promoter deletions in

transgenic plants arise from differences in the time of active transcription during development. The transient expression assay therefore provides access to the functional role of *cis*-regulatory elements in germinating pollen, but does not provide a direct assessment of their absolute activity during pollen maturation.

Chapter 4

Multiple co-dependent activator elements regulate pollen-specific transcription

4.1 Introduction

Functional analysis of the *lat*52 promoter presented in chapter 3 showed that pollen-specific transcription was developmentally regulated by the activator regions -492 to -101, -100 to -72 and -71 to -42. This chapter details experiments designed to further elucidate the precise sequences and organisation of *cis*-regulatory elements required for the pollen-specific activation of the *lat*52 promoter. Previous analysis using the transient expression assay system defined an active role in the control of pollen-specificity for the region -100 to -86, which contains the *cis*-regulatory motif PBI (Twell *et al.*, 1991). In addition, mutation of the central GG residues dramatically reduced the expression of the region -100 to +110 in germinating pollen. To define the 5' and 3' boundaries of the *cis*-regulatory element PBI and to identify the sequences of other *cis*-regulatory elements within the region -100 to +110 a series of substitution mutations were introduced throughout the region -99 to -42 of the 5' promoter deletion mutant -100 to +110 and analysed in germinating pollen. Previous gain of function analysis demonstrated that the region -194 to -176 containing PBII was sufficient to confer pollen-specific expression when linked to a heterologous *CaMV35S* core promoter (Twell *et al.*, 1991). This type of analysis was utilised to identify the minimal promoter regions sufficient to enhance the *CaMV35S* core promoter specifically in pollen.

4.2 The *lat*52 promoter contains three autonomous activator domains

As a result of the extended and more systematic analysis of the *lat*52 promoter presented in this thesis it was thought necessary to redefine the functional promoter map proposed by Twell *et al.*, (1991) (fig. 4.1 A) to that presented in figure 4.1 B.

Results presented in chapter 3 strongly suggested that the major sequences controlling pollen-specific transcription reside within the region -492 to -42. The region -492 to -52 was chosen for further analysis and sub-divided into three domains: domain A, -492 to -226, domain B -225 to -101 and domain C -100 to -52. To investigate their individual and combinatorial activities single, mutually exclusive domain pairs and all three linked domains were cloned upstream of the -90 *CaMV35S* promoter.

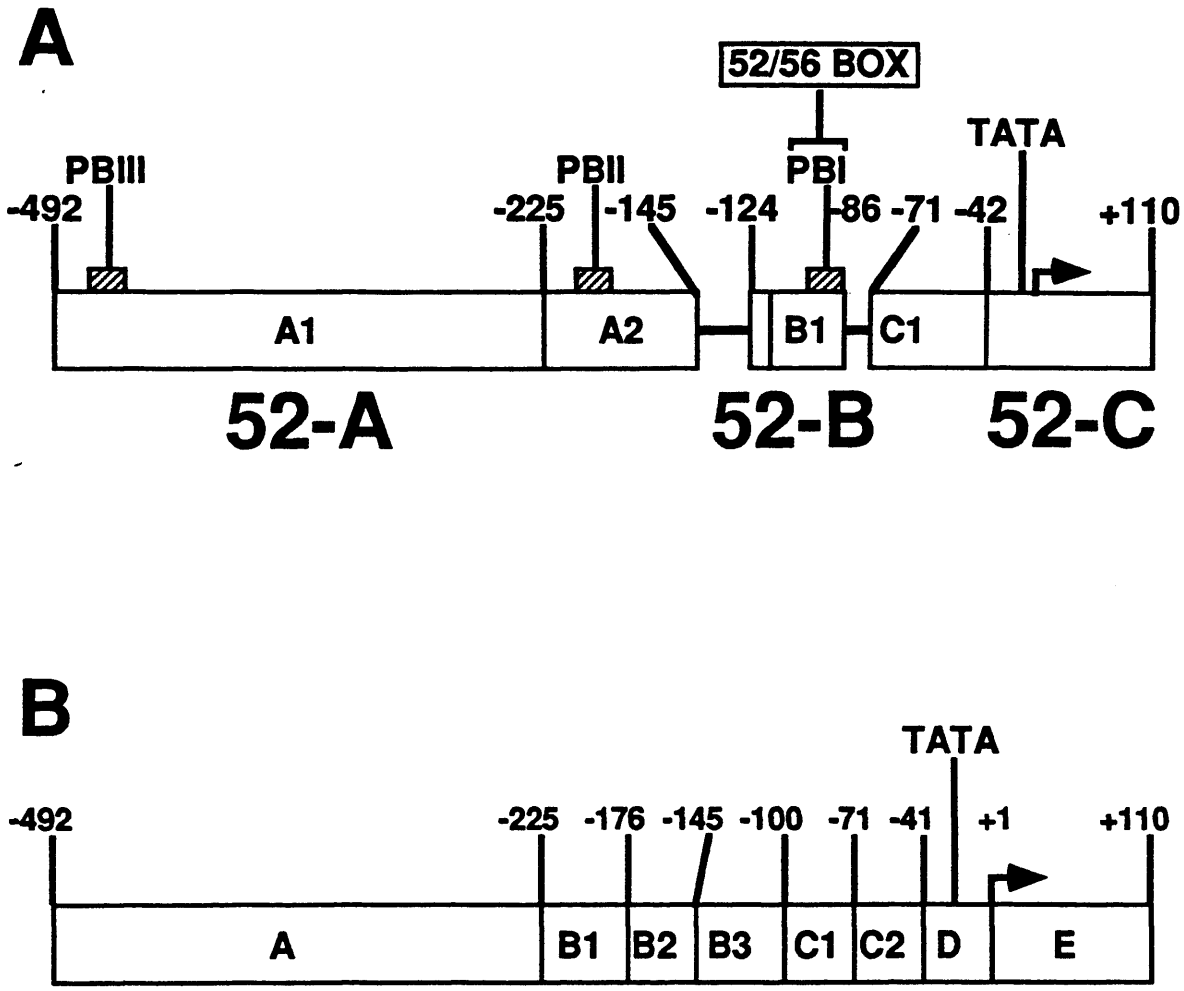


Figure 4.1. A) Functional map of the *lat*52 promoter as defined by Twell *et al.*, 1991. B) Functional map of the *lat*52 promoter redefined as a result of the new transgenic and transient expression analyses presented in this thesis.

4.2.1 Construction of hybrid *lat52/CaMV35S* promoter fusions

The first step in the construction of hybrid *lat52/CaMV35S* promoter fusions was the creation of cassettes which contained either the -45 or -90 5' deletion mutants of the *CaMV35S* promoter fused to the *luc* reporter gene. The steps involved in the construction of the *-45 CaMV35S-TEV-luc-CaMV35S3' (C3')* plasmid fusion is shown schematically in figure 4.2. The *CaMV35S* promoter region -45 to +1 fused to the *TEV 5'-UTR* was amplified from the template pRT2-TEV-LUC (section 3.3.1) using PCR (section 2.5.2) in conjunction with a specific 5' oligonucleotide containing the multiple cloning site *KpnI-HindIII-SalI* and a 3' oligonucleotide complementary to the nonsense strand of *luc*. The generated PCR product was digested with *KpnI and NcoI* and subcloned into the plasmid pNBL-0 (section 3.3.1) displacing the TEV 5'-UTR. The plasmid constructed was termed pNBC-45 LUC.

The steps involved in the construction of the *-90 CaMV35S-TEV-luc-C3'* plasmid fusion is shown schematically in figure 4.3. The *CaMV35S* promoter region -90 to +1 fused to the *TEV 5'-UTR* was amplified from the template pRT2-TEV-LUC (section 3.3.1) using PCR (section 2.5.2) in conjunction with a specific 5' oligonucleotide containing a *SalI* site and a 3' oligonucleotide complementary to the nonsense strand of *luc*. The generated PCR product was digested with *SalI and NcoI* and subcloned into the plasmid pNBC-45 LUC displacing the -45 *CaMV35S-TEV 5'-UTR* fragment. The plasmid constructed was termed pNBC-90 LUC.

The steps involved in the construction of hybrid *lat52/CaMV35S* promoter fusions are shown schematically in figure 4.4. The *lat*52 promoter fragments -492 to -52, -492 to -101, -492 to -226, -225 to -52 and -225 to -101 were amplified from the template pNBL52-5 (section 3.3.2) using PCR (section 2.5.2) in conjunction with specific 5' oligonucleotides each containing a *KpnI* site at the relevant *lat*52 5' border sequence and specific 3' oligonucleotides each containing a *SalI* site at the relevant *lat*52 3' border sequence. The generated PCR products were digested with *KpnI* and *SalI* and subcloned into the plasmid pNBC-90 LUC. The plasmids constructed were termed pNBL52-113 (-492 to -52), -101 (-492 to -101), -82 (-492 to -226), -114 (-225 to -52), -81 (-225 to -101). Construction of the plasmid pNBL52-41 containing the *lat*52 promoter region -100 to -52 fused to -90 *CaMV35S* is discussed in section 4.4.3.

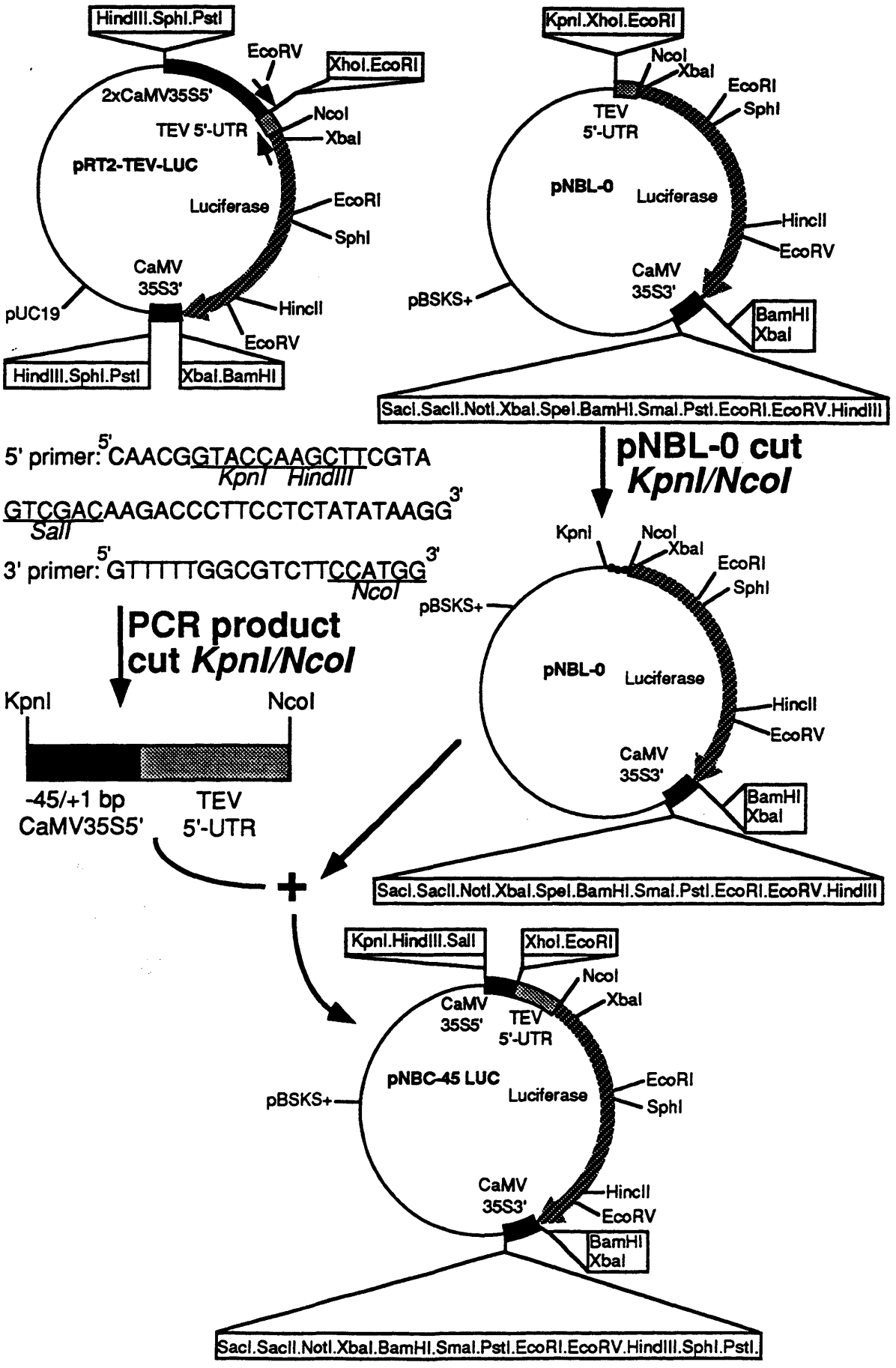


Figure 4.2. Construction of the *-45 CaMV35S-TEV-luc-C3'* plasmid fusion. The *CaMV35S* promoter region -45 to +1 fused to the *TEV 5'-UTR* was amplified from the template pRT2 TEV-LUC (section 3.3.1) and cloned into the plasmid pNBL-0 (section 3.3.1) as a *KpnI/NcoI* fragment.

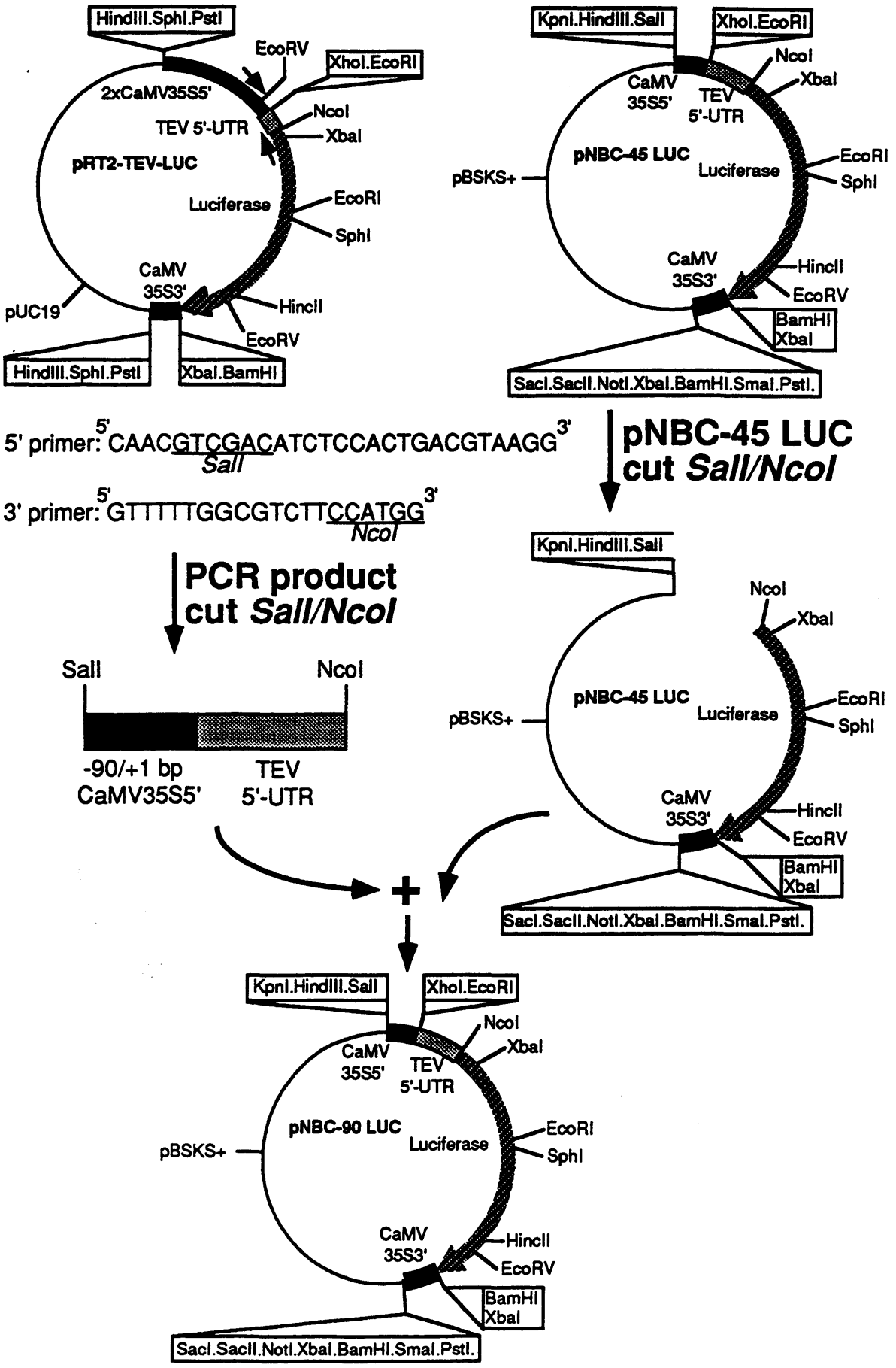


Figure 4.3. Construction of the *-90 CaMV35S-TEV-luc-C3'* plasmid fusion. The *CaMV35S* promoter region -90 to +1 fused to the *TEV 5'-UTR* was amplified from the template pRT2 TEV-LUC (section 3.3.1) and cloned into the plasmid pNBC-45 LUC (section 4.2.1) as a *SalI/NcoI* fragment displacing the -45 *CaMV35S-TEV 5'-UTR* fragment.

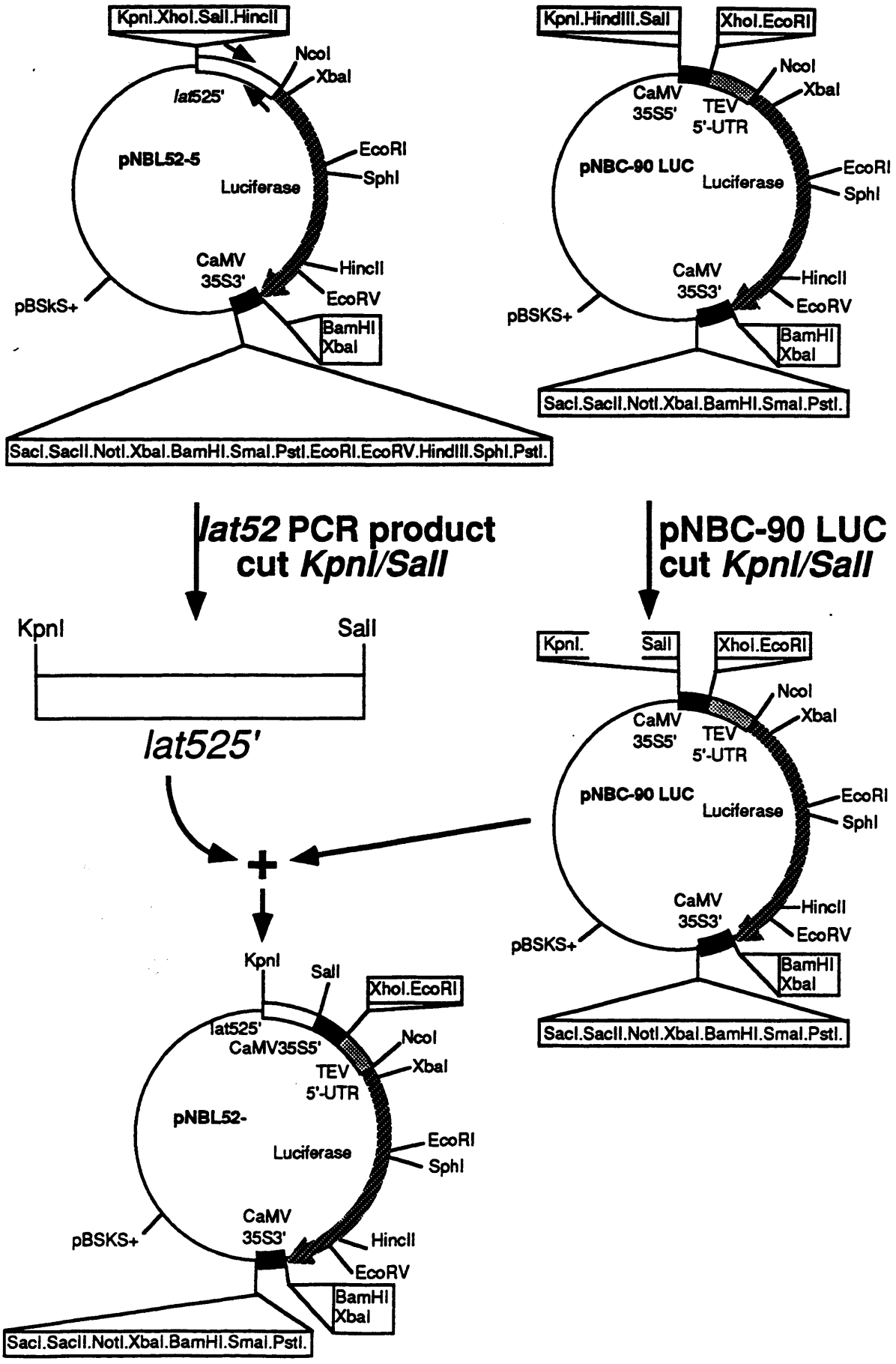


Figure 4.4. Construction of hybrid *lat52/CaMV35S* promoter fusions. The *lat52* promoter fragments -492 to -52, -492 to -101, -492 to -226, -225 to -52 and -225 to -101 were amplified from the template pNBL52-5 (section 3.3.2) and cloned into the plasmid pNBC-90 LUC (section 4.2.1) as *KpnI/SalI* fragments.

4.2.2 Pollen-specific activation of the *CaMV35S* core promoter by *lat52* promoter domains

The test plasmids pNBL52-113, -101, -82, -114, -81, -41 and pNBC-90 LUC were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5) and leaves (section 2.9.6). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the - 90 *CaMV35S* 5' promoter deletion mutant. The results of this analysis are presented in figure 4.5.

In pollen domains A+B+C (-492 to -52) together enhanced the activity of -90 *CaMV35S* by 24-fold. Removal of domain C leaving domains A+B (-492 to -101) dramatically decreased the level of enhancement from 24- to 8-fold. In contrast, removal of domain A showed that domains B+C (-225 to -52) together were sufficient to enhance -90 *CaMV35S* to similar levels (26-fold) seen for domains A+B+C (-492 to -52) together (24-fold). The individual domains A (-492 to -226), B (-225 to -101) and C (-100 to -52) enhanced -90 *CaMV35S* by 6-, 4- and 11-fold respectively. In leaves domains A, B or C individually or in combination did not significantly enhance -90 *CaMV35S*. These analyses demonstrate that domains A (-492 to -226), B (-225 to -101) and C (-100 to -52) alone are sufficient to enhance transcription in a pollen-specific manner. They also suggest that in germinating pollen domain A (-492 to -226) is transcriptionally redundant when both domains B+C (-225 to -52) are present.

4.3 Functional analysis of the activator domain B

4.3.1 Dissection of domain B reveals a redundant sub-domain B1 and two new novel activator sub-domains, B2 and B3

Data shown in figure 4.5 indicated that domain B, -225 to -101 contains functional sequences which regulate pollen-specificity of the *lat52* promoter. Previous analysis showed that the region -194 to -176 containing PBII activated a truncated *CaMV35S* promoter in a pollen-specific manner (Twell *et al.*, 1991). To confirm the proposed active role for the PBII motif, and to define the location of any other *cis*-regulatory elements which contribute to pollen-specificity within domain B, targeted 5' deletion and substitution mutational analyses were performed on the region -225 to -101.

The domain B promoter fragments -145 to -101, -176 to -101, -194 to -101 and -194 to -101 containing a 4 bp substitution mutation of the PBII motif were amplified and fused directly upstream of the -90 *CaMV35S* promoter as described for the construction of the plasmid pNBL52-81 (section 4.2.1). The plasmids were termed pNBL52-77 (-145

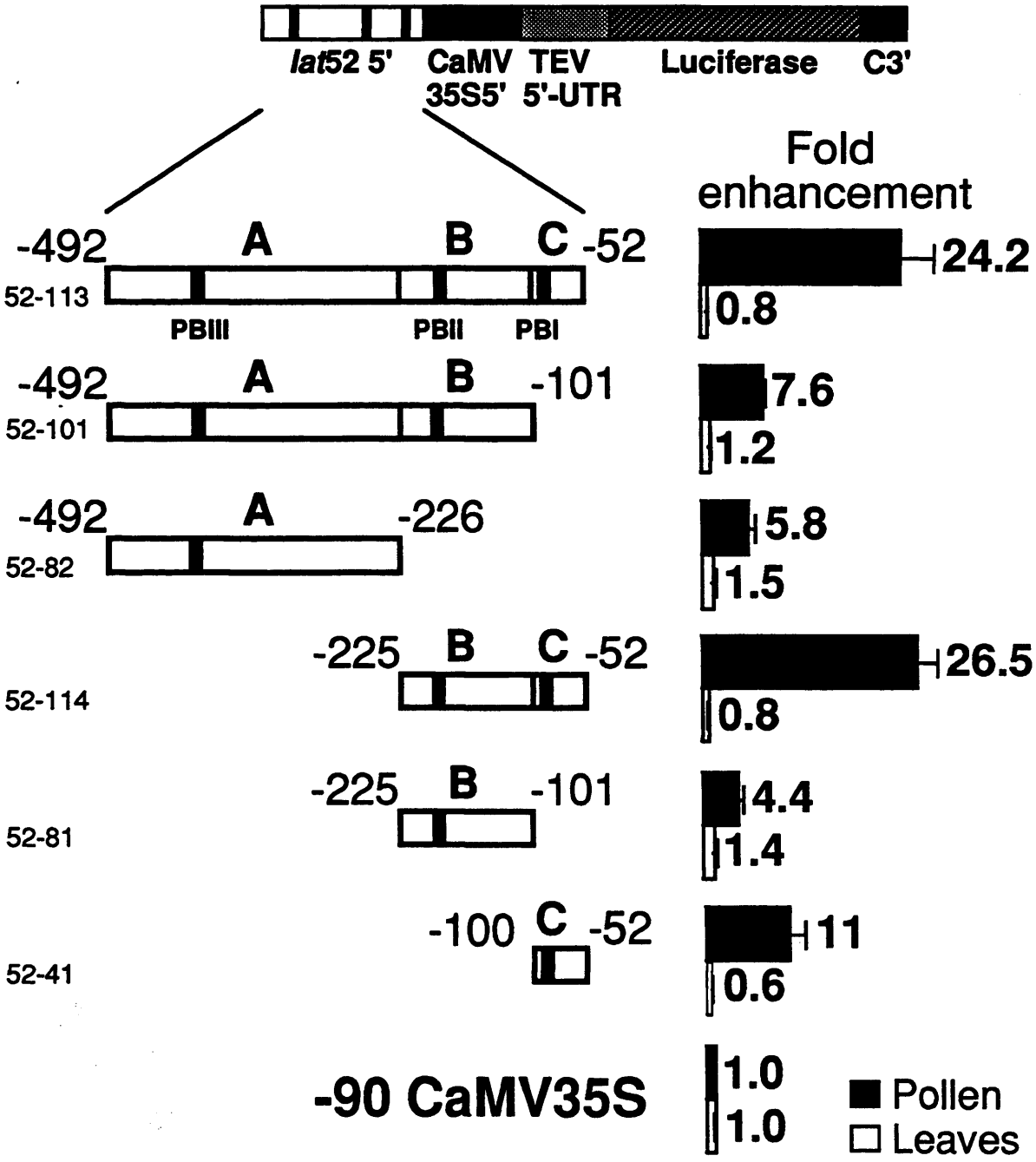


Figure 4.5. Activation of the -90 *CaMV35S* core promoter by *lat*52 promoter fragments. The relative activities of test plasmids containing chimeric *lat52/-90 CaMV35S5'-TEV-luc-C3'* gene fusions were determined in pollen and leaves in the transient expression assay. pLAT52-7 was used as the reference plasmid for pollen, and pRTL2GUS was used as the reference plasmid for leaves. The fold enhancement of each test plasmid was calculated from the ratio of LUC/GUS normalised to that of the *-90 CaMV35S-TEV-luc-C3'* control (average light units per bombardment was 1841597 and 163598 in pollen and leaves respectively). Data from two independent experiments are shown, representing six independent bombardments for each test plasmid. Error bars represent the standard error.

to -101), -78 (-176 to -101), -79 (-194 to -101) and -80 (-194 to 101, with 4 bp PBII mutation).

The test plasmids pNBL52-77, -78, -79, -80 , -81 and pNBC-90 LUC were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5) and leaves (section 2.9.6). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the - 90 *CaMV35S* 5' promoter deletion mutant. The results of this analysis are presented in figure 4.6.

Progressive 5' deletion of domain B from -225 to -194 and -194 to -176 (deleting PBII) did not significantly reduce the enhancement of -90 *CaMV35S* by domain B in pollen. Mutation of the central 4 bp core motif of PBII from CCAC to AATA also had no effect on pollen-specific enhancement. Further deletion of domain B from -176 to -145 reduced enhancement of -90 *CaMV35S* from 4-fold to 2.5-fold in pollen. Domain B promoter fragments did not enhance -90 *CaMV35S* in leaves. This analysis divided domain B of the *lat*52 promoter into three sub-domains: a redundant sub-domain B1 (-225 to -177), containing PBII, and two new novel sub-domains; B2 (-176 to -146) and B3 (-145 to -101).

4.3.2 Sub-domains B2 and B3 are essential for full functional activity of the *lat*52 promoter region -492 to +110

5' deletion analysis of the *lat*52 promoter in the transient expression assay showed that the region -492 to -101 (domains:A+B) contributed ~ 60 % of the total promoter activity in germinating pollen (fig. 3.10). To determine whether the domains identified using the gain of function approach above (fig. 4.6) directly contributed to the 60 % of *lat*52 promoter activity, each region was progressively removed.

The domain B 5' deletion mutants -225, -194, -176 and -145 to +110 were constructed as follows. The 5' deletion mutants -225 to +110 and -145 to +110 were excised from *lat*52 promoter-*gus-nos3'* gene fusions (Twell *et al.*, 1991) and fused to the *luc* coding region as described for the construction of the plasmid pNBL52-5 (section 3.3.2). The 5' deletion mutants -176 to +110 and -194 to +110 of the *lat*52 promoter were amplified by PCR (section 2.5.2) and fused to the *luc* coding region in the same way as described for the construction of the plasmid pNBL52-10 (section 3.3.2). The plasmids were termed pNBL52-3 (-145), -4 (-225), -89 (-176) and -90 (-194).

The test plasmids pNBL52-2, -3, -4, -5, -89 and -90 were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the -492 to +110 *lat*52 5' promoter deletion mutant. The results of this analysis are presented in figure 4.7.

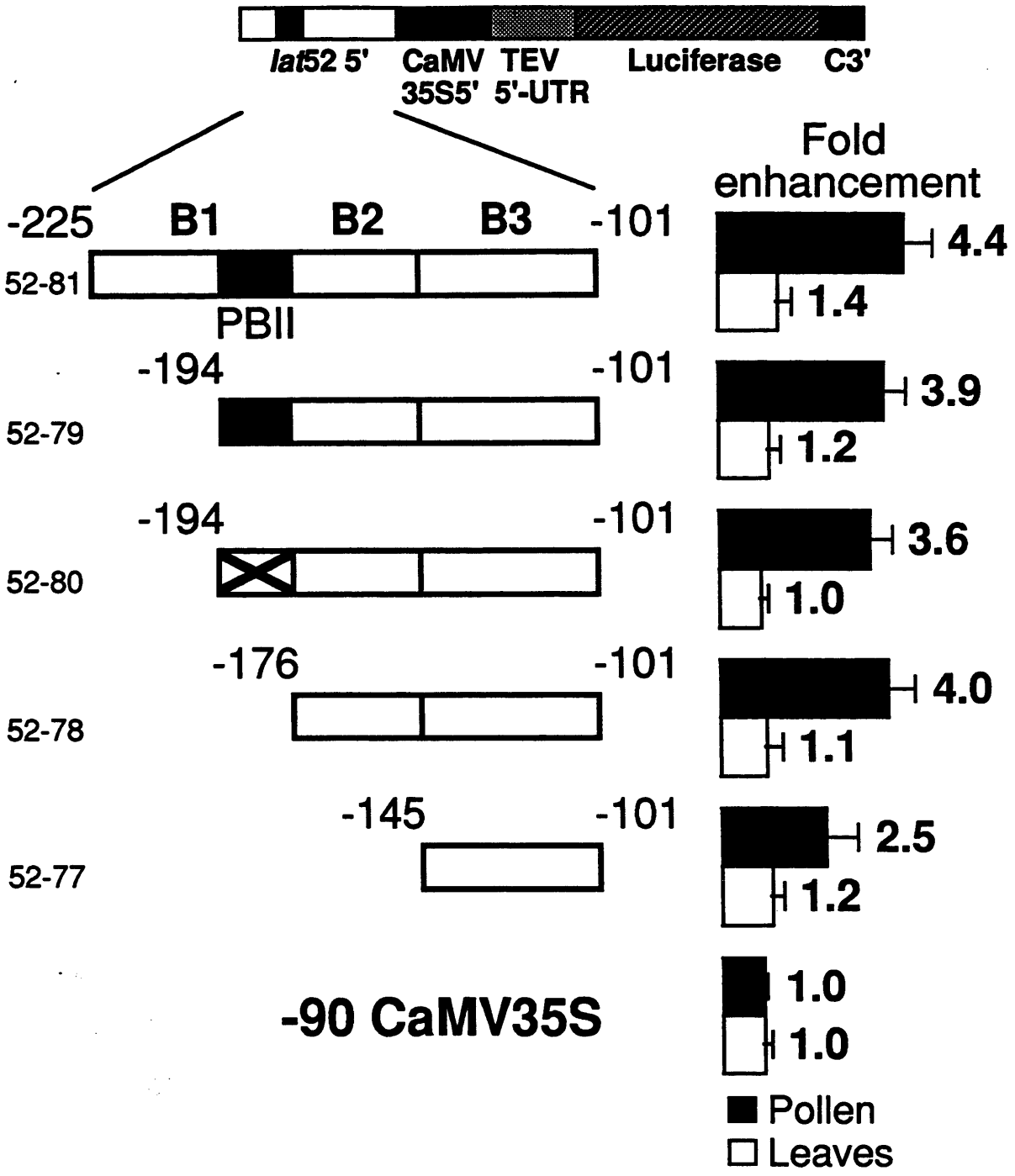


Figure 4.6. Targeted mutagenesis of the *cis*-regulatory element PBII and 5' deletion analysis of the *lat*52 promoter domain B. The relative activities of test plasmids containing chimeric *lat52/ 90 CaMV35S5'-TEV-luc-C3'* gene fusions were determined in pollen and leaves in the transient expression assay. pLAT52-7 was used as the reference plasmid for pollen, and pRTL2GUS was used as the reference plasmid for leaves. The fold enhancement of each test plasmid was calculated from the ratio of LUC/GUS normalised to that of the *-90 CaMV35S-TEV-luc-C3'* control (average light units per bombardment was 5690233 and 70574 in pollen and leaves respectively). Data from two independent experiments are shown, representing six independent bombardments for each test plasmid. Error bars represent the standard error.

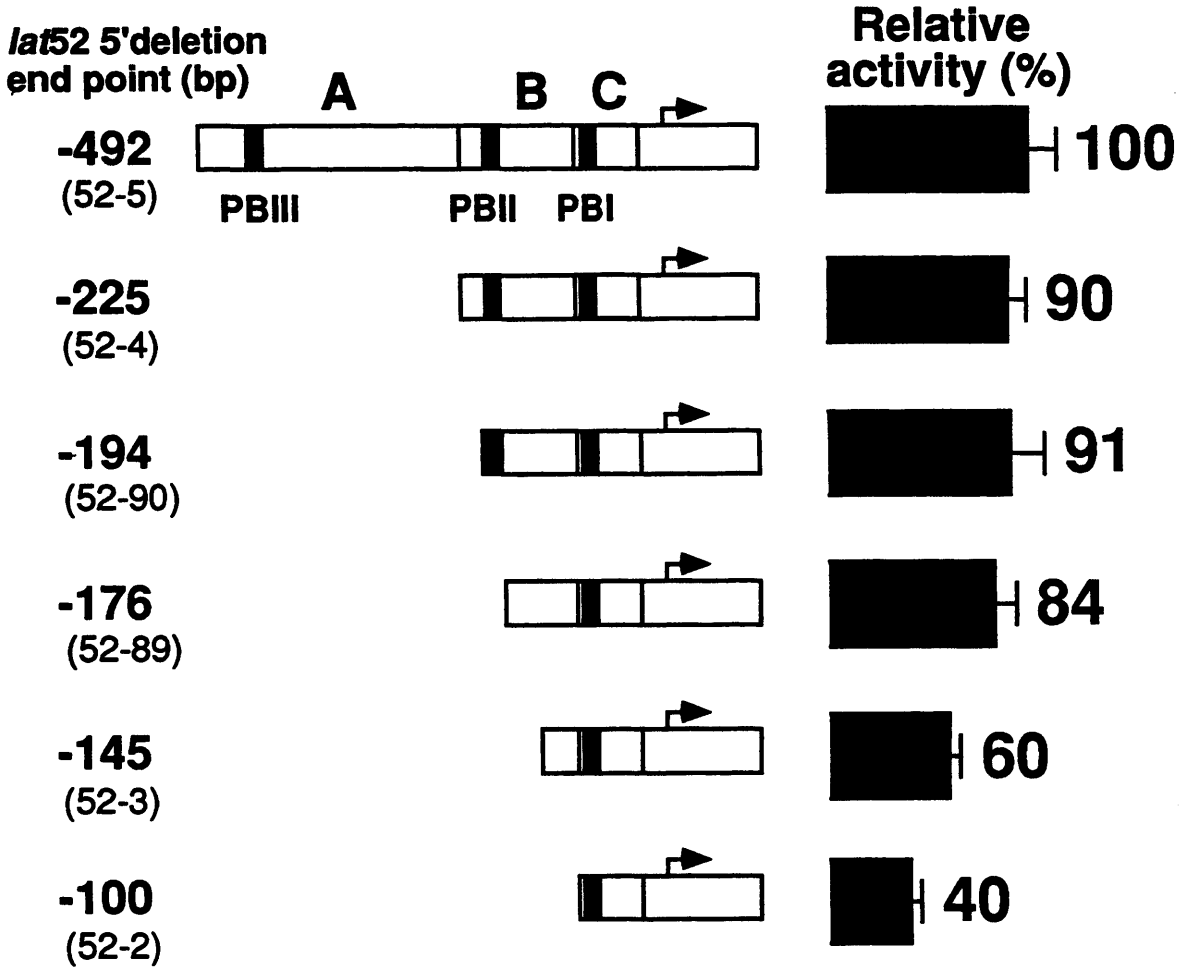


Figure 4.7. Transient expression analysis of *lat*52 5' promoter deletion mutants. Test plasmids containing *lat*52 5' promoter deletion-*luc-C3'* gene fusions were co-bombarded into pollen with the reference plasmid pLAT52-7. Relative activities of test plasmids containing *lat52-luc-C3'* fusions were calculated from a ratio of LUC:GUS normalised to the activity of the largest 5' deletion mutant (-492 to +110; average light units per bombardment was 53099556). Data from two independent experiments are shown, representing six independent bombardments for each plasmid. Error bars represent the standard error.

Deletion of domain A (-492 to -226), or domain B1 (-225 to -177) did not significantly reduce promoter activity. However, further deletion of domain B2 (-176 to -146) and domain B3 (-145 to -101) progressively decreased promoter activity by 24 % and 20 % respectively. These results confirm an active role for the sub-domains B2 and B3 in the pollen-specific regulation of the *lat*52 promoter and further confirm redundancy in the function of domain A and sub-domain B1 (PBII) when linked to downstream domains in germinating pollen.

4.4 Functional analysis of the activator domain C

4.4.1 Mutational analysis of the *cis*-regulatory element PBI reveals a central GTGG motif as essential for wild type activity

5' deletion analysis in transgenic tobacco pollen showed that deletion of the region -100 to -72 reduced promoter activity by ~ 15-fold (fig. 3.2), highlighting the importance of this region in the control of pollen-specific transcription. This region contains the sequence motif PBI (TGTGGTT). The PBI motif is flanked by AT-rich sequences, TTATTGTGGTTAT (Twell *et al.*, 1989b) such that the 3' half differs by only 1 bp from the box II core motif (GGTTAA) of the light responsive pea *rbcS-3A* gene (Green *et al.*, 1988). The box II core motif GGTTAA has been defined as critical for the binding of the transcription factor GT-1a (Gilmartin *et al.*, 1992; Perisic and Lam, 1992). To define the 5' and 3' boundaries of the *cis*-regulatory element PBI and to functionally test if PBI contains the same critical consensus sequence as box II, two base pair linker scan mutations were introduced throughout the region -99 to -88.

The steps involved in the construction of the two base pair linker scan mutations introduced throughout the region -99 to -88 are shown schematically in figure 4.8. The *lat*52 promoter region -100 to +110 was amplified from the template pNBL52-2 (section 3.3.2) using PCR (section 2.5.2) in conjunction with specific 5' degenerate oligonucleotides (containing a *KpnI* site and the desired two base pair linker scan mutation) and a 3' oligonucleotide complementary to the nonsense strand of *luc*. The generated PCR product was digested with *KpnI and NcoI* and subcloned into the plasmid pNBL52-5 (section 3.3.2) displacing the *lat*52 promoter region -492 to +110. The 5' deletion -100 to +110 carrying the GG to CA mutation at positions -93 and -92 respectively, was excised from a *lat*52 promoter-*gus-nos3'* gene fusion (Twell *et al.*, 1991) and fused to the *luc* coding region as described for the construction of the plasmid pNBL52-5 (section 3.3.2). The plasmids were termed pNBL52-19 (TT to GG: -99,-98), -20 (AT to CG: -97, -96), -21 (GT to TA: -95, -94), - 14 (GG to CA: -93, -92), -22 (TT to GG: -91, -90) and -23 (AT to CG: -89, -88).

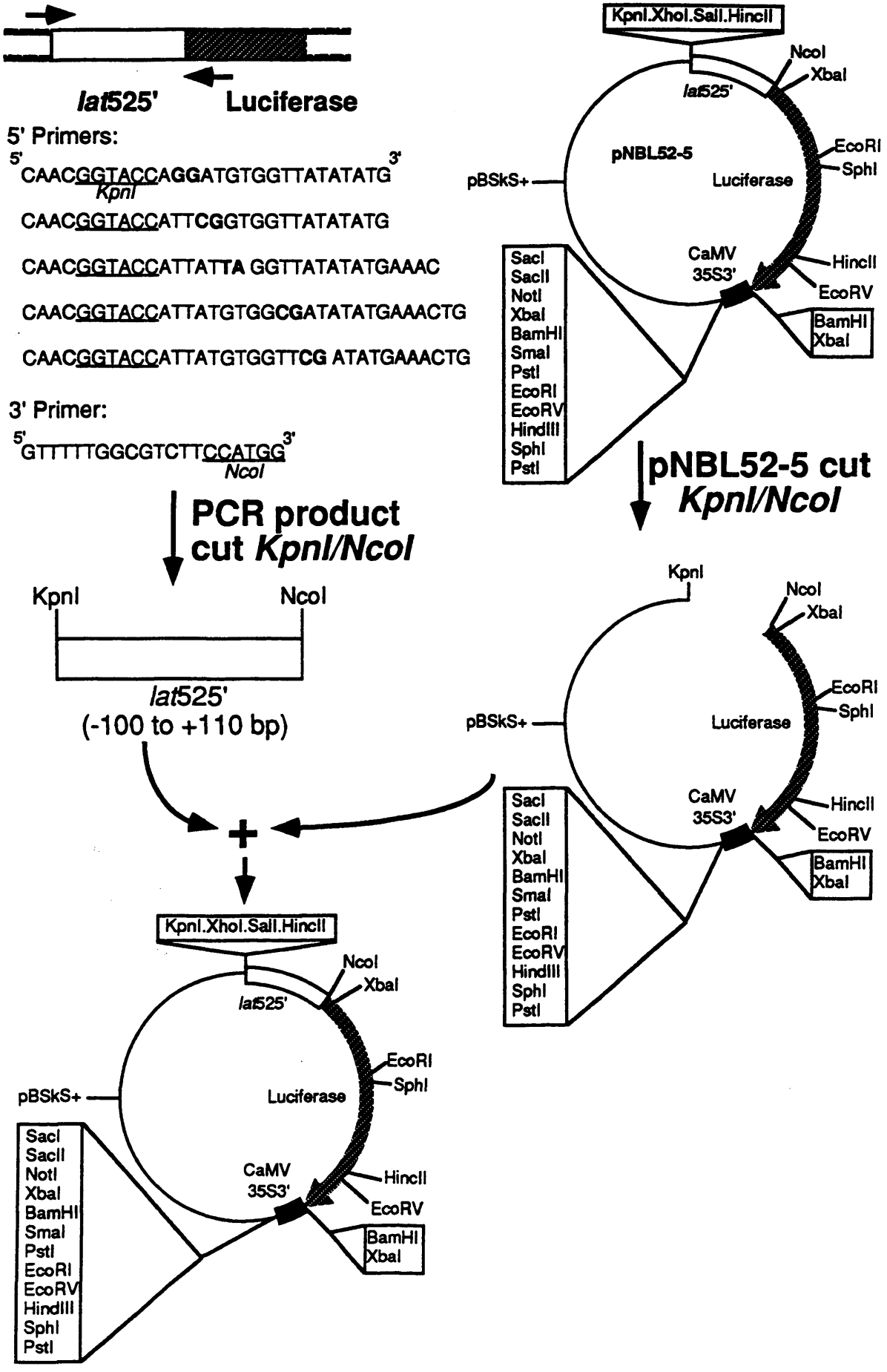


Figure 4.8. Construction of the two base pair linker scans throughout the region -99 to -88 (PBI). A series of 2 bp mutations were introduced into the *lat*52 promoter region -100 to +110 by PCR, and cloned into the plasmid pNBL52-5 (section 3.3.2) as *KpnI/NcoI* fragments.

The test plasmids pNBL52-2, -14, -19, -20, -21, -22 and -23 were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the wild type -100 to +110 *lat*52 5' promoter deletion mutant. The results of this analysis are presented in figure 4.9.

Mutation of the central GG residues of PBI reduced expression 10-fold, whilst only a 2-fold reduction was observed when the adjacent 5' GT residues were mutated. In contrast, mutation of the flanking pairs of nucleotides, TT and AT, either side of the central GTGG core resulted in a 2- to 3-fold stimulation of LUC activity. If it can be assumed that a mutation within a *cis*-element which prevents a protein binding *in vitro* would also inactivate the *cis*-element *in vivo*, then these results suggest that the *cis*-regulatory element PBI contains a different functional core sequence to that observed for box II (Green *et al.*, 1988). Since the only functional similarity between the two *cis*-regulatory elements is the mutation of the two central GG residues, mutation either side of GG gave different results for both boxes. In summary, the pollen-specific regulatory element PBI contains a core GTGG motif required for wild type levels of activity, flanked on either side by TTAT which appear to negatively regulate the activity of PBI.

4.4.2 The region -84 to -41 contains two novel *cis*-regulatory sequences

The results presented in figure 3.2 showed that deletion of the *lat*52 promoter from -71 to -41 resulted in a 3-fold drop in GUS activity in pollen of transgenic plants. In the transient expression assay in pollen deletion of this region led to only a 1.5-fold drop in LUC activity (fig. 3.10). Analysis of additional 5' deletion mutants within the region -71 to -41 highlighted the presence of at least two other *cis*-regulatory sequences which appeared to regulate *lat*52 promoter activity (fig. 3.10). Therefore, to confirm the position of these pollen-specific *cis*-regulatory elements within the region -71 to -42 and to investigate their possible interactions with upstream proximal sequences, substitution mutations (SM1 to 7) were introduced throughout the region -84 to -42 of the 5' promoter deletion mutant -100 to +110.

The block substitution mutants SM1 to 7 fused to the *gus* coding region were constructed by D. Twell. The *lat*52 promoter region containing the block substitution mutations were excised from the *lat*52 promoter-*gus-nos3'* gene fusions and fused to the *luc* coding region as described for the construction of the plasmid pNBL52-5 (section 3.3.2). Plasmids constructed were termed pNBL52-12 (SM3), -13 (SM7), -15 (SM1), -16 (SM6), -17 (SM2), -18 (SM5) and -61 (SM4).

The test plasmids pNBL52-2, -12, -13, -15, -16, -17, -18 and -61 were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in

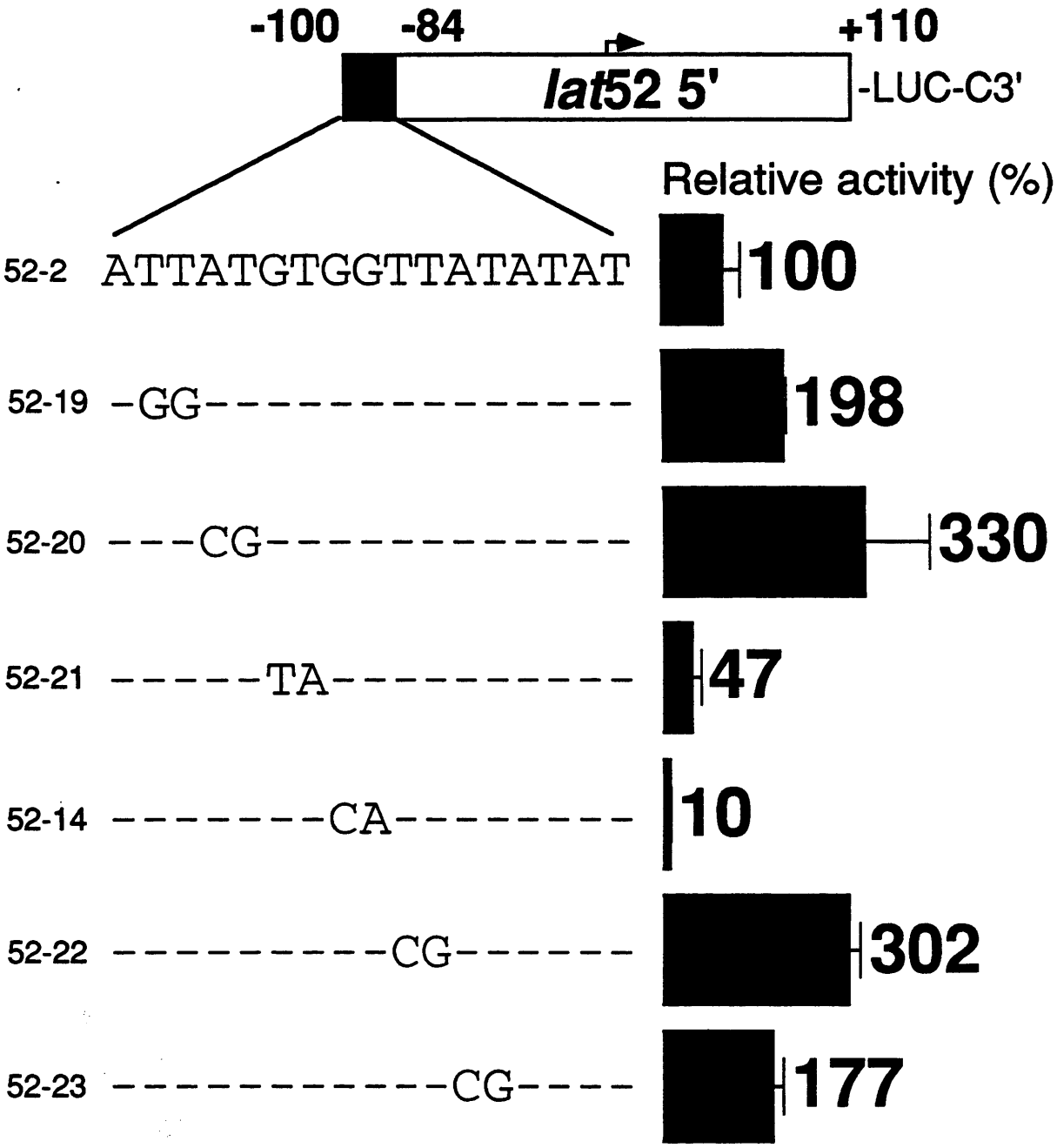


Figure 4.9. Two base pair linker scan analysis of the pollen-specific *cis*-regulatory element PBI. Test plasmids containing single 2 bp linker scan mutations throughout the region -99 to 88 of the -100 to +110 *lat*52 5' promoter deletion-*luc-C3'* gene fusion were co-bombarded into pollen with the reference plasmid pLAT52-7. Relative activities of test plasmids containing mutant *lat52-luc-C3'* fusions were calculated from a ratio of LUC:GUS normalised to the activity of the wild type -100 to +110 *lat*52 5' promoter deletion-*luc-C3'* gene fusion (average light units per bombardment was 55592667). Data from two independent experiments are shown, representing six independent bombardments for each plasmid. Error bars represent the standard error.

germinating pollen (section 2.9.5). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the wild type -100 to +110 *lat*52 5' promoter deletion mutant. The results of this analysis are presented in figure 4.10.

Substitution mutants SM1, 2, 3, and 5 decreased WT (-100) promoter activity by 2-fold or less, where as SM7 slightly increased activity. The most dramatic effects were observed for SM4 and SM6 which decreased promoter activity by 10- and 4-fold respectively. These analyses confirmed the location of two new functional *cis*-regulatory sequences AGAAA (-72 to -68) and TCCACCATA (-60 to -52), both of which are flanked by sequences which positively influence promoter activity but to a lesser extent.

4.4.3 Domain C of the *lat*52 promoter activates the minimal *CaMV35S* promoter independent of distance

Functional analysis of domain C (-100 to -52) showed that the *cis*-regulatory sequences GTGG (PBI) (fig. 4.9), AGAAA (-72 to -68) and TCCACCATA (-60 to -52) (fig. 4.10) were essential for full activity of the *lat*52 promoter region -100 to +110. In addition, data presented in figure 4.5 showed that domain C activated -90 *CaMV35S* by 11-fold. The activating properties of domain C were further investigated by its addition directly upstream of the -45 5' deletion mutant of the *CaMV35S* promoter, thus allowing it to be tested in its original position with respect to the TATA box. Also the PBI core sequence GTGG defined as essential in section 4.4.1 was mutated to TACA (MPBI) in both the domain C/-90 and /-45 *CaMV35S* hybrid promoter fusions.

The steps involved in the construction of the above hybrid *domain C/CaMV35S* promoter fusions are shown schematically in figure 4.11. Domain C fragments containing wild type or mutated PBI sequences were inserted directly upstream of either the -45 or -90 *CaMV35S* promoters essentially as described for the construction of the plasmid pNBL52-113 (section 4.2.1). Except that the *lat*52 promoter region -100 to -52 was obtained in the following way. The *lat*52 promoter region -100 to +110 was amplified from the template pNBL52-13 (section 4.4.2) using PCR (section 2.5.2) in conjunction with specific 5' oligonucleotides containing a 5' *KpnI* site, with and without the 4 bp PBI mutation, and a 3' oligonucleotide complementary to the nonsense strand of *luc*. The generated PCR products were digested with *KpnI* and *SalI* and the *lat*52 promoter fragment -100 to -52 was cloned upstream of -45 and -90 *CaMV35S* promoters. The plasmids constructed were termed pNBL52-40 (-100 to -52/ -45 *CaMV35S*), -41 (-100 to -52/ -90 *CaMV35S*), -59 (MPBI: -100 to -52/ -45 *CaMV35S*) and -60 (MPBI: -100 to -52/ -90 *CaMV35S*).

The test plasmids pNBL52-40, -41, -59, -60, pNBC-90 LUC and pNBC-45 LUC were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently

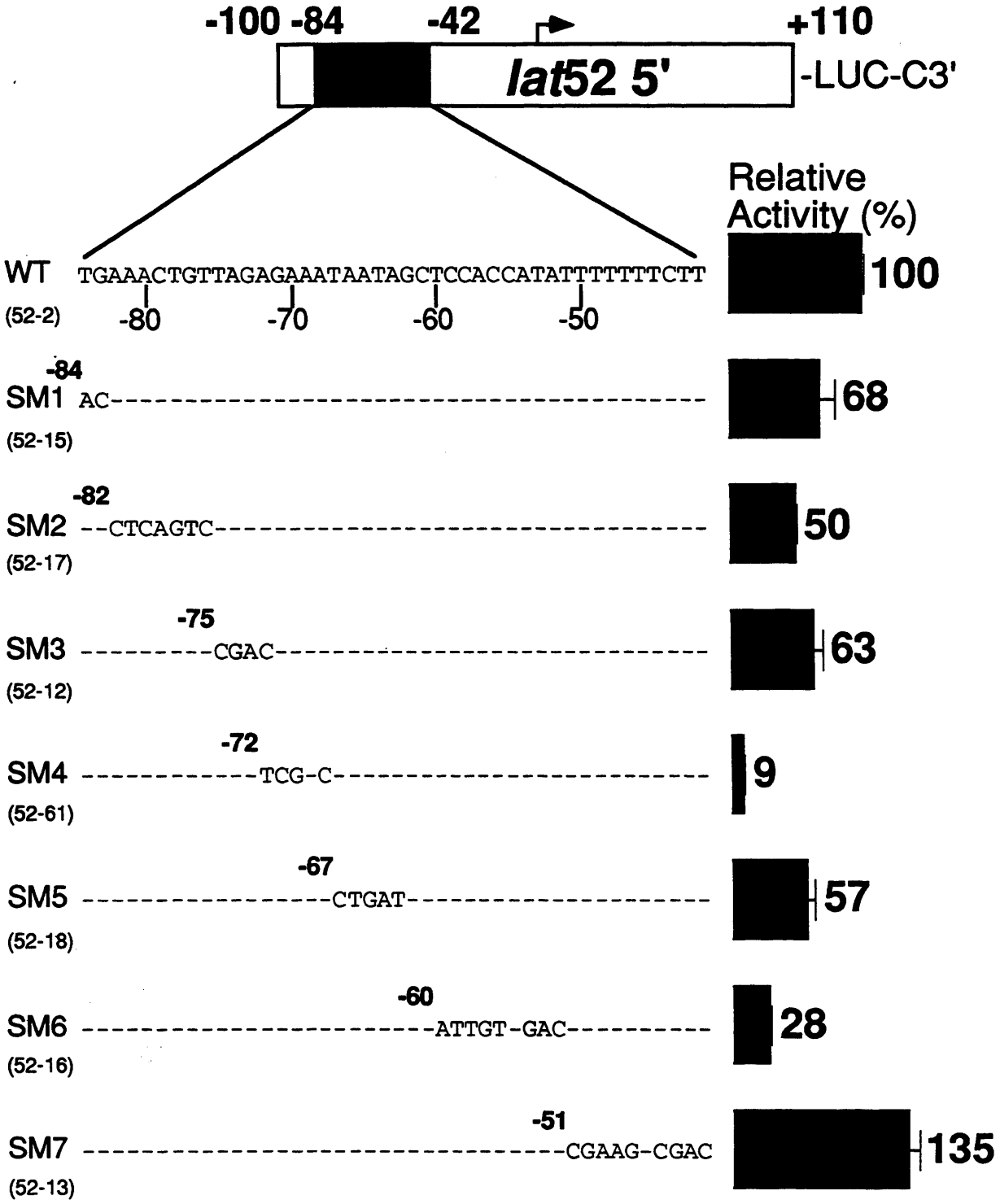


Figure 4.10. Targeted substitution mutagenesis of domain C of the *lat*52 promoter. Test plasmids containing single block substitution mutations throughout the region -84 to -42 of the 100 to +110 *lat*52 5' promoter deletion-*luc-C3'* gene fusion were co-bombarded into pollen with the reference plasmid pLAT52-7. Relative activities of test plasmids containing mutant *lat52-luc-C3'* fusions were calculated from a ratio of LUC:GUS normalised to the activity of the wild type -100 to +110 *lat*52 5' promoter deletion-*luc-C3'* gene fusion (average light units per bombardment was 72675625). Data from two independent experiments are shown, representing six independent bombardments for each plasmid. Error bars represent the standard error.

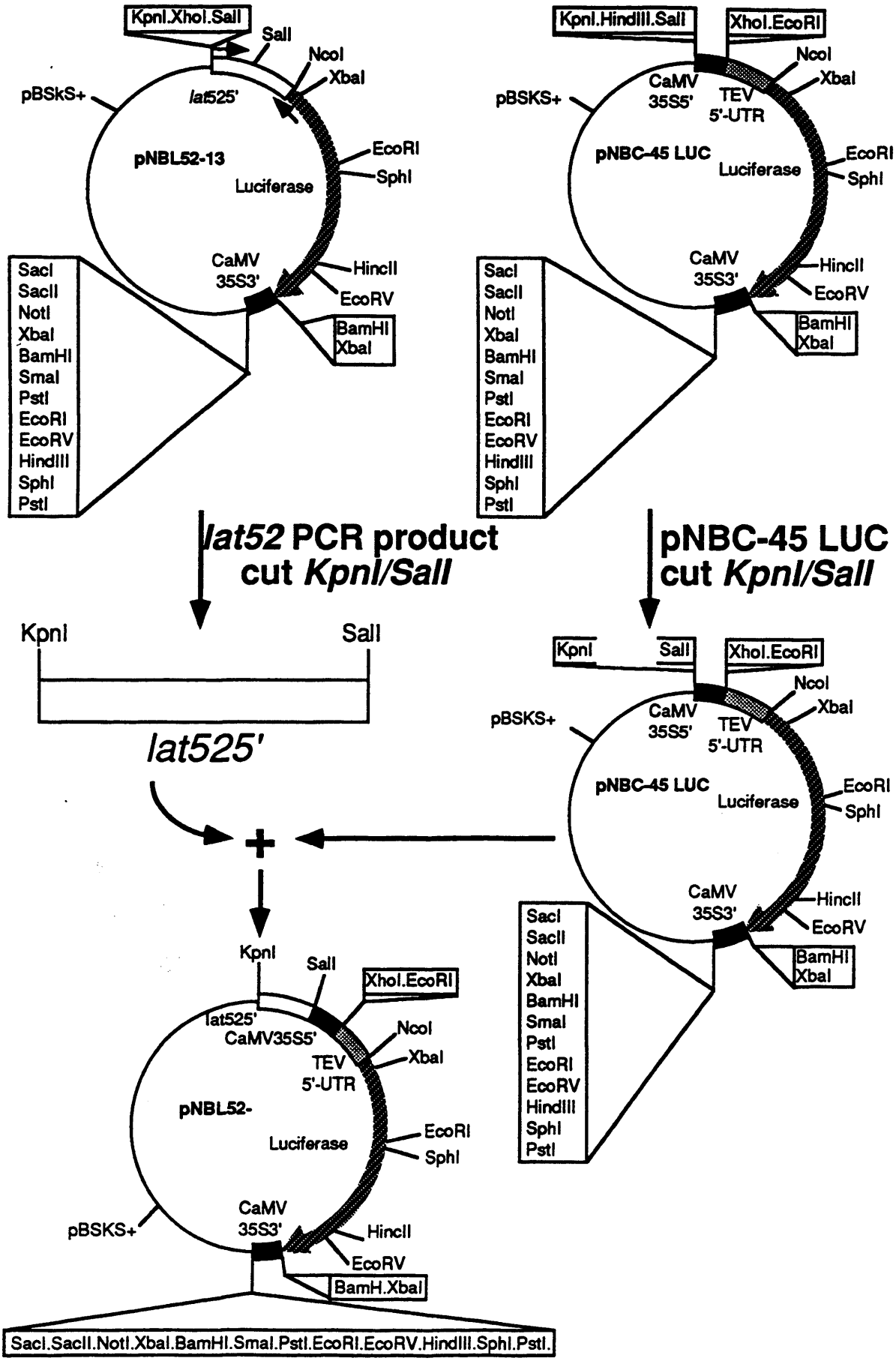


Figure 4.11. Construction of hybrid *domain C/CaMV35S* promoter fusions. Domain C fragments containing wild type or mutated PBI sequences were amplified from the template pNBL52-13 (section 4.4.2) and cloned into the plasmid pNBC-45 LUC as KpnI/SalI fragments.

expressed in germinating pollen (section 2.9.5) and leaves (section 2.9.6). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the -45 *CaMV35S* 5' promoter deletion mutant. The results of this analysis are presented in figure 4.12.

In pollen domain C activated both -45 and -90 *CaMV35S* by 12.5- and 11-fold respectively. Mutation of PBI resulted in a 2- to 3-fold decrease in activity. In leaves neither the wild type or mutated domain C constructs affected the expression of the -45 *CaMV35S* promoter, but addition onto -90 *CaMV35S* resulted in a 2-fold decrease in activity. In summary domain C activated a heterologous core promoter independent of distance in a pollen-specific manner. Since mutation of PBI did not abolish the activating properties of domain C this further suggests a functional role for the *cis*-regulatory sequences AGAAA and TCCACCATA.

4.4.4 The region -71 to -52 of domain C is the minimal pollen-specific functional unit

Pollen-specific activation of the -45 *CaMV35S* promoter by domain C was reduced but not abolished by mutation of PBI (fig. 4.12). Domain C also contains the *cis*-regulatory sequences AGAAA and TCCACCATA. To determine if these sequences were responsible for the residual activity of domain C, and to further investigate the interdependence of *cis*-regulatory sequences within domain C, deletion and substitution mutants of domain C were fused to the -45 *CaMV35S* core promoter.

The domain C fragments -100 to -72 and -71 to -51 were inserted directly upstream of either the -45 or -90 *CaMV35S* promoters as described for the construction of the plasmid pNBL52-113 (section 4.2.1). The steps involved in the fusion of the *lat*52 promoter fragment -61 to -52 upstream of the -45 *CaMV35S* promoter is shown schematically in figure 4.13. Oligonucleotides were designed to the region -61 to -52 of the *lat*52 promoter such that when they were annealed a *KpnI* overhang was created at the 5' end and a *SalI* overhang was created at the 3' end. In addition, the distance between the promoter fragment was maintained with respect to the TATA box. The double stranded oligonucleotide was then inserted upstream of the -45 *CaMV35S* promoter as described for the plasmid pNBL52-40. This method was also used to insert the *lat*52 promoter fragments -86 to -52 and -71 to -62 upstream of the -45 *CaMV35S* promoter. The plasmids constructed were termed pNBL52-85 (-86 to -52/ -45 *CaMV35S*), -112 (-71 to -52/ -45 CaMV35S), -84 (-61 to -52/ -45 *CaMV35S*), -121 (-71 to -61/ -45 CaMV35S), -111 (-100 to -72/ -45 *CaMV35S*) and -98 (-100 to -72/ -90 *CaMV35S*).

The test plasmids pNBL52-40, -84, -85, -98, -111, -112, -121, pNBC-90 LUC and pNBC-45 LUC were isolated (section 2.4.1), equalised to each other (section 2.9.1)

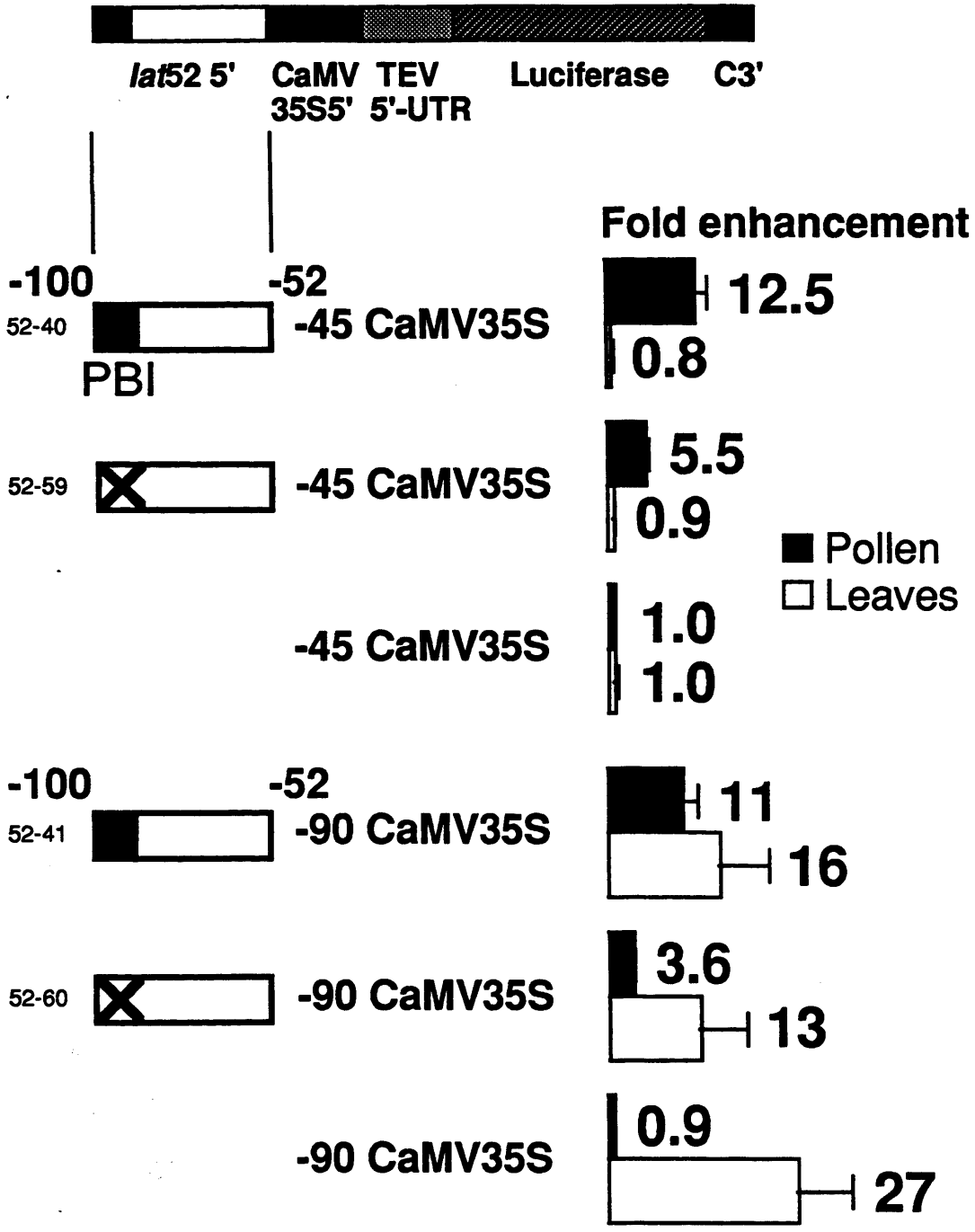


Figure 4.12. Pollen-specific activation of the -45 and -90 *CaMV35S* core promoters by *lat*52 promoter domain C. *lat*52 promoter fragments containing domain C (-100 to -52) or domain C with the *cis*-regulatory element PBI mutated (black cross) were fused upstream of the -45 and 90 *CaMV35S* minimal promoters. The relative activities of test plasmids containing chimeric *lat52/-45-* or *lat52/-90 CaMV35S'-TEV-luc-C3'* gene fusions were determined in pollen and leaves in the transient expression assay. pLAT52-7 was used as the reference plasmid for pollen, and pRTL2GUS was used as the reference plasmid for leaves. The fold enhancement of each test plasmid was calculated from the ratio of LUC/GUS normalised to that of the *-45 CaMV35S TEV-luc-C3'* control (average light units per bombardment was 2111872 and 5290 in pollen and leaves respectively). Data from two independent experiments are shown, representing six independent bombardments for each test plasmid. Error bars represent the standard error.

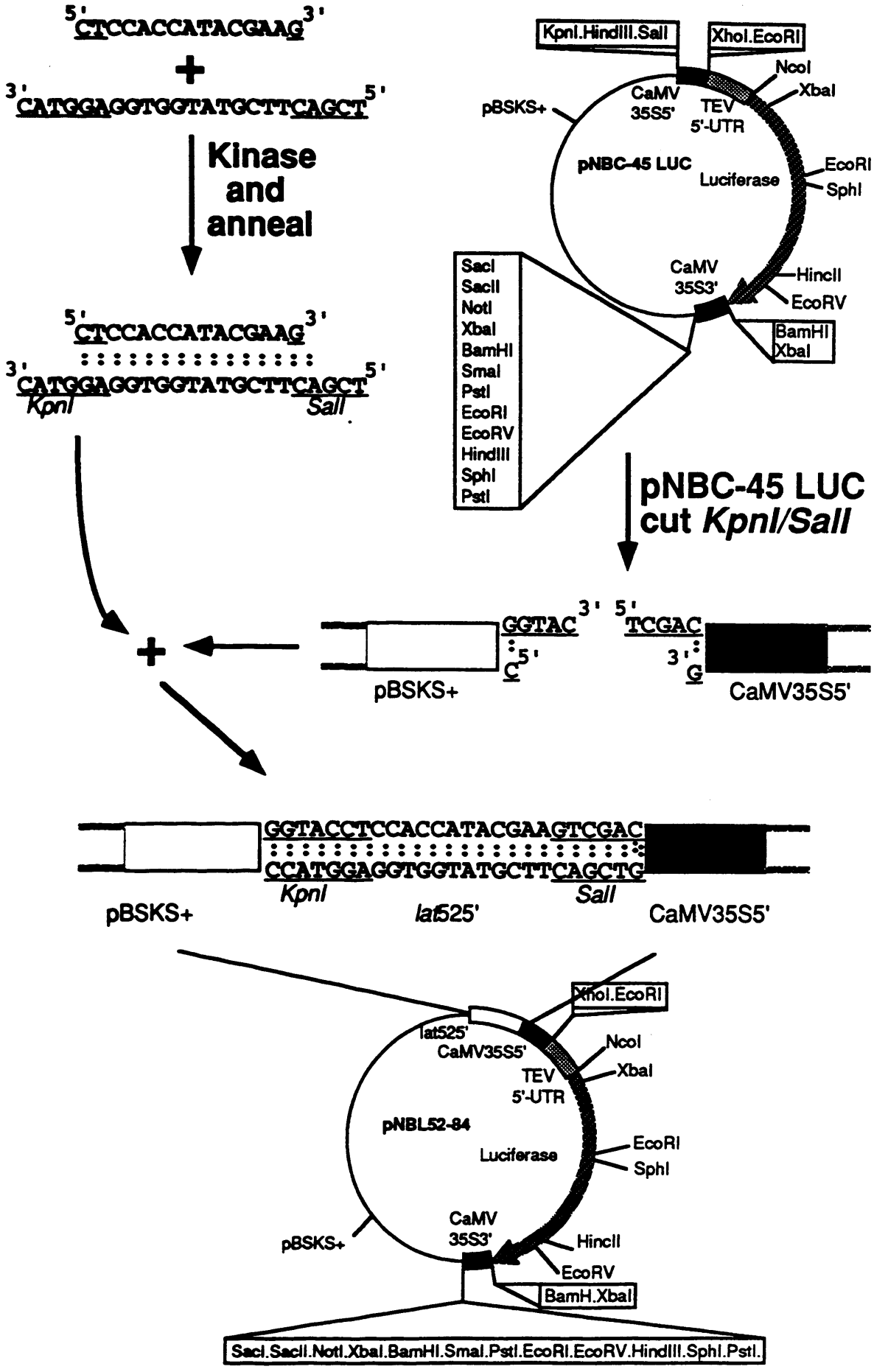


Figure 4.13. Fusion of the *lat52* promoter fragment -61 to -52 to the -45 *CaMV35S* promoter. Oligonucleotides complementary to the region -61 to -52 of the *lat52* promoter containing *KpnI* and *SalI* overhangs were inserted upstream of the -45 *CaMV35S* promoter.

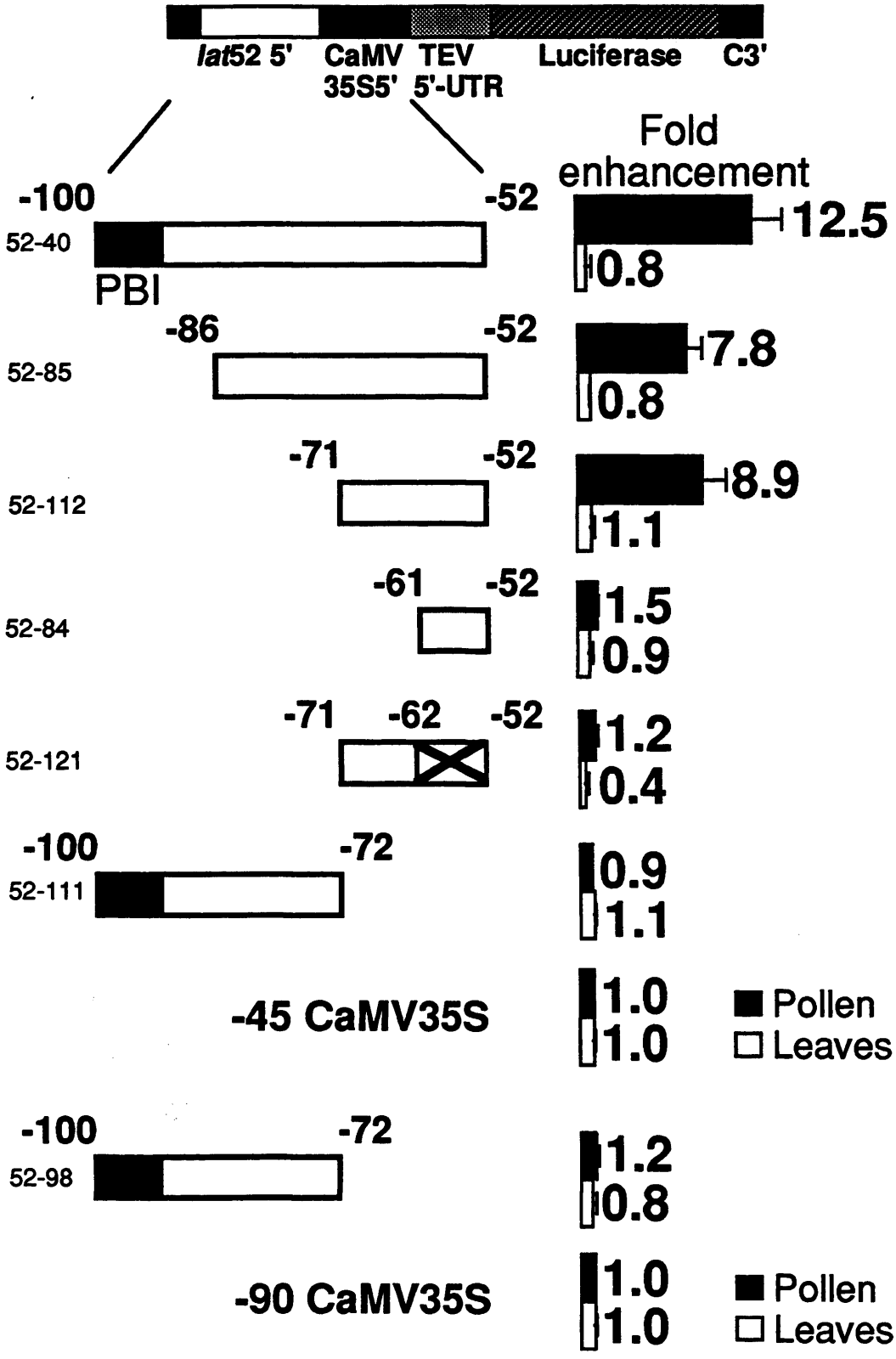


Figure 4.14. Activation of the *CaMV35S* core promoter by proximal *lat52* promoter fragments. The relative activities of test plasmids containing chimeric *lat52/CaMV35S5'-TEV-luc-C3'* gene fusions were determined in pollen and leaves in the transient expression assay. pLAT52-7 was used as the reference plasmid for pollen, and pRTL2GUS was used as the reference plasmid for leaves. The fold enhancement of each test plasmid was calculated from the ratio of LUC/GUS normalised to that of either the -45 (average light units per bombardment was 4091345 and 11072 in pollen and leaves respectively) or -90 (average light units per bombardment was 1833050 and 230513 in pollen and leaves respectively) *CaMV35S5'-TEV-luc-C3'* controls. Data from two independent experiments are shown, representing six independent bombardments for each test plasmid. Error bars represent the standard error.

and transiently expressed in germinating pollen (section 2.9.5) and leaves (section 2.9.6). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the -45 *CaMV35S* 5' promoter deletion mutant. The results of this analysis are presented in figure 4.14.

In pollen removal of PBI from domain C by deleting to -86 decreased LUC activity by ~ 33%, further deletion to -71 had no effect, but deletion to -61 reduced activity to the basal level of -45 *CaMV35S*. The region -100 to -72 alone, containing PBI, did not activate -45 *CaMV35S* or -90 *CaMV35S* . Substitution mutation of the region -60 to -52 within the region -71 to -52 abolished enhancement. In leaves none of the domain C fragments significantly affected expression of the -45 or -90 *CaMV35S* promoters. These data demonstrate that within domain C the function of PBI is strictly dependent on the presence of downstream sequences between -71 to -52, and the regulatory sequence TCCACCATA requires the presence of the upstream AGAAA motif for activity. These data also define the minimal functional pollen-specific unit of domain C to the 20 bp sequence GAAATAATAGCTCCACCATA, -71 to -52, which contains the underlined bipartite regulatory sequences.

4.5 Discussion

To investigate the functional organisation of *cis*-regulatory elements which coordinate the pollen-specific and developmentally regulated transcription of the *lat*52 gene, progressive 5' deletion, gain of function and targeted substitution mutational analyses were employed.

4.5.1 Organisation of *cis*-regulatory elements within domain A (-492 to -226)

Deletion of domain A (-492 to -226) did not significantly reduce *lat*52 promoter activity in germinating pollen (fig. 4.7), but a gain of function approach demonstrated that domain A alone activated a heterologous core promoter in germinating pollen (fig. 4.5). This data suggests that *cis*-regulatory sequences within domain A are recognised by transcription factors. Previous sequence analysis within domain A identified the core motif TGTGGTT at position -438 (PBIII) (Twell *et al.*, 1991). 5' deletion or substitution mutational analysis of PBIII in the presence of down stream sequences did not significantly reduce *lat*52 promoter activity in germinating pollen (Weterings, 1994). These results are not unexpected since the region -492 to -226 containing PBIII is redundant in germinating pollen. Therefore, this data does not rule out a functional role for the PBIII motif. Also present within domain A is the sequence

ACGTAGAAAAATAATACCATT (-332 to -312) which shows significant homology (16/21 bp: highlighted in bold) with the box VI motif AAATAGATAAATAAAAACATT from the pea *rbcS-3A* gene, which has been shown to bind the transcription factor 3AF1 (Lam *et al.*, 1990). Moreover, 2 out of the 3 nucleotides defined as essential for box VI-protein complex formation *in vitro* are conserved within the *lat*52 promoter sequence (underlined). Northern blot analysis demonstrated that 3AF1 transcript was detectable in all tissues examined including anthers and non-photosynthetic tissues such as roots (Lam *et al.*, 1990). The tomato homolog of 3AF1, E4/E8BP-1 has been shown to be involved in ethylene-regulated gene transcription (Coupe and Deikman, 1997). Furthermore, activity of the 3AF1/ E4/E8BP-1 recognition sequences appears to involve cooperative interaction with a second promoter element which suggests a requirement for other transcription factors (Lam *et al.*, 1990: Xu *et al.*, 1996). Therefore, it is possible that the transcription factor 3AF1 binds to the *lat*52 promoter domain A and cooperatively interacts with a pollen-specific factor which could bind to the PBIII motif (or elsewhere) to activate transcription specifically in pollen.

4.5.2 Organisation of *cis*-regulatory elements within the region -225 to -101

Previous analysis demonstrated that a fragment (-194 to -176) containing the PBII motif (AACCACA) activated a heterologous core promoter specifically in germinating pollen (Twell *et al.*, 1991). In contrast, a fragment (-100 to -72) containing the PBI motif (TGTGGTT) failed to activate -45 or -90 *CaMV35S* in germinating pollen (fig. 4.14). These analyses strongly indicate that the motifs PBII and PBI bind different transcription factors even though PBII is a reverse copy of the PBI motif. Interestingly, the motif **AACCACAC** which shares 7/8 bp homology (highlighted in bold) with PBII, has been implicated as responsible for the activity of the spinach plastocyanin promoter in pollen (Lubberstedt *et al.*, 1994).

Results presented in figure 4.6 showed that in the presence of downstream sequences (-176 to -101) PBII is redundant in germinating pollen since mutation or 5' deletion of PBII did not significantly decrease the pollen-specific activity of domain B. Although PBII appears to be redundant in germinating pollen, other pollen-specific regulatory elements within the region -176 to -101 are functional within domain B. Examination of this region revealed the sequence TCAC at position -126 to -123 which in reverse orientation (GTGA) contains the four nucleotides defined as essential for the functional activity of the pollen-specific 56/59 box (Twell *et al.*, 1991). Previous analysis in transgenic tomato showed that deletion of the *lat*52 promoter from -145 to -124 had little effect on promoter activity. This deletion removes the TC nucleotides of the reverse

56/59 box and could therefore be sufficient to inactivate it since all four core nucleotides are required for pollen-specific activity of the 56/59 box (Twell *et al.*, 1991). Therefore, it is unlikely that the reverse copy of the 56/59 box (TCAC) is functional in domain B of the *lat*52 promoter. In transgenic tomato deletion from -124 to -71 decreased activity by 100-fold (Twell *et al.*, 1991), whereas data presented in figure 3.2 showed that deletion of the region -100 to -71 in transgenic tobacco decreased promoter activity by only 15-fold. These results indicate that functional sequences are located between -124 to -101.

The region -176 to -101 (sub-domains B2 and B3) is extremely AT-rich and contains no previously detailed pollen-specific *cis*-regulatory sequences. The region -176 to -146 (sub-domain B2) which enhanced *lat*52 promoter activity in pollen (fig.s 4.6 & 4.7) is 97 % AT-rich (29/30 A or T nucleotides) and therefore could be a binding site for high mobility group protein 1 (HMG 1) -like proteins. HMG1 is a non histone, chromatin-associated nuclear protein which as been shown to enhance the sequence-specific DNA binding activity *in vitro*, and the transcriptional activation *in vivo*, of the human HOXD9 protein (Zappavigna *et al.*, 1996). In plants HMG1-like proteins isolated from soybean bind to AT-rich regions in soybean nodulin promoters. Their proposed role is suggested to be the modification of chromatin structure in nodulin promoters as a prerequisite for the organ-specific interaction with other factors (Jacobsen *et al.*, 1990). Such factors binding within the sub-domain B2 (-176 to -146) could stimulate the binding of pollen-specific activators to the downstream sub-domain B3 (-124 to -101).

4.5.3 Organisation of *cis*-regulatory elements within the region -100 to -52

Results presented in this chapter and in chapter 3 have demonstrated a significant role for domains A+B (-492 to -101) in contributing to high level pollen-specific transcription. Although this region is important for directing pollen-specific expression it is by no means the only critical functional unit involved. Domain C (-100 to -42) is also essential for high level *lat*52 transcriptional activity during pollen maturation and germination (figs. 3.2, 3.4 & 3.10). To define the location and determine the putative interactions of other pollen-specific *cis*-regulatory elements within the region -100 to -42 with PBI a series of substitution mutations were introduced progressively throughout the region -84 to -42 in context of the promoter deletion -100 to +110. Analysis of these substitution mutations revealed that mutation of sequences between -84 to -52 led to a substantial decrease in the activity of the promoter deletion mutant -100 to +110 in pollen (fig 4.10). The differential effects of each mutation suggests that there are at least two other *cis*-regulatory elements contained within the region -84 to -52. Which are flanked by sequences that may contain a minor part of each element or are involved in controlling

the flexibility of the promoter fragment, and hence possible physical interactions between *trans*-acting factors.

A further interesting property of domain C is that it conferred high level pollen-specific expression on the *CaMV35S* promoter independent of distance (fig. 4.12). These data define domain C as an independent moveable functional unit which is sufficient to direct pollen-specific transcription. This result confirms the data presented by Eyal *et al.*, (1995), but in the experiments described here the absolute level of pollen-specific enhancement of the minimal *CaMV35S* promoter mediated by the region -100 to -52 was at least 100-fold lower. This major difference in the activation potential demonstrated by the region -100 to -52 may be explained as follows. The plasmid pJO64 (*-64 CaMV35S-luciferase*) used as a base control to test for the ability of the region -100 to -55 to enhance minimal *CaMV35S* in the gain of function experiments presented by Eyal *et al.*, 1995 was stated to be inactive in both pollen and TXD cell culture. This would prevent a direct quantitative comparison of the activities shown by the test plasmids and the control -64 *CaMV35S* plasmid. In contrast, a translational enhancer from tobacco etch virus (TEV 5'-UTR) (Carrington and Freed, 1990) was fused to the minimal *CaMV35S* promoter to elevate the activity of each control construct well above background levels, thus allowing a direct measurement of the functional activity exhibited by each *lat52* promoter region relative to the -90 or -45 *CaMV35S* control plasmids.

The location and organisation of regulatory elements within domain C was further investigated. Mutation of the PBI core region GTGG decreased but did not abolish the pollen-specific enhancement of the minimal -45 or -90 *CaMV35S* promoters (fig. 4.12). The region -100 to -72, containing PBI alone was unable to activate either the -45 or -90 *CaMV35S* in a pollen-specific manner (fig. 4.14). It has been proposed that the DNA-binding protein GT-1a could bind to PBI. Mutation of PBI to that of a GT-1a binding site did not reduce promoter activity in pollen (Eyal *et al.*, 1995). However, data presented in figure 4.9 functionally defined the central GG residues of PBI as the only functional similarity between PBI and the GT-1 binding site (*rbcS-3a* boxII). Although PBI is unlikely to bind GT-1a, this does not rule out the possibility that it could be recognised by a closely related factor. Indeed, GT-1-related RT-PCR products were detected in tobacco pollen (Eyal *et al.*, 1995). In addition, nuclear GT-1-like factors have been shown to interact with non light-responsive genes such as the tobacco *PR-1a* gene (Buchel *et al.*, 1996) and with the pollen-specific *NTP303* gene (Hochstenbach *et al.*, 1996).

Since PBI alone could not activate minimal *CaMV35S* in pollen, this would indicate that transcription factors recognising PBI require other downstream transcription factors localised to the region -71 to -52 for their activation or binding. The region - 61 to -52 (CTCCACCATA) contains the sequence CCAC which is an inverted copy of the PBI

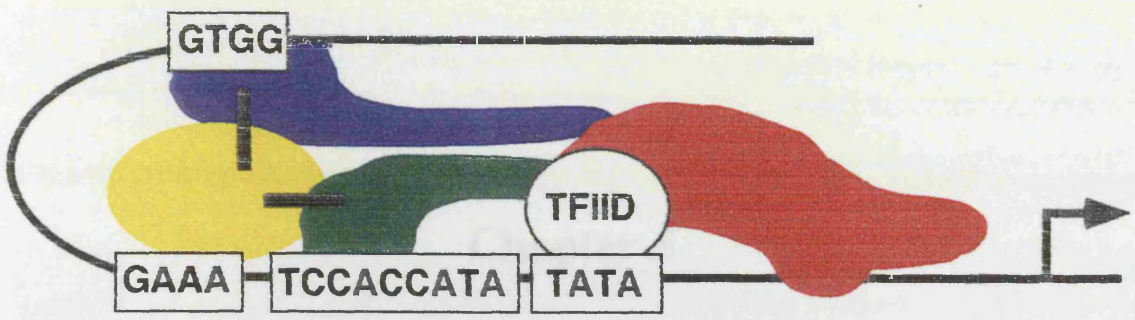
core region GTGG. Mutation of the region -60 to -52 decreased pollen-specific activity (fig. 4.10), but in the gain of function analysis the region -61 to -52 containing only the CCAC motif could not activate minimal *CaMV35S* (fig. 4.14). Substitution analysis of the region -84 to -42 identified a further functional motif AGAAA (-72 to -68), which when mutated virtually abolished promoter activity (fig. 4.10). This not only highlights the functional importance of this motif but also shows that PBI and TCCACCATA alone are not sufficient to form an active transcription complex. PBI in conjunction with the AGAAA motif does not require the presence of the downstream TCCACCATA motif since mutation of the TCCACCATA motif in the region -100 to +110 decreased but did not abolish promoter activity (fig. 4.10). *Vice versa* the TCCACCATA motif does not require the functional presence of PBI since the region -71 to -52 (GAAA + TCCACCATA) activated minimal *CaMV35S* in a pollen-specific manner (fig. 4.14). The region -71 to -62 containing only the motif GAAA was not sufficient to activate minimal *CaMV35S* in pollen (fig. 4.14). Therefore, the minimal pollen-specific functional unit was defined as the 20 bp sequence between -71 and -52 which contains the putative binding sites (GAAA + TCCACCATA) for two transcription factors.

Even though the TCCACCATA motif contains the sequence CCAC which is an inverted copy of the PBI core region GTGG and both motifs are dependent on the GAAA motif for function it is possible that different transcription factors bind to both motifs. Comparison of the developmental GUS accumulation profiles between plants harbouring either the -100 or -71 *lat52-gus* fusions (fig. 3.4), showed that for -100, GUS activity accumulated steadily during early pollen development (bud length 16-30 mm) followed by a sharp increase in GUS activity at the later stages of pollen maturation (bud length ~ 30 mm). In contrast, -71 only showed a detectable increase in GUS activity in buds at or over 30 mm in length. The motif GAAA is essential for promoter activity. Since the -100 deletion is active from pollen mitosis I (bud length 16 mm) onwards it can be assumed that the factor binding this motif is present throughout pollen development. The GAAA motif alone is inactive (fig. 4.14), therefore one other motif must be functional during early pollen development. It is proposed that the activity of -100 during the early stages of pollen development arises from transcription factors binding to the motifs PBI and GAAA, to form a weak pollen-specific transcriptional activation complex. The sharp increase in GUS activity at the later stages of pollen maturation (bud length ~ 30 mm) for -100 is achieved by a new transcription factor binding to the TCCACCATA motif, which together with the transcription factors already binding to the PBI and GAAA motifs forms a strong pollen-specific transcriptional activation complex.

Two models for the action of putative transcription factors binding within the region -100 to -52 are presented in figure 4.15. The first model involves a factor binding at the GAAA (-71 to -68) motif which facilitates and stabilises the binding of further

transcription factors to the GTGG (-95 to -92) and TCCACCATA (-60 to -52) to contact TFIID and RNA polymerase II (fig. 4.15 A). The second model involves factors binding to the GTGG (-95 to -92) and TCCACCATA (-60 to -52) motifs which facilitate and stabilise the binding of a transcription factor to the GAAA (-71 to -68) motif.

A



B

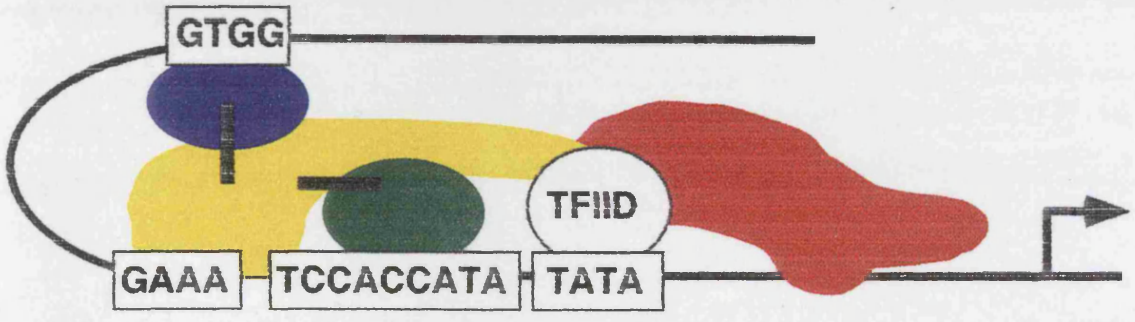


Figure 4.15. Two models for the action of putative transcription factors binding within the region 100 to -52. A) The first model involves a factor binding at the GAAA motif (yellow) which facilitates and stabilises the binding of further transcription factors to GTGG (blue) and TCCACCATA (green) to contact TFIID (white) and RNA polymerase II (red). B) The second model involves factors binding to GTGG (blue) and TCCACCATA (green) motifs which facilitate and stabilise the binding of a transcription factor to the GAAA (yellow) motif

Chapter 5

Maturation-specific translational enhancement mediated by the *lat*52 5'-UTR

5.1 Introduction

Analysis of *lat*52 5' promoter deletion mutants in transgenic tobacco demonstrated that the minimal region -41 to +110, was preferentially expressed in pollen (chapter 3). This region contains a TATA box (-28 to -25) which is sufficient to support similar basal levels of transcription in both pollen and leaves. Analysis of the region -41 to + 110 in the transient expression assay system indicated that sequences controlling pollen-specific gene expression were localised within the 5' untranslated region (UTR), +1 to +110. This chapter details experiments designed to establish the role of the *lat*52 5'-UTR in controlling the pollen-specific expression of *lat*52. A series of promoter/5'-UTR fusions were constructed and analysed in the transient expression assay system and in transgenic plants. In this analysis the activity of constructs containing the *lat*52 5'-UTR were compared to otherwise identical constructs containing a synthetic polylinker 5'-UTR. The activity of the *lat*52 5'-UTR was quantified in its native context with respect to the *lat*52 promoter, and also when fused downstream of the *CaMV35S* promoter. In addition, fusion to the *CaMV35S* promoter allowed the activity of the *lat*52 5'-UTR to be analysed in sporophytic tissues.

5.2 The *lat*52 5'-UTR preferentially enhances gene activity in pollen independent of fused promoter, reporter gene and 3'-UTR sequences

5.2.1 The *lat*52 5'-UTR is essential for maximal levels of gene expression in pollen

In order to determine if the *lat*52 5'-UTR contributed to the pollen-specific activity of the region -41 to +110 the *lat*52 5'-UTR was replaced with the 5'-UTR from tobacco etch virus (TEV) and also with a synthetic polylinker sequence (SYN). To investigate any possible promoter/5'-UTR interactions the *lat*52 5'-UTR was replaced with the TEV and SYN 5'-UTRs in context of the full promoter region -492 to -1. The steps involved in the construction of the plasmid *lat52-TEV 5'-UTR-luc-C3'* (pNBLT-1) are shown schematically in figure 5.2. The *lat*52 promoter region -492 to -1 was amplified from the plasmid pNBL52-5 (section 3.3.2), digested with *XhoI* and cloned into the plasmid pNBL-0 (section 3.3.1). The steps involved in the construction of the plasmid *lat52-SYN 5'-UTR-luc-C3'* (pNBL52-42) are shown schematically in figure 5.3. The *SalI/NcoI* polylinker fragment from pGEM 5Z(-) (Promega) was cloned into pSL301 (Invitrogen). The synthetic polylinker 5'-UTR was excised as a *BamHI/NcoI* fragment and cloned into

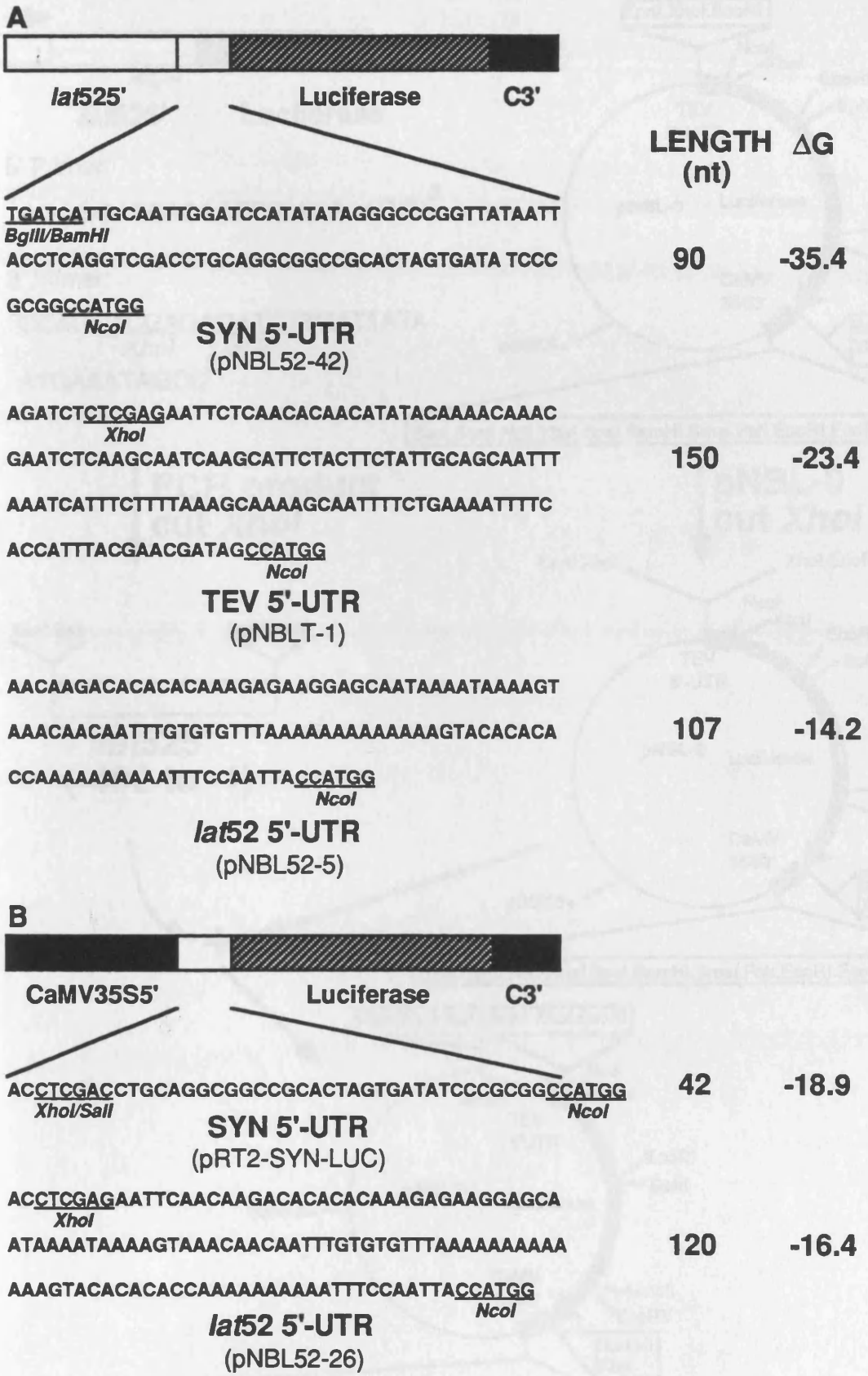


Figure 5.1. Test gene fusion constructs containing different 5'-UTRs used in transient expression analysis. Constructs present in test plasmids containing the *lat*52 promoter (A) or *CaMV35S* promoter (B) are shown, with the nucleotide sequences beginning at +1 of the different 5'-UTRs. Promoter sequences were fused to the *luciferase* reporter gene and the 3'-UTR (C3') from *CaMV35S*. The length, in nucleotides, and the calculated stability (ΔG in kcal/mol) of secondary structure within the 5'-UTRs are shown on the right. Restriction sites used in plasmid construction and initiator ATGs are underlined.

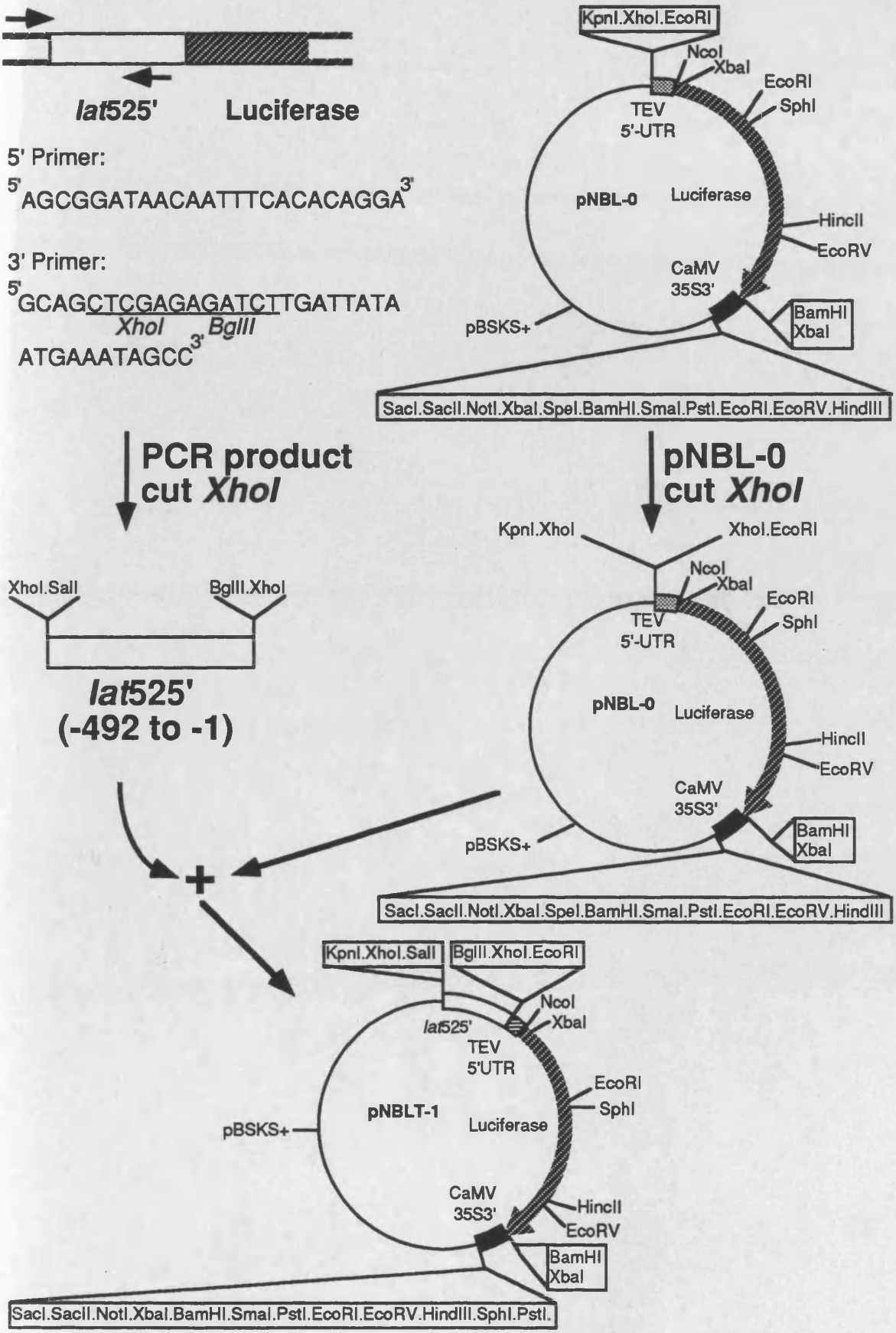


Figure 5.2. Construction of the plasmid pNBLT-1. The *lat*52 promoter region -492 to -1 was amplified by PCR, cut with *XhoI* and cloned into the plasmid pNBL-0.

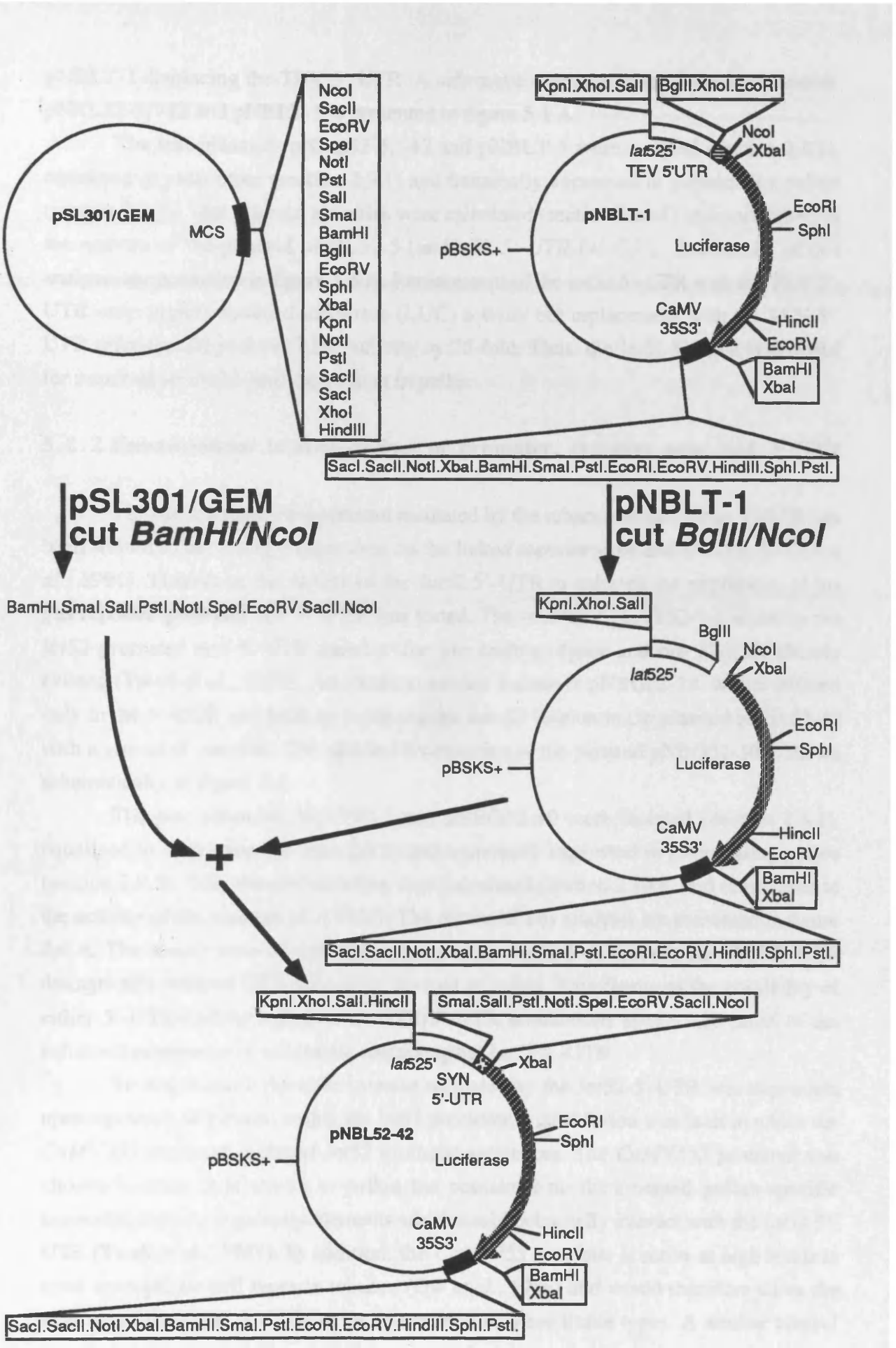


Figure 5.3. Construction of the plasmid pNBL52-42. The *SalI/NcoI* polylinker fragment from pGEM 5Z(-) was cloned into pSL301. The synthetic polylinker 5'-UTR was excised as a *BamHI/NcoI* fragment and cloned into pNBLT-1 replacing the TEV 5'-UTR.

pNBLT-1 displacing the TEV 5'-UTR. A schematic representation of the test plasmids pNBL52-5, -42 and pNBLT-1 is presented in figure 5.1 A.

The test plasmids pNBL52-5, -42 and pNBLT-1 were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the plasmid pNBL52-5 (*lat52-52 5'-UTR-luc-C3'*). The results of this analysis are presented in figure 5.6 A. Replacement of the *lat*52 5'-UTR with the TEV 5'-UTR only slightly reduced luciferase (LUC) activity but replacement with the SYN 5'-UTR dramatically reduced LUC activity by 25-fold. Thus, the *lat*52 5'-UTR is essential for maximal levels of gene expression in pollen.

5.2.2 Enhancement is independent of promoter, reporter gene and 3'-UTR

The translational enhancement mediated by the tobacco mosaic virus 5'-UTR has been shown to be strongly dependent on the linked reporter gene and 3'-UTR (Gallie *et al.*, 1991). Therefore, the ability of the *lat*52 5'-UTR to enhance the expression of the *gus* reporter gene and *nos* 3'-UTR was tested. The construct pLAT52-7 containing the *lat*52 promoter and 5'-UTR fused to the *gus* coding region and *nos* 3'-UTR already existed (Twell *et al.*, 1991). An identical control construct pNBG52-10, which differed only in the 5'-UTR was built by replacing the *luc-C3'* region in the plasmid pNBL52-42 with a *gus-nos3'* cassette. The detailed construction of the plasmid pNBG52-10 is shown schematically in figure 5.4.

The test plasmids pLAT52-7 and pNBG52-10 were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the plasmid pLAT52-7. The results of this analysis are presented in figure 5.6 A. The results showed that replacement of the *lat*52 5'-UTR with the SYN 5'-UTR dramatically reduced GUS activity by 60-fold in pollen. This dismisses the possibility of either 5'-UTR/coding region or 5'-UTR/3'-UTR interactions as the sole cause of the enhanced expression of constructs containing the *lat52* 5'-UTR.

To determine if the enhancement mediated by the *lat*52-5'-UTR was dependent upon upstream sequences within the *lat*52 promoter, a gene fusion was built in which the *CaMV35S* promoter replaced *lat*52 promoter sequences. The *CaMV35S* promoter was chosen because it is active in pollen but contained no documented pollen-specific transcriptional *cis*-regulatory elements which could potentially interact with the *lat*52 5'-UTR (Twell *et al.*, 1991). In addition, the *CaMV35S* promoter is active at high levels in most sporophytic cell types in tobacco (Ow *et al.*, 1986) and would therefore allow the activity of the *lat*52 5'-UTR to be determined in other tissue types. A similar control

construct containing a synthetic polylinker sequence *CaMV35S-SYN 5'-UTR-luc-C3'* (pRT2-SYN-LUC) was built by R. Turner (Turner *et al.*, 1994). The steps involved in the construction of the plasmid *CaMV35S-52 5'-UTR-luc-C3'* (pNBL52-26) are shown schematically in figure 5.5. The *lat*52 5'-UTR was amplified from the plasmid pNBL52-5 (section 3.3.2) and ligated to the *luc* coding region from the plasmid pRT2-TEV-LUC (section 3.3.1) and the *CaMV35S 3'-pUC19-2 x CaMV35S 5'* fragment from the plasmid pRTL2GUS. A schematic representation of the test plasmids pRT2-SYN-LUC and pNBL52-26 is shown in figure 5.1 B.

The test plasmids pRT2-SYN-LUC and pNBL52-26 were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the plasmid pRT2-SYN-LUC. The results of this analysis are presented in figure 5.6 A. The results showed that the *lat*52 5'-UTR dramatically enhanced LUC activity 14-fold above the SYN 5'-UTR. Levels of enhancement in pollen (14-fold) were less than 2-fold lower than those observed when the *lat*52 5'-UTR was present in its native context downstream of the *lat*52 promoter (Fig. 5.6 A). Therefore, enhancement was largely independent of upstream transcriptional regulatory sequences within the *lat*52 promoter.

5.2.3 Enhancement mediated by the *lat*52 5'-UTR is significantly reduced in sporophytic cells

Data presented in section 5.2.2 showed that the *lat*52 5'-UTR when fused downstream of the *CaMV35S* promoter enhanced LUC activity by 14-fold in pollen. To investigate the ability of the *lat*52 5'-UTR to enhance gene expression in sporophytic cell types, the activity of the *luc*-based constructs driven by the *CaMV35S* promoter (pNBL52-26 and pRT2-SYN-LUC: section 5.2.2) were transiently expressed in leaves (section 2.9.6) and suspension cultured cells (section 2.9.7). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the plasmid pRT2-SYN-LUC. The results of this analysis are presented in figure 5.6 B, C. In contrast to the results obtained in pollen, the presence of the *lat*52 5'-UTR led only to a modest 2-fold enhancement of gene expression above the SYN 5'-UTR control. This low level of enhancement was reproducible and similar in both leaves and suspension cells. These data demonstrated a strong preferential activity of the *lat*52 5'-UTR in pollen.

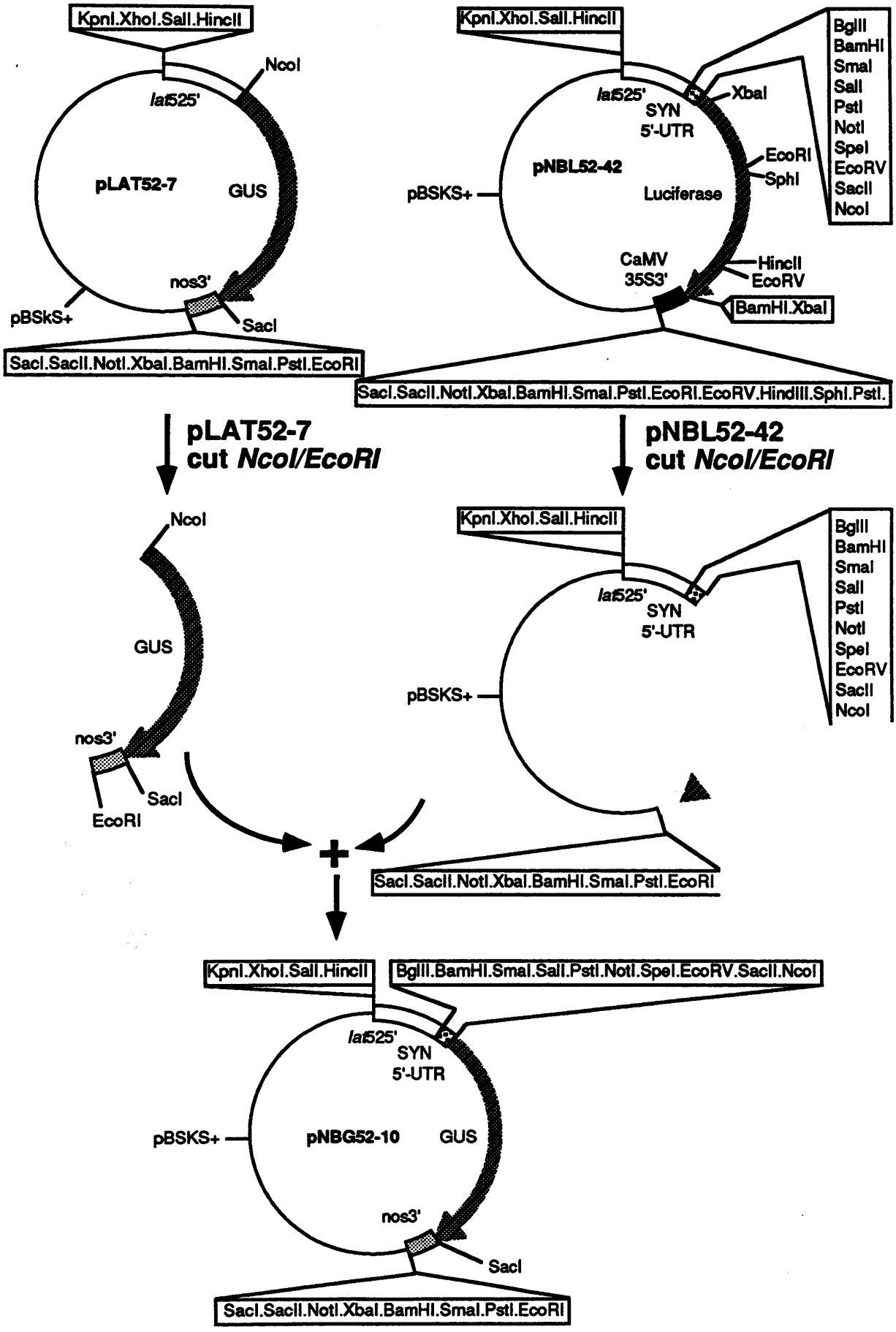


Figure 5.4. Construction of the plasmid pNBG52-10. *gus-nos3'* was excised as a *NcoI/EcoRI* fragment from pLAT52-7 and inserted into the plasmid pNBL52-42 displacing the *luc-C3'* region.

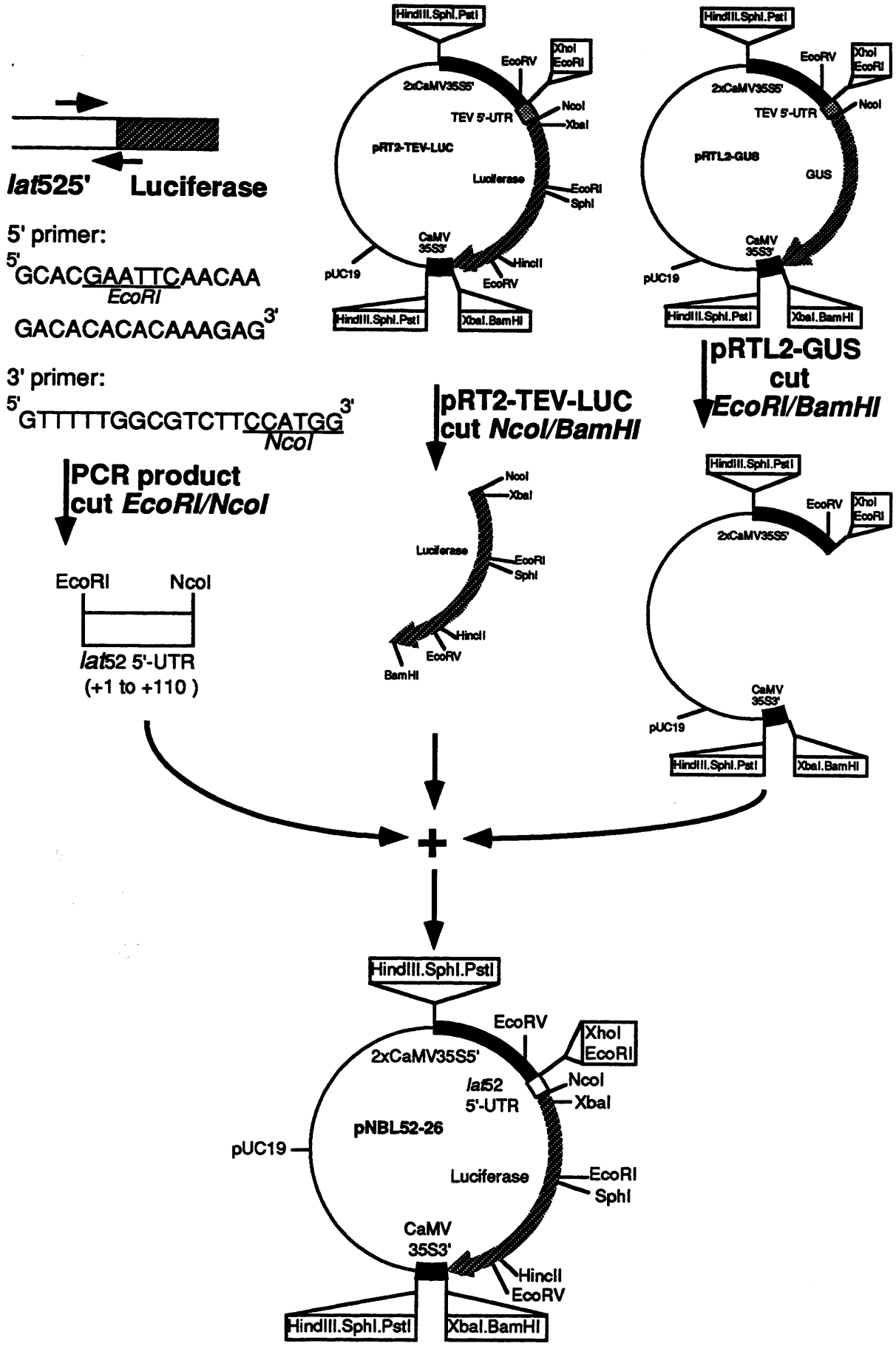


Figure 5.5. Construction of the plasmid pNBL52-26. The *lat*52 5'-UTR was amplified by PCR, cut with *EcoRI/NcoI* and ligated to the *luc* coding region from pRT2-TEV-LUC and the *CaMV35S 3'-pUC19-2 x CaMV35S 5'* fragment from pRTL2GUS.

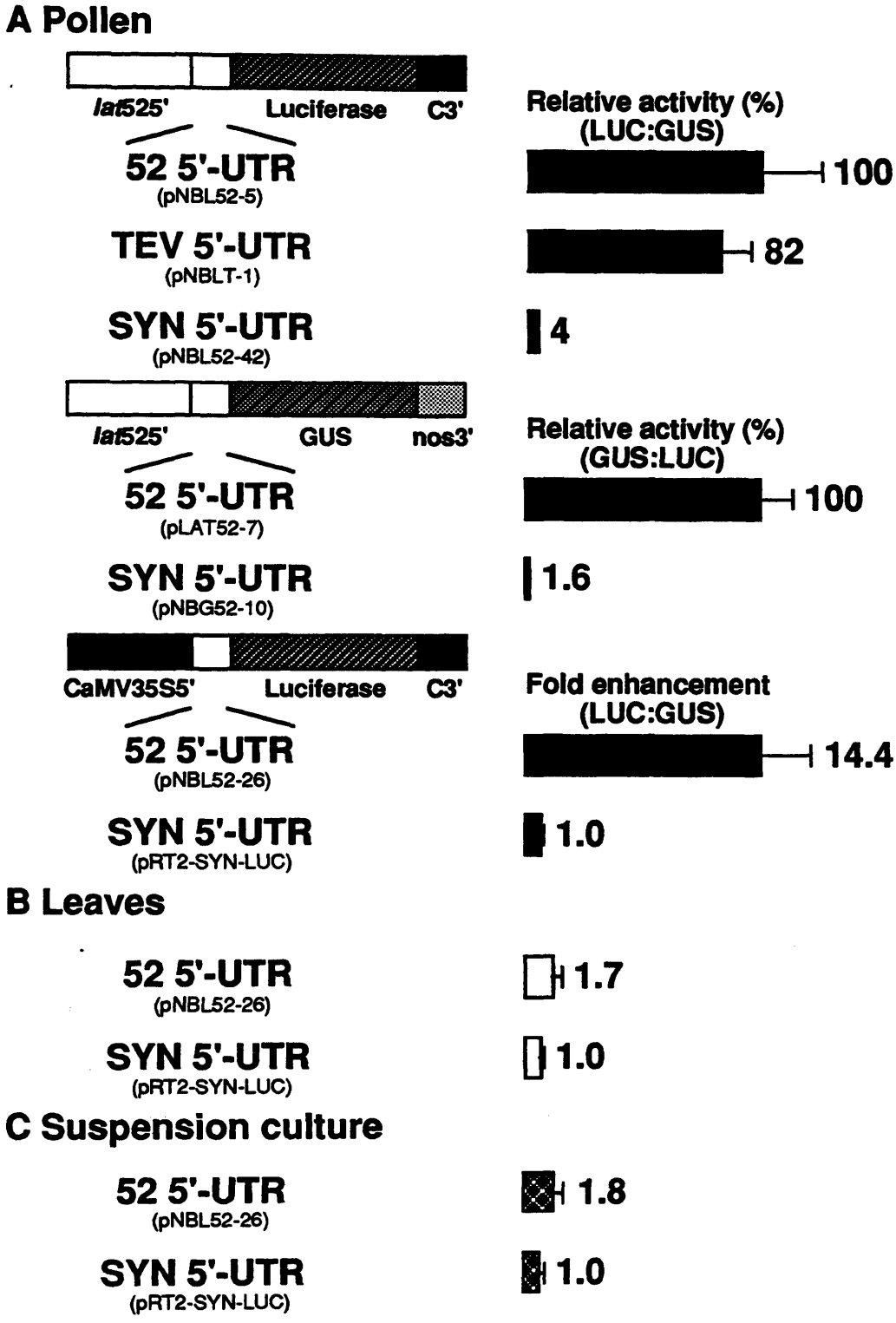


Figure 5.6. Transient expression analysis of test gene fusion constructs containing different combinations of promoter and 5'-UTR. Test plasmids containing *luc* gene fusion constructs shown were co-bombarded into pollen (A) together with the reference plasmid pLAT52-7. Test plasmids containing *gus* gene fusion constructs were co-bombarded into pollen (A) together with the reference plasmid pNBL52-5. Test plasmids were co-bombarded into leaves (B) or suspension cells (C) together with the reference plasmid pRTL2GUS. Relative activities of plasmids containing the *lat*52 5'-UTR represent the fold increase in LUC or GUS activity normalized to the activity of the equivalent construct in which the *lat*52 5'-UTR was replaced with a synthetic polylinker (SYN). Average light units per bombardment for the plasmids *lat52-SYN-luc-C3'* and *CaMV35S-SYN-luc-C3'* in polen was 9822000 and 894205 respectively. Average light units per bombardment for the plasmid *CaMV35S-SYN-luc-C3'* in leaves and suspension culture was 360475 and 2251167 respectively. The combined data from two independent experiments are shown, representing six independent bombardments for each plasmid. Error bars represent the standard error.

5.2.4 The *lat*52 5'-UTR enhances gene expression when stably integrated into the plant genome

To determine whether the *lat*52 5'-UTR-mediated enhancement observed in the transient expression assay (fig 5.6) also occurred *in planta*, the four gene fusion constructs, *lat52-SYN-luc-C3'*, *lat52-52-luc-C3'*, *CaMV35S-SYN-luc-C3'* and *CaMV35S-52-luc-C3'* (Fig. 5.1) were inserted into the binary vector pBIN19 (Bevan, 1984) and stably introduced into tobacco by *Agrobacterium*-mediated transformation. The steps involved in the construction of the *lat52-5'-UTR-luc-C3'* binary plasmid fusions are shown schematically in figure 5.7. *lat52-5'-UTR-luc-C3'* cassettes were excised from the plasmids pNBL52-5 and pNBL52-42 as *Kpn*I/*Hind*III fragments and subcloned into pBIN19 to give the plasmids pNBL52-43 and pNBL52-44 respectively. The steps involved in the construction of the *CaMV35S-5'-UTR-luc-C3'* binary plasmid fusions are shown schematically in figure 5.8. *CaMV35S-5'-UTR-luc-C3'* cassettes were excised from the plasmids pNBL52-26 and pRT2-SYN-LUC as *Hind*III fragments and subcloned into pBIN19 to give the plasmids pNBL52-45 and pNBL52-45 respectively.

Thirty to thirty-five independent transgenic lines were generated for each construct. LUC activity was determined (section 2.10.5) in extracts prepared from roots, leaves and mature pollen for primary transformants containing constructs driven by the *CaMV35S* promoter. LUC activity was also determined (section 2.10.5) in extracts prepared from leaves and mature pollen for primary transformants containing constructs driven by the *lat*52 promoter. The results of this analysis are presented in figure 5.9. Results showed that in pollen the *lat*52 5'-UTR in combination with the *lat*52 or *CaMV35S* promoters strongly enhanced LUC activity by 26- and 12-fold respectively above the SYN 5'-UTR control. In leaves the *lat*52 5'-UTR in combination with the *lat*52 or *CaMV35S* promoters enhanced LUC activity by 2.4- and 4.7-fold respectively above the SYN 5'-UTR control. Similarly, the presence of the *lat*52 5'-UTR enhanced LUC activity by 3-fold in roots relative to the SYN 5'-UTR control. Mean levels of enhancement detected in transgenic plants were similar to those observed in the transient expression assay system (compare fig. 5.6 with fig. 5.9). These data confirm the ability of the *lat*52 5'-UTR to act as an enhancer of gene expression when stably integrated into the plant genome. They also further substantiate the results of the transient expression experiments, in that enhancement mediated by the *lat*52 5'-UTR was highly preferential in pollen.

Interestingly, levels of enhancement mediated by the *lat*52 5'-UTR were similar in pollen and in leaves for constructs driven by the *lat*52 promoter or the *CaMV35S* promoter. This suggests that the enhancement observed was independent of the intrinsic strength of transcriptional regulatory sequences and therefore, of transcript abundance.

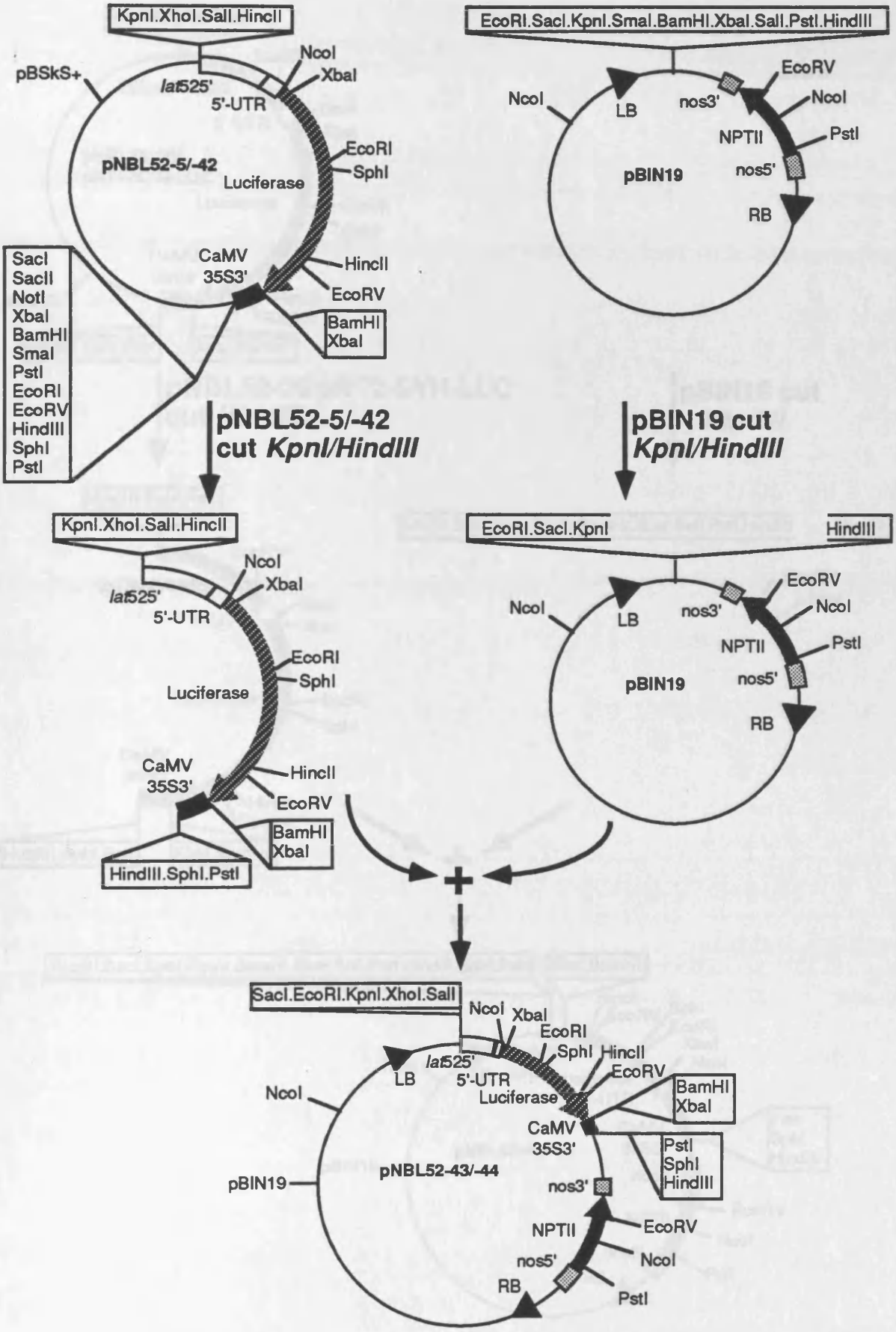


Figure 5.7. Construction of *lat52-5'-UTR-luc-C3'* binary plasmid fusions. *lat52-5'-UTR luc-C3'* cassettes were excised from the plasmids pNBL52-5 and pNBL52-42 as *KpnI/HindIII* fragments and subcloned into pBIN19 to give the plasmids pNBL52-43 and pNBL52-44 respectively.

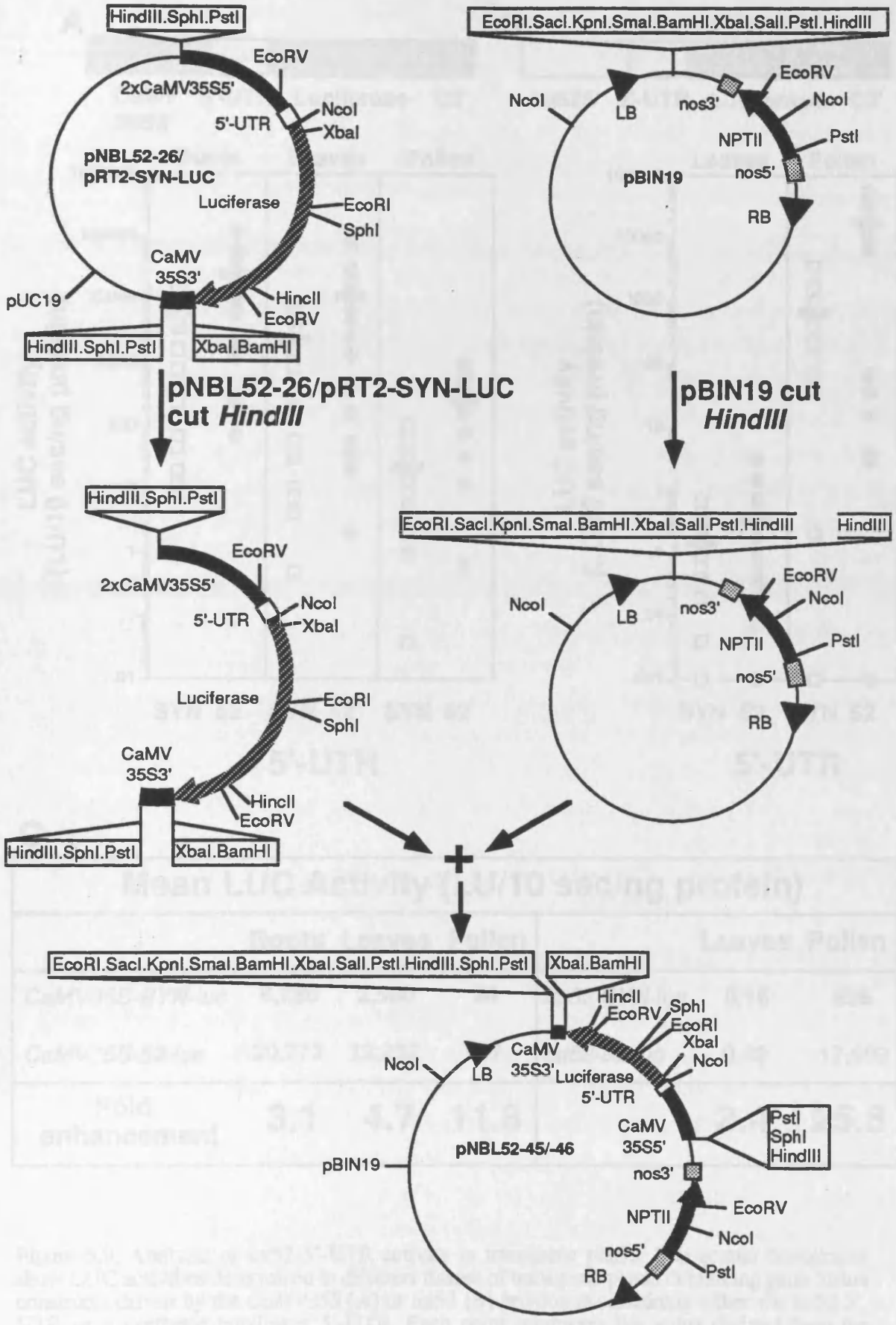
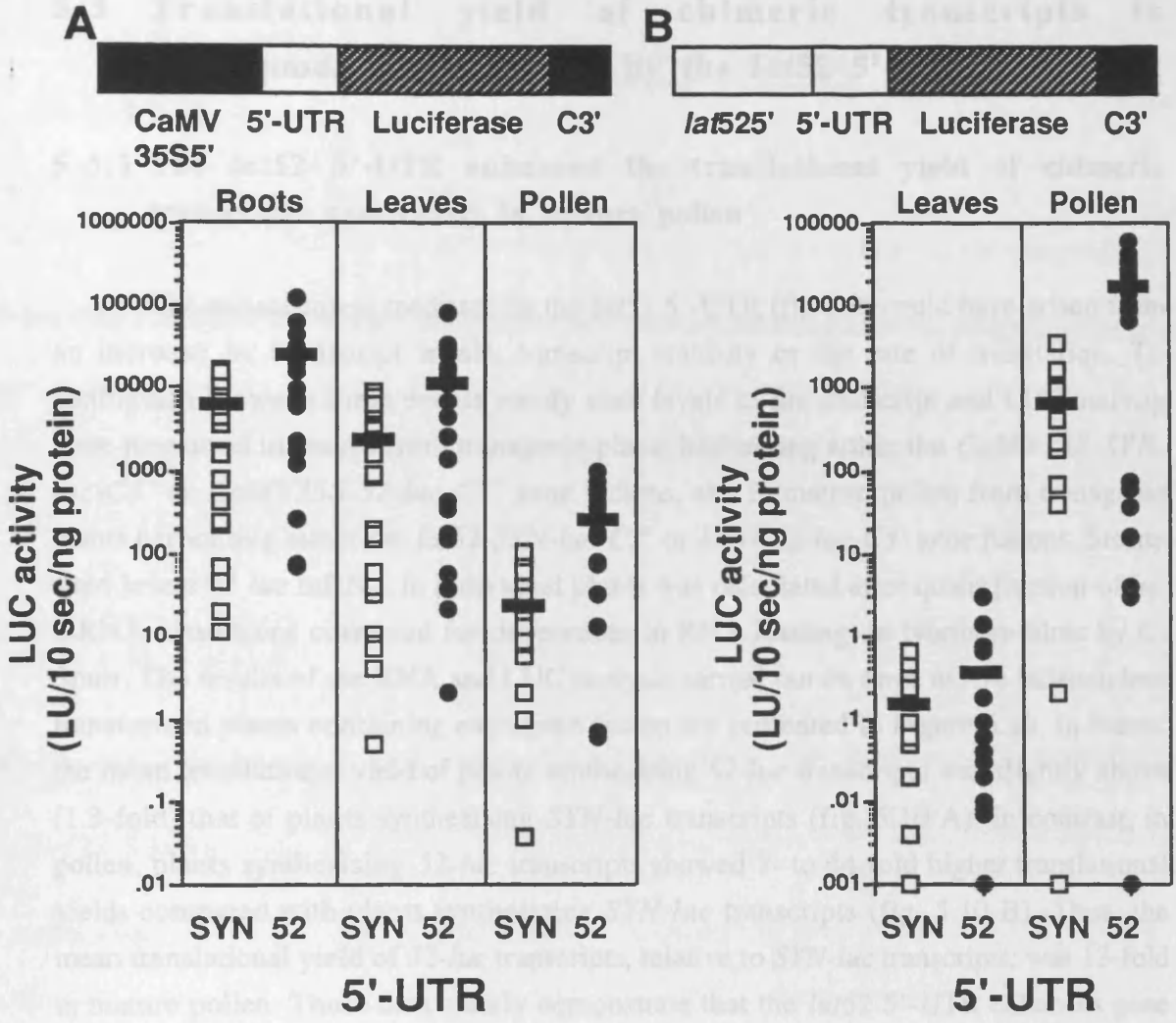


Figure 5.8. Construction of the *CaMV35S-5'-UTR-luc-C3'* binary plasmid fusions. *CaMV35S-5'-UTR-luc-C3'* cassettes were excised from the plasmids pNBL52-26 and pRT2-SYN-LUC as *HindIII* fragments and subcloned into pBIN19 to give the plasmids pNBL52-45 and pNBL52-45 respectively.



C

Mean LUC Activity (LU/10 sec/ng protein)						
	Roots	Leaves	Pollen		Leaves	Pollen
CaMV35S-SYN-luc	6,780	2,590	24	*lat52-SYN-luc*	0.16	696
CaMV35S-52-luc	20,772	12,232	287	*lat52-52-luc*	0.39	17,989
Fold enhancement	**3.1**	**4.7**	**11.8**		**2.4**	**25.8**

Figure 5.9. Analysis of *lat*52-5'-UTR activity in transgenic plants. The scatter histograms show LUC activities determined in different tissues of transgenic plants containing gene fusion constructs driven by the *CaMV35S* (A) or *lat*52 (B) promoters containing either the *lat*52 5' UTR or a synthetic polylinker 5'-UTR. Each point represents the value derived from the analysis of a single independent transgenic plant, with the mean values indicated by a black bar. The data summarized below (C) shows the mean values and the relative enhancement mediated by the *lat*52-5'-UTR in each tissue.

5.3 Translational yield of chimeric transcripts is developmentally regulated by the *lat*52 5'-UTR

5.3.1 The *lat*52 5'-UTR enhanced the translational yield of chimeric transcripts specifically in mature pollen

The enhancement mediated by the *lat*52 5'-UTR (fig 5.9) could have arisen from an increase in: transcript levels, transcript stability or the rate of translation. To distinguish between these events steady state levels of *luc* transcript and LUC activity were measured in leaves from transgenic plants harbouring either the *CaMV35S-SYN-luc-C3'* or *CaMV35S-52-luc-C3'* gene fusions, and in mature pollen from transgenic plants harbouring either the *lat52-SYN-luc-C3'* or *lat52-52-luc-C3'* gene fusions. Steady state levels of *luc* mRNA in individual plants was calculated after quantification of *luc* mRNA abundance corrected for differences in RNA loading on Northern blots by C. Spurr. The results of the RNA and LUC analysis carried out on three to five independent transformed plants containing each gene fusion are presented in Figure 5.10. In leaves the mean translational yield of plants synthesising *52-luc* transcripts was slightly above (1.3-fold) that of plants synthesising *SYN-luc* transcripts (fig. 5.10 A). In contrast, in pollen, plants synthesizing *52-luc* transcripts showed 7- to 44-fold higher translational yields compared with plants synthesizing *SYN-luc* transcripts (fig. 5.10 B). Thus, the mean translational yield of *52-luc* transcripts, relative to *SYN-luc* transcripts, was 13-fold in mature pollen. These data clearly demonstrate that the *lat*52 5'-UTR enhances gene expression in pollen primarily as a result of an increase in translational yield.

5.3.2 Developmental regulation of the translational activity of the *lat*52 5'-UTR

The *lat*52 5'-UTR specifically enhanced the translational yield of chimeric transcripts in mature pollen (fig 5.10 B). To investigate whether translational enhancement was uniform throughout pollen development or became apparent at specific developmental stages, LUC activity was determined at precise developmental stages defined by bud length (Twell *et al.*, 1993). Three independent transgenic lines harbouring the construct *lat52-SYN-luc-C3'* and three independent transgenic lines harbouring the construct *lat52-52-luc-C3'* (section 5.2.4) were selected for detailed developmental analyses. LUC activity was assayed (section 2.10.5) in extracts prepared from anthers at 15-20 developmental stages spanning microspore release to mature dehisced pollen (bud lengths 10 to ≥52 mm). The results of this analysis are presented in figure 5.11.

The developmental profiles of LUC accumulation for *lat52-SYN-luc-C3'* and

A

LEAVES	CaMV 35S	SYN 5'-UTR	Luciferase	C3'	CaMV 35S	52 5'-UTR	Luciferase	C3'
LUC activity	1,350	1,510	1,670	2,580	371	4,320	4,370	14,100
Rel. LUC activity	3.63	4.07	4.5	6.95	1.0	11.64	11.78	38.0
luc mRNA counts	64,492	61,428	150,157	57,973	32,922	172,473	99,257	355,674
rRNA counts	7,931	5,936	10,437	8,741	8,443	8,959	9,987	8,156
Rel. *luc* mRNA	8.1	10.3	14.4	6.6	3.9	19.3	9.9	43.6
Rel. translational yield	0.5	0.4	0.3	1.1	0.3	0.6	1.2	0.9
Mean TL Yield		**0.6 ± 0.4**				**0.8 ± 0.4**		
Rel. Mean TL Enhancement		**1.0**				**1.3**		

B

POLLEN	*lat*52 5'-UTR	SYN	Luciferase	C3'	*lat*52 5'-UTR	52	Luciferase	C3'
LUC activity	170	600	610	4,000	6,300	19,000	57,500	59,000
Rel. LUC activity	1.0	3.52	3.59	23.5	37.1	118.8	338.2	347.1
luc mRNA counts	6,840	5,070	6,169	4,408	7,456	21,571	83,245	81,032
rRNA counts	69,014	46,999	68,990	57,307	73,711	79,811	72,936	66,370
Rel. *luc* mRNA	0.10	0.11	0.09	0.08	0.10	0.27	1.14	0.94
Rel. translational yield	10	32	40	294	371	440	297	369
Mean TL Yield		**27.3 ± 15.5**				**354 ± 61**		
Rel. Mean TL Enhancement		**1.0**				**13.0**		

Figure 5.10. Analysis of the translational yield from chimeric transcripts synthesized in mature pollen and leaves of transgenic plants. Levels of LUC activity (light units/10 secs/mg protein x10^6) were determined for leaf (A) or pollen (B) tissue sample extracts of several independent transformants containing the gene fusion constructs *CaMV35S-SYN-luc-C3'*, *CaMV35S-52-luc-C3'* or *lat52-SYN-luc-C3'*, *lat52-52-luc-C3'* respectively. LUC activity is also shown as Relative LUC activity after normalizing to the value of the lowest individual. RNA was isolated from the remainder of the same tissue samples and northern blot analysis used to determine *luc* transcript and rRNA (after re hybridization) abundance using phosphorimagery. Arbitrary phosphorimagery counts are shown for *luc* mRNA and rRNA abundance. To account for any differences arising from RNA transfer, *luc* transcript abundance from each primary transformant was normalized to its corresponding rRNA value to give relative *luc* mRNA abundance. The relative translational yield for each individual was then calculated as the ratio of Rel. LUC activity:Rel. *luc* mRNA abundance.

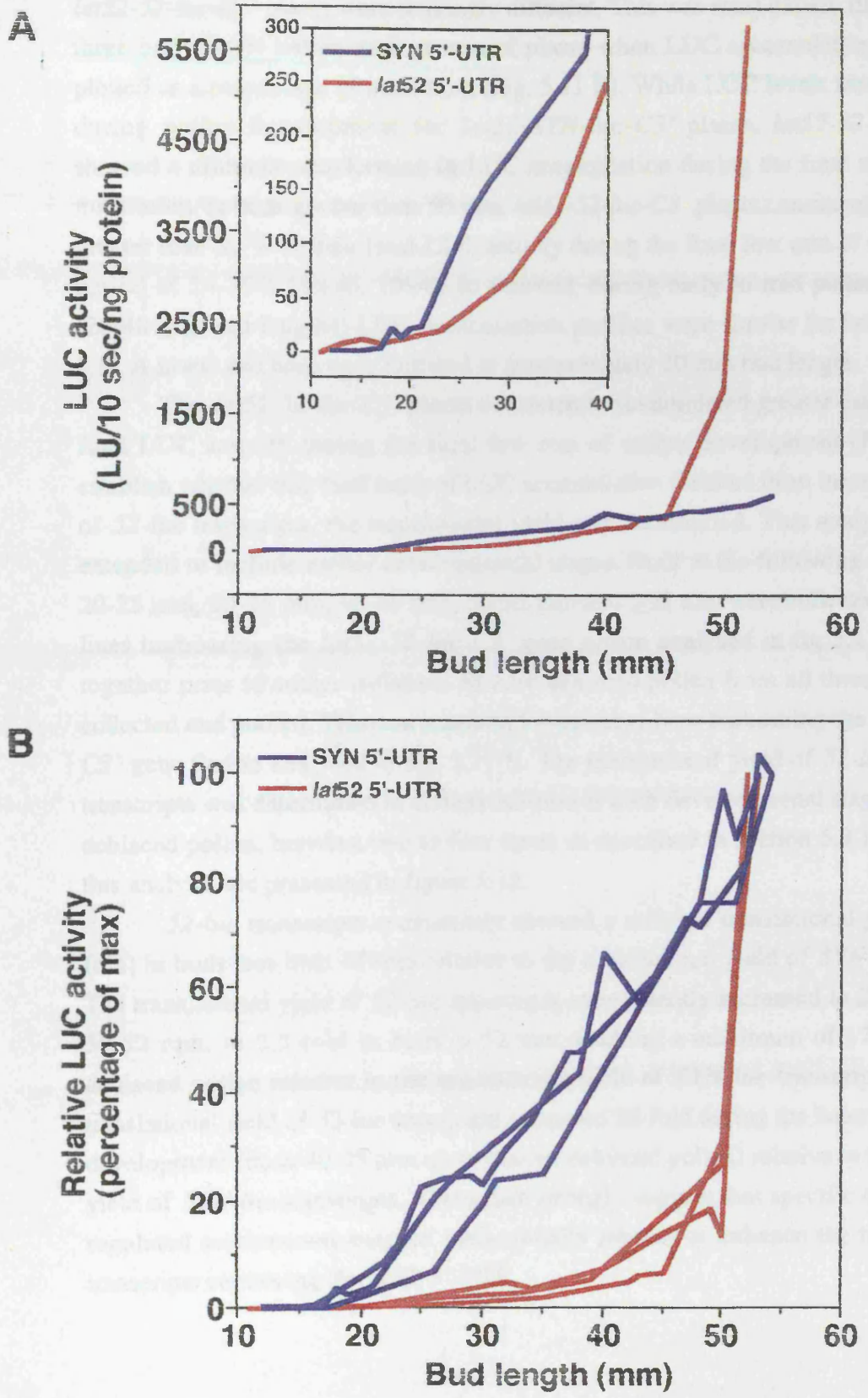


Figure 5.11. Analysis of LUC accumulation profiles during anther development in transgenic tobacco plants synthesizing SYN-LUC and 52-LUC transcripts driven by *lat*52 promoter sequences. A. LUC activity accumulation profiles for one representative SYN-LUC (blue line) and one 52 LUC (red line) plant. The inset shows the same data with an extended Y axis to illustrate similar point of initiation of LUC synthesis and similar initial rate of LUC accumulation in both plants. B. LUC accumulation profiles for three SYN-LUC (blue lines) and three 52-LUC (red lines) plants plotted as the percentage of maximum LUC activity.

lat52-52-luc-C3' plants were strikingly different. This was most clearly illustrated for all three individuals within both groups of plants when LUC accumulation profiles were plotted as a percentage of maximum (Fig. 5.11 B). While LUC levels increased steadily during pollen development for *lat52-SYN-luc-C3'* plants, *lat52-52-luc-C3'* plants showed a dramatic acceleration in LUC accumulation during the final stages of pollen maturation, in buds greater than 50 mm. *lat52-52-luc-C3'* plants consistently accumulated greater than 80 % of their final LUC activity during the final few mm of development, a period of 24-36 h (Twell, 1994). In contrast, during early to mid pollen development (20-40 mm bud lengths) LUC accumulation profiles were similar for both groups (Fig. 5.11 A inset) and both were initiated at approximately 20 mm bud length.

The *lat52-52-luc-C3'* plants consistently accumulated greater than 80 % of their final LUC activity during the final few mm of anther development (fig 5.11 B). To establish whether this final burst of LUC accumulation resulted from increased translation of *52-luc* transcripts, the translational yield was determined. This analysis was further extended to include earlier developmental stages. Buds at the following discrete stages; 20-25 mm, 30-35 mm, 40-45 mm, 50-52 mm and ≥52 mm were collected from the three lines harbouring the *lat52-52-luc-C3'* gene fusion analysed in fig 5.11 B and pooled together prior to anther isolation. Mature dehisced pollen from all three lines was also collected and pooled. This was repeated for the three lines harbouring the *lat52-SYN-luc-C3'* gene fusion analysed in fig. 5.11 B. The translational yield of *52-luc* and *SYN-luc* transcripts was determined in anthers isolated at each developmental stage and in mature dehisced pollen, between two to four times as described in section 5.3.1. The results of this analysis are presented in figure 5.12.

52-luc transcripts consistently showed a reduced translational yield (2- to 3.5-fold) in buds less than 45 mm relative to the translational yield of *SYN-luc* transcripts. The translational yield of *52-luc* transcripts subsequently increased to 2-fold in buds of 50-52 mm, to 3.5-fold in buds ≥ 52 mm reaching a maximum of 17-fold in mature dehisced pollen relative to the translational yield of *SYN-luc* transcripts. Overall, the translational yield of *52-luc* transcripts increased 36-fold during the latter stages of pollen development (buds 40-45 mm up to mature dehisced pollen) relative to the translational yield of *SYN-luc* transcripts. These data strongly suggest that specific developmentally regulated mechanisms exist to preferentially repress or enhance the translatability of transcripts containing the *lat52* 5'-UTR.

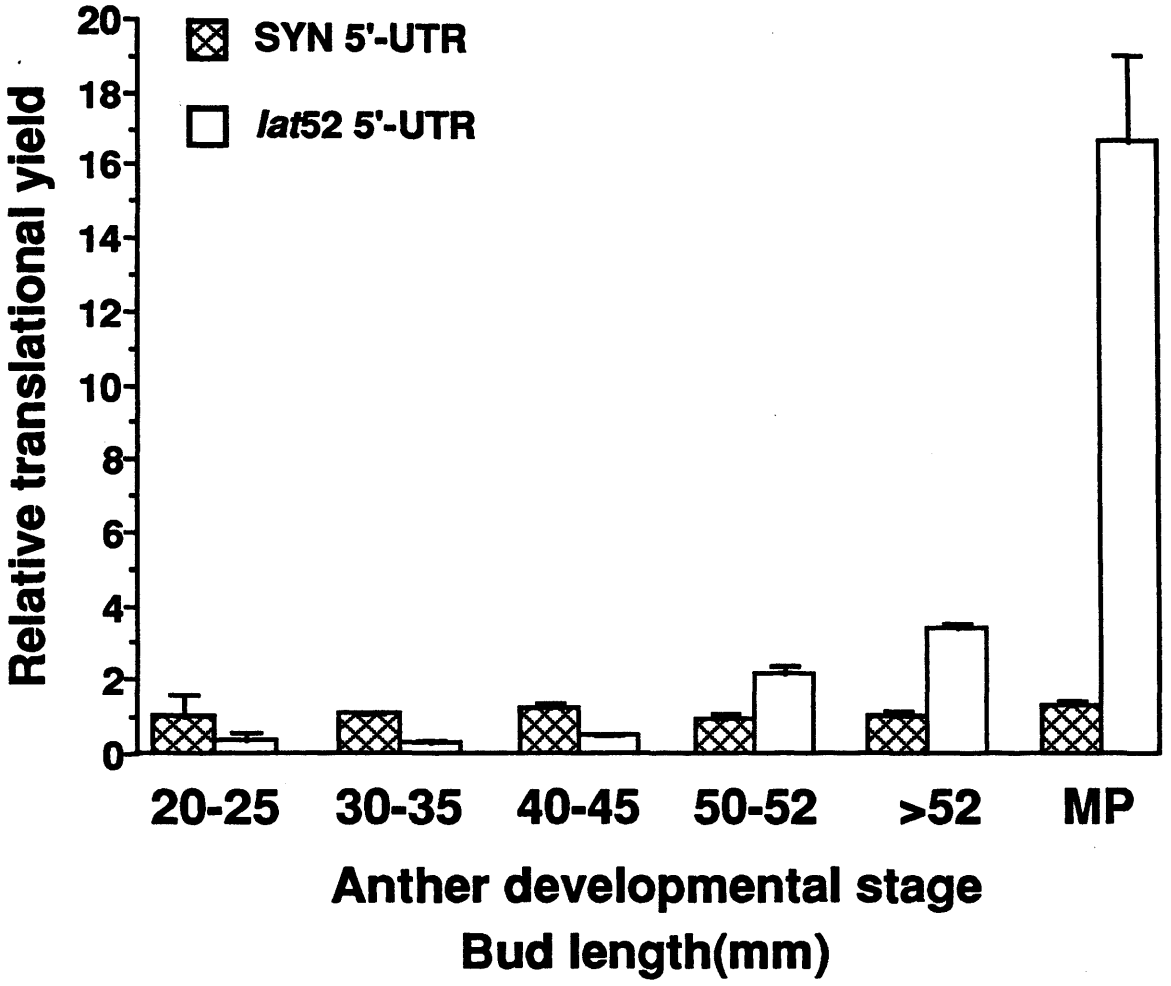


Figure 5.12. Analysis of translational yield of *SYN-luc* and *52-luc* transcripts at different stages during pollen development. Relative translational yield of *SYN-luc* and *52-luc* transcripts in anthers harvested from buds of 20-25 mm, 30-35 mm, 40-45 mm, 50-52 mm, ≥52 mm and from mature pollen (MP) was determined as described in Fig. 5.10. Each bar represents data from two to four determinations at each stage with error bars representing the standard error.

5.4 Viral and native plant 5'-UTRs enhance gene activity in pollen and leaves

5.4.1 Viral 5'-UTRs enhance gene activity in germinating pollen

Results presented in figure 5.6 A, demonstrated that the 5'-UTR from tobacco etch virus (TEV) when fused downstream of the *lat*52 promoter enhanced LUC activity 20-fold above a similar control construct containing a synthetic polylinker (SYN) in germinating pollen. Previous analysis showed that in transgenic tobacco leaves the TEV 5'-UTR when fused downstream of the *CaMV35S* promoter enhanced GUS activity 5-fold above a similar control construct (Carrington and Freed, 1990). These data would indicate that the TEV 5'-UTR preferentially enhanced gene activity in pollen. It was therefore decided to further investigate the activity of the TEV 5'-UTR in pollen and leaves and also to establish if other viral translational enhancers preferentially increased gene activity in pollen. The 5'-UTRs from alfalfa mosaic virus (AMV: Jobling and Gehrke, 1987), tobacco mosaic virus (TMV: Gallie *et al.*, 1987) and potato virus S (PVS and VTE: Turner, 1995) were inserted between the *CaMV35S* promoter and *luc* coding region to give the plasmids pRT2-AMV-LUC, pRT2-TMV-LUC and pRT2-PVS-LUC respectively. The plasmids pRT2-AMV-LUC and pRT2-TMV-LUC were constructed by M-K. Cheung. The plasmids pRT2-PVS-LUC and pRT2-VTE-LUC were constructed by R. Turner.

The test plasmids pRT2-SYN-LUC (section 5.2.2), pRT2-AMV-LUC, pRT2-TEV-LUC (section 3.3.1) pRT2-TMV-LUC, pRT2-PVS-LUC and pRT2-VTE-LUC were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5) and leaves (section 2.9.6). Transient expression analysis in leaves was performed by M-K. Cheung. Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the plasmid pRT2-SYN-LUC. The results of this analysis are presented in figure 5.13.

The 5'-UTRs AMV, TMV, PVS, TEV and VTE enhanced LUC activity in leaves by 2.5-, 4.6-, 5.1-, 4.6- and 1.3-fold respectively, and in pollen by 15.7-, 18.7-, 4.7-, 21.6 and 2.4-fold respectively. The level of enhancement detected in leaves for each viral 5'-UTR was similar to previously published results (reviewed by Turner and Foster, 1995). In pollen, the viral 5'-UTRs AMV, TMV and TEV preferentially enhanced LUC activity by 4- to 6-fold. In contrast, the PVS 5'-UTR enhanced LUC activity to the same degree in both pollen and leaves.

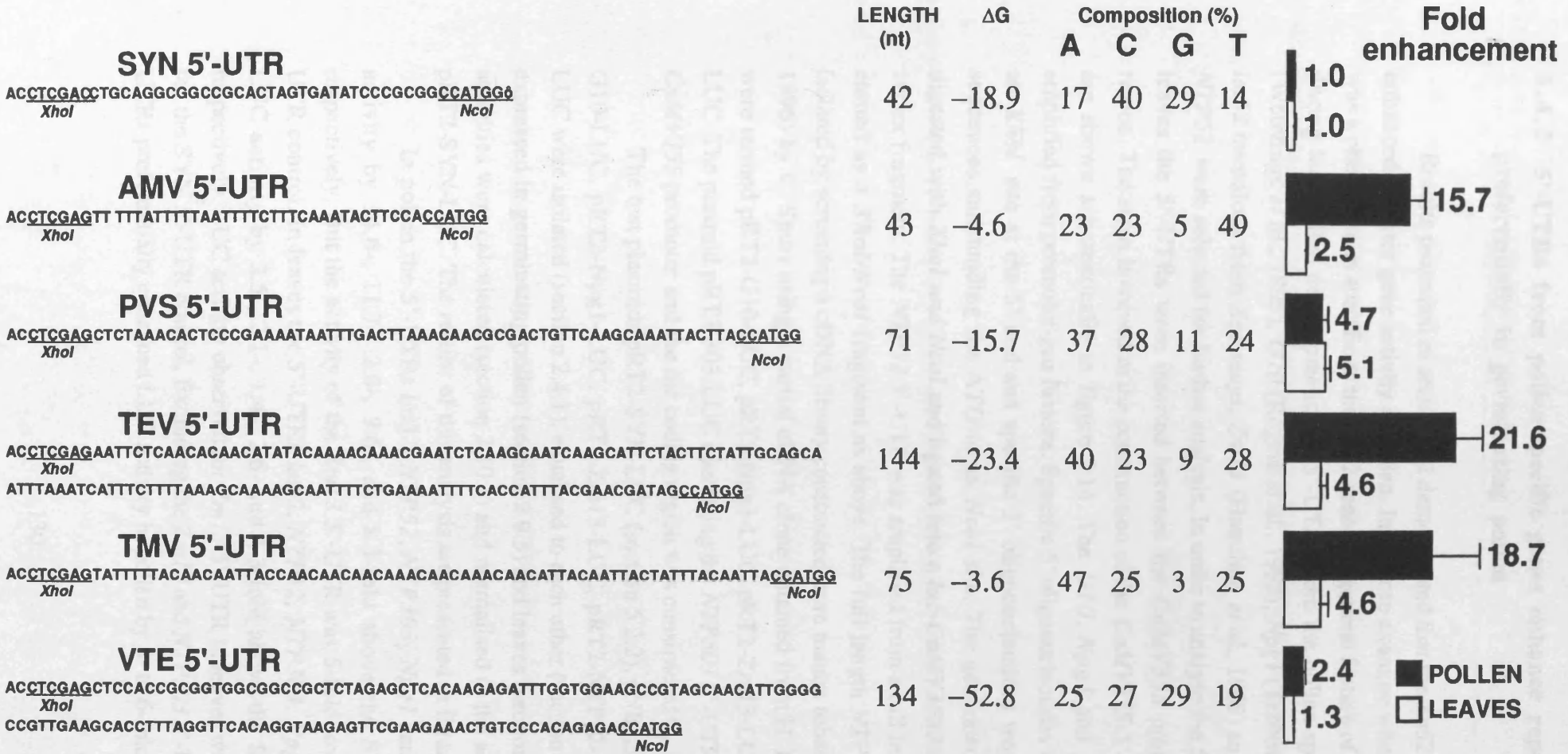


Figure 5.13. Transient expression analysis of the viral 5'-UTRs in pollen and leaves. The nucleotide sequence of each 5'-UTR beginning at +1 is shown on the left and the restriction sites used in plasmid construction and initiator ATGs are underlined. The length, in nucleotides, the calculated stability (ΔG in kcal/mol) of secondary structure and the nucleotide composition (%) of each 5'-UTR are shown in the middle. Test plasmids containing *CaMV35S-5'-UTR luc-C3'* gene fusions were co-bombarded into pollen with the reference plasmid pLAT52-7 and into leaves with the reference plasmid pRTL2GUS. Relative activities of test plasmids were calculated from a ratio of LUC:GUS normalised to the activity of the construct containing the SYN 5'-UTR (average light units per bombardment was 476144 in pollen). Data from two independent experiments are shown, representing six independent bombardments for each test plasmid. Error bars represent the standard error.

5.4.2 5'-UTRs from pollen-specific genes enhance reporter gene activity preferentially in germinating pollen

Results presented in section 5.2 demonstrated that the *lat*52 5'-UTR preferentially enhanced reporter gene activity in pollen. In order to examine whether this enhancement was a phenomenon specific to the *lat*52 gene or a general feature of other genes expressed during late pollen development, the 5'-UTRs from the pollen-specific genes *NTP303* (Weterings *et al.*, 1995), *G10* (Rogers *et al.*, 1992), *Npg1* (Tebbutt *et al.*, 1994) and the *lat*52 homologs from *Zea mays*, *Zm13* (Hamilton *et al.*, 1989) and *Nicotiana tabacum*, *NTP52* were selected for further analysis. In order to analyse the 5'-UTRs in pollen and leaves the 5'-UTRs were inserted between the *CaMV35S* promoter and *luc* coding region. The steps involved in the construction of the *CaMV35S-5'-UTR-luc-C3'* fusions are shown schematically in figure 5.14. The *G10*, *Npg1* and *Zm13* 5'-UTRs were amplified from promoter-*gus* fusions. Specific 5' oligonucleotides were used to introduce an *XhoI* site at the 5' end and specific 3' oligonucleotides were used to mutate the sequences surrounding the ATG to an *NcoI* site. The generated PCR products were digested with *XhoI and NcoI* and ligated into a *luc-CaMV35S3'-pUC19-CaMV35S5'* vector fragment. The *NTP52* 5'-UTR was amplified from a full length cDNA clone and cloned as a *XhoI/NcoI* fragment as above. The full length *NTP52* cDNA clone was isolated by screening a cDNA library constructed from mature tobacco pollen (Sweetman, 1996) by C. Spurr using a partial cDNA clone obtained from H. Rogers. The plasmids were termed pRT2-G10-LUC, pRT2-Npg1-LUC, pRT2-Zm13-LUC and pRT2-NTP52-LUC. The plasmid pRT2-303-LUC containing the *NTP303* 5'-UTR inserted between the *CaMV35S* promoter and the *luc* coding region was constructed by K. Weterings.

The test plasmids pRT2-SYN-LUC (section 5.2.2), pNBL52-26 (section) pRT2-G10-LUC, pRT2-Npg1-LUC, pRT2-Zm13-LUC, pRT2-NTP52-LUC and pRT2-303-LUC were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5) and leaves (section 2.9.6). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the plasmid pRT2-SYN-LUC. The results of this analysis are presented in figure 5.15.

In pollen the 5'-UTRs *lat*52, *NTP52*, *NTP303*, *Npg1* and *G10* enhanced LUC activity by 14.6-, 11.7-, 2.8-, 9.6-, and 8.3-fold above the SYN 5'-UTR control respectively. But the activity of the *Zm13* 5'-UTR was 5-fold lower than the SYN 5'-UTR control. In leaves the 5'-UTRs *lat*52, *NTP52*, *NTP303*, *Npg1* and *G10* enhanced LUC activity by 2.5-, 2.1-, 1.4-, 2.6- and 1.3-fold above the SYN 5'-UTR control respectively. LUC activity observed for *Zm13* 5'-UTR in leaves was similar to that seen for the SYN 5'-UTR control. Excluding the *Zm13* and *NTP303* 5'-UTRs all the other 5'-UTRs preferentially enhanced LUC activity in pollen by 4- to 6-fold.

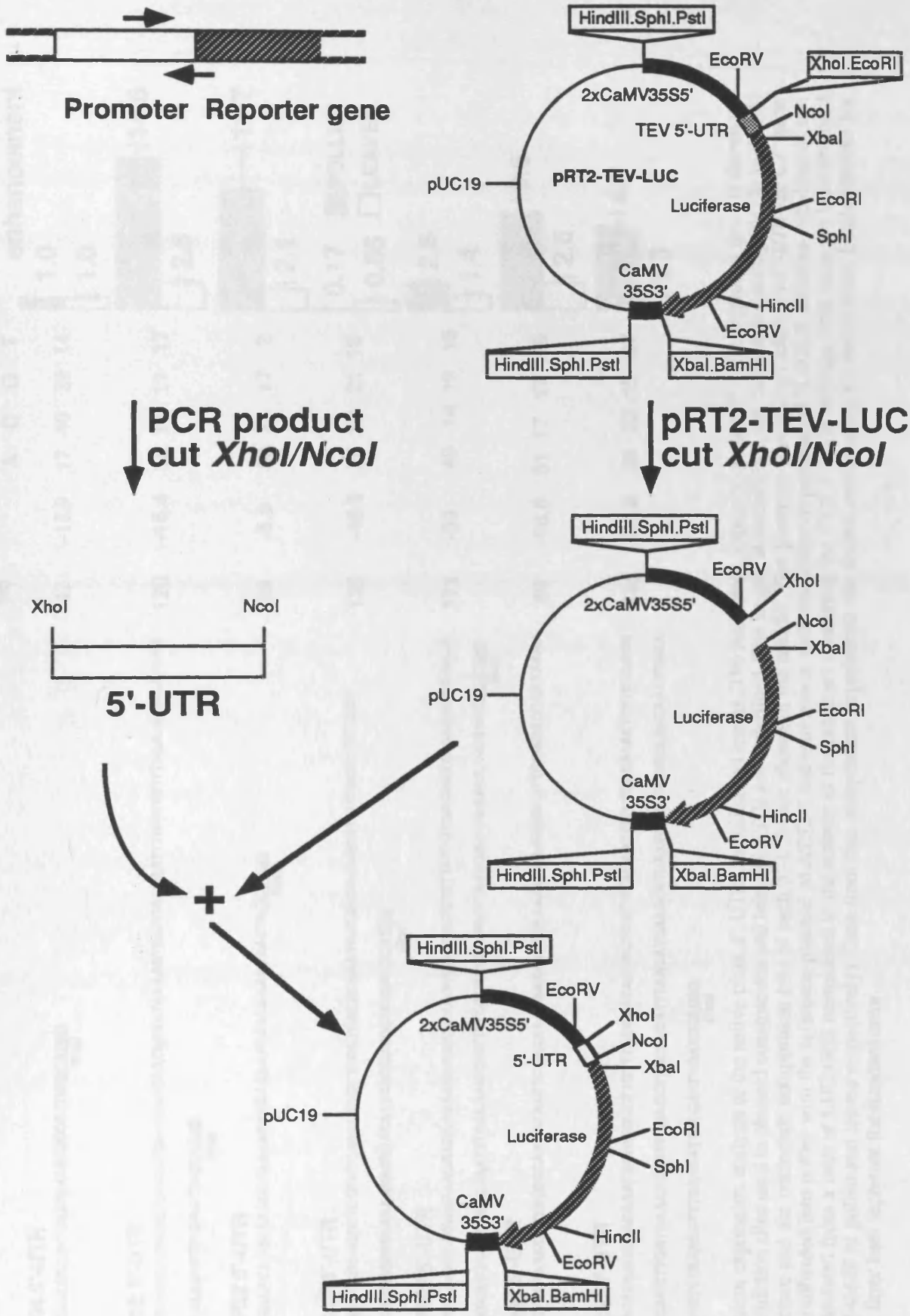


Figure 5.14. Construction of the *CaMV35S-5'-UTR-luc-C3'* gene fusions. The *G10*, *Npg1* and *Zm13* 5'-UTRs were amplified from promoter-*gus* fusions by PCR cut with *XhoI and NcoI* and cloned into pRT2-TEV-LUC displacing the TEV 5'-UTR.

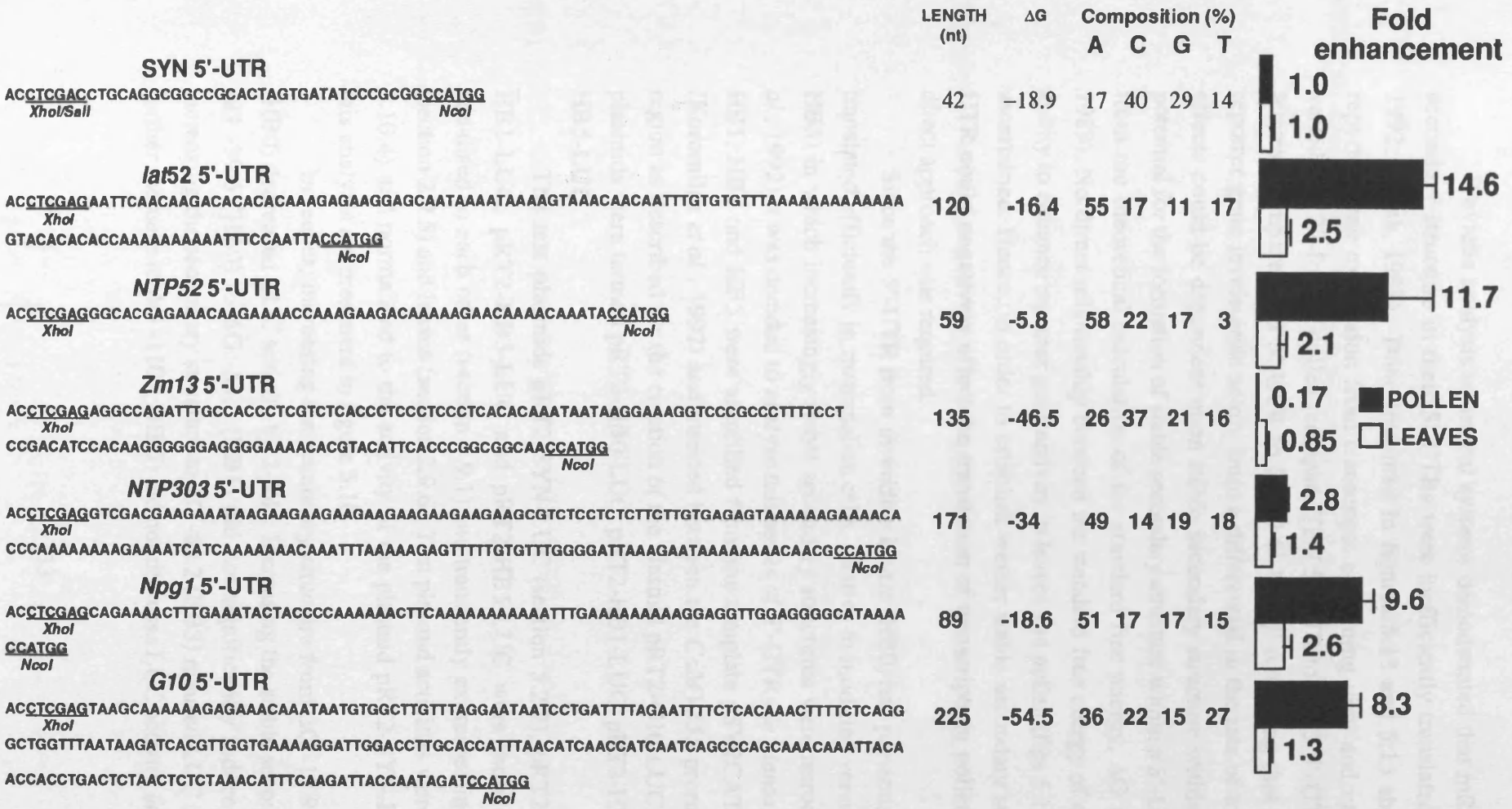


Figure 5.15. Transient expression analysis of the native plant 5'-UTRs in pollen and leaves. The nucleotide sequence of each 5'-UTR beginning at +1 is shown on the left and the restriction sites used in plasmid construction and initiator ATGs are underlined. The length, in nucleotides, the calculated stability (ΔG in kcal/mol) of secondary structure and the nucleotide composition (%) of each 5'-UTR are shown in the middle. Test plasmids containing *CaMV35S-5'-UTR-luc-C3'* gene fusions were co-bombarded into pollen with the reference plasmid pLAT52-7 and into leaves with the reference plasmid pRTL2GUS. Relative activities of test plasmids were calculated from a ratio of LUC:GUS normalised to the activity of the construct containing the SYN 5'-UTR (average light units per bombardment was 836005 and 146628 in pollen and leaves respectively). Data from two independent experiments are shown, representing six independent bombardments for each test plasmid. Error bars represent the standard error.

5.4.3 Translation in pollen is strongly inhibited by stable secondary structures within the 5'-UTR

Previous analysis in animal systems demonstrated that mRNAs with extensive secondary structure in their 5'-UTRs were inefficiently translated (Koromilas *et al.*, 1992: Kozak, 1988). Data presented in figures 5.13 and 5.15 showed that levels of reporter gene expression from transcripts containing viral and native plant 5'-UTRs varied by 127-fold in pollen (compare TEV 5'-UTR to *Zm13* 5'-UTR) and by 6-fold in leaves (compare PVS 5'-UTR to *Zm13* 5'-UTR). Assuming that these differences in reporter gene levels arise solely from a differential in the rate of translation then these effects could be dependent upon stable secondary structure within each 5'-UTR. The potential for the formation of stable secondary structure within a 5'-UTR can be predicted from the theoretical calculation of the standard free energy, ΔG in Kcal/mol (Zuker, 1989). No direct relationship between the standard free energy of each 5'-UTR and its ability to enhance reporter gene activity in leaves and pollen (figs 5.13 and 5.15) could be ascertained. Hence, in order to establish wether stable secondary structure within a 5'-UTR could negatively effect the translation of transcripts in pollen and leaves a more direct approach was required.

Since the 5'-UTR from *thymidine kinase* (HB0) had previously been shown to be translated efficiently in mammalian cells, relative to modified versions (HB1, HB3 and HB5) in which increasingly stable secondary structures were introduced (Koromilas *et al.*, 1992) it was decided to analyse this series of 5'-UTRs *in planta*. The 5'-UTRs HB0, HB1, HB3 and HB5 were amplified from the template pSV2CAT and its derivatives (Koromilas *et al.*, 1992) and inserted between the *CaMV35S* promoter and *luc* coding region as described for the creation of the plasmid pRT2-G10-LUC (section 5.4.2). The plasmids were termed pRT2-HB0-LUC, pRT2-HB1-LUC, pRT2-HB3-LUC and pRT2-HB5-LUC.

The test plasmids pRT2-SYN-LUC (section 5.2.2), pRT2-HB0-LUC, pRT2-HB1-LUC, pRT2-HB3-LUC and pRT2-HB5-LUC were isolated (section 2.4.1), equalised to each other (section 2.9.1) and transiently expressed in germinating pollen (section 2.9.5) and leaves (section 2.9.6). Test plasmid activities were calculated (section 2.10.4) and normalised to the activity of the plasmid pRT2-SYN-LUC. The results of this analysis are presented in figure 5.16.

In leaves, increasing the secondary structure from ΔG -18.9 (SYN) to ΔG -39.5 (HB0) decreased LUC activity by 2-fold. Increasing the stable secondary structure from ΔG -39.5 (HB0) to ΔG -54.3 (HB1) did not significantly reduce LUC activity. But increasing the secondary structure to ΔG -82.2 (HB3) reduced LUC activity by 3-fold. A further increase to ΔG -110.1 (HB5) did not decrease LUC activity any further. An

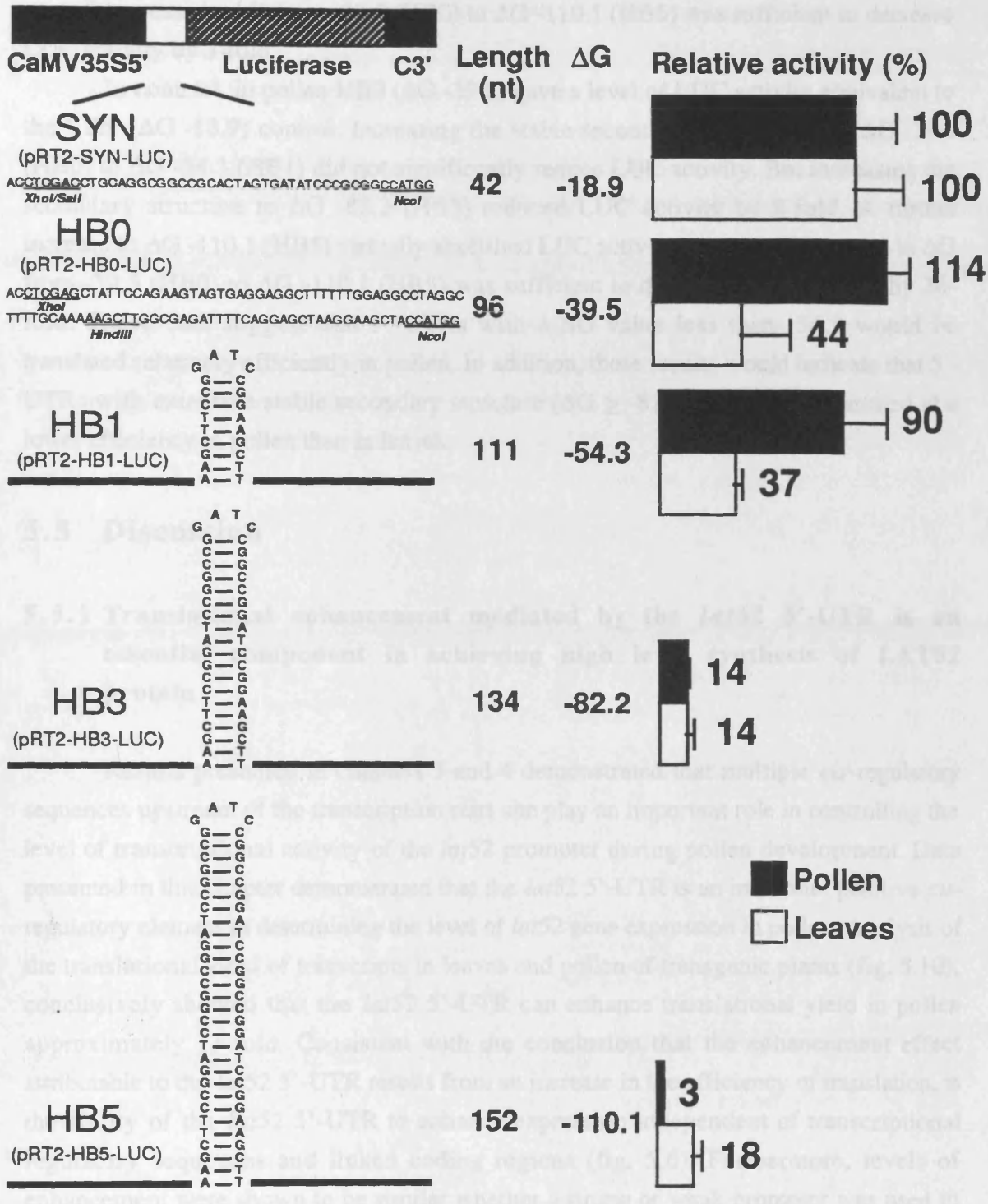


Figure 5.16. Analysis of increasing stable secondary structure on reporter gene activity in pollen and leaves. The nucleotide sequence of the HB0 5'-UTR beginning at +1 is shown on the left. The restriction sites used in plasmid construction are underlined. Also shown are the predicted stable stem loop structures created from the introduction of *BamHI* linkers into the *HindIII* site of the HB0 5'-UTR. The length, in nucleotides and the calculated stability (ΔG in kcal/mol) of secondary structure within the 5'-UTRs are shown in the middle. Test plasmids containing *CaMV35S-5'-UTR-luc-C3'* gene fusions were co-bombarded into pollen with the reference plasmid pLAT52-7 and into leaves with the reference plasmid pRTL2GUS. Relative activities of test plasmids were calculated from a ratio of LUC:GUS normalised to the activity of the construct containing the SYN 5'-UTR (average light units per bombardment was 360475 and 782338 in pollen and leaves respectively). Data from two independent experiments are shown, representing six independent bombardments for each test plasmid. Error bars represent the standard error.

overall increase in ΔG from -39.5 (HB0) to ΔG -110.1 (HB5) was sufficient to decrease LUC activity by 3-fold.

In contrast, in pollen HB0 (ΔG -39.5) gave a level of LUC activity equivalent to the SYN (ΔG -18.9) control. Increasing the stable secondary structure from ΔG -39.5 (HB0) to ΔG -54.3 (HB1) did not significantly reduce LUC activity. But increasing the secondary structure to ΔG -82.2 (HB3) reduced LUC activity by 8-fold. A further increase to ΔG -110.1 (HB5) virtually abolished LUC activity. An overall increase in ΔG from -39.5 (HB0) to ΔG -110.1 (HB5) was sufficient to decrease LUC activity by 36-fold. These data suggest that 5'-UTRs with a ΔG value less than -54.3 would be translated relatively efficiently in pollen. In addition, these results would indicate that 5'-UTRs with extensive stable secondary structure (ΔG $\geq$ -82.2) would be translated at a lower efficiency in pollen than in leaves.

5.5 Discussion

5.5.1 Translational enhancement mediated by the *lat*52 5'-UTR is an essential component in achieving high level synthesis of LAT52 protein

Results presented in chapters 3 and 4 demonstrated that multiple *cis*-regulatory sequences upstream of the transcription start site play an important role in controlling the level of transcriptional activity of the *lat*52 promoter during pollen development. Data presented in this chapter demonstrated that the *lat*52 5'-UTR is an important positive *cis*-regulatory element in determining the level of *lat*52 gene expression in pollen. Analysis of the translational yield of transcripts in leaves and pollen of transgenic plants (fig. 5.10), conclusively showed that the *lat*52 5'-UTR can enhance translational yield in pollen approximately 13-fold. Consistent with the conclusion that the enhancement effect attributable to the *lat*52 5'-UTR results from an increase in the efficiency of translation, is the ability of the *lat*52 5'-UTR to enhance expression independent of transcriptional regulatory sequences and linked coding regions (fig. 5.6). Furthermore, levels of enhancement were shown to be similar whether a strong or weak promoter was used to drive expression, which suggests that transcripts containing the *lat*52 5'-UTR, are translated more efficiently, regardless of their abundance. The transient expression experiments, utilizing pollen germinated *in vitro*, demonstrated the ability of the *lat*52 5'-UTR to enhance gene expression during pollen tube growth (fig. 5.6), and the transgenic analyses demonstrated enhancement during pollen maturation (fig. 5.9). Thus, the mechanisms which mediate such enhancement appear to operate both prior to, and following pollen germination. Regardless of the mechanisms involved, these data

demonstrate that the *lat*52 5'-UTR is an important component in controlling the high level of LAT52 protein synthesised during pollen development and pollen tube growth (Villalba *et al.*, 1993; Muschietti *et al.*, 1994).

5.5.2 Translational enhancement mediated by the *lat*52 5'-UTR is not a result of length, optimal translational context and lack of predicted secondary structure

Prior to discussing the mechanisms which increase the translational yield of transcripts containing the *lat*52 5'-UTR it is important to dismiss the one possibility that could falsely indicate the presence of a translational enhancer within the *lat*52 5'-UTR. This possibility is based upon the synthetic control 5'-UTRs (SYN and HB0) being translated with an artificially lower efficiency. If this were the case then the *lat*52 5'-UTR would be translated with significantly greater efficiency in all tissue types and at all stages during pollen development, which was not observed.

With regard to the mechanisms which increase the translational yield of transcripts containing the *lat*52 5'-UTR in pollen several possibilities exist. Perhaps the most simple explanation is a structural one in which the low stability of secondary structure present within the *lat*52 5'-UTR would allow efficient progress of scanning ribosomes to the initiator AUG. Calculation of the stability of secondary structure within the *lat*52 and SYN 5'-UTRs tested showed that none had predicted ΔG values above -35.4 kcal/mol (fig. 5.1). Data presented in figure 5.16 clearly demonstrated that in pollen the construct containing the HB0 5'-UTR (ΔG -39.5 kcal/mol) was expressed as efficiently as the construct containing the SYN 5'-UTR (ΔG -18.9 kcal/mol). It was not until a ΔG value of -82.2 kcal/mol (HB3) was reached that a significant 8-fold reduction in gene expression was observed. Since an increase in ΔG from -18.9 (SYN) to -82.2 (HB3) kcal/mol is required to decrease gene expression by 8-fold in pollen, it would be extremely unlikely that an increase in ΔG from -16.4 (*lat*52) to -18.9 (SYN) kcal/mol could account for the 14-fold difference in expression levels observed between the SYN and *lat*52 5'-UTRs in pollen (fig. 5.6).

Other factors that could influence the translational yield of transcripts containing the *lat*52 5'-UTR in pollen are 5'-UTR length and the nucleotide sequence (context) surrounding the initiator AUG. Increasing the length of the 5'-UTR from 42 nucleotides (SYN) to 96 nucleotides (HB0) did not affect the level of reporter gene expression in pollen (fig 5.16). This suggests that the length of the 5'-UTR does not influence gene expression levels. Previous analysis in a mammalian system demonstrated that the nucleotide sequence surrounding the initiator AUG was an important determinant of the level of gene expression (Kozak, 1986). This analysis defined ACCAUGG as the

optimal sequence for initiation by eukaryotic ribosomes, and also identified the nucleotides at positions -3 (A) and +4 (G) as essential for high level gene expression (Kozak, 1986). In addition, Lonsdale *et al.* (1995) demonstrated that a G residue at position +4 was required for high level gene expression *in planta*. The 5'-UTRs analysed in this chapter were fused to the *luc* coding region using a *NcoI* site (CCATGG), the only difference observed between the optimal translational context was the nucleotide at -3. Since analysis of transcripts with mutations at -3 revealed no differences in translational efficiency in an *in vitro* wheat germ system (Lutcke *et al.*, 1987) it is unlikely that translational context, as defined by Kozak, 1986 is a significant factor in accounting for the observed differences in translational yield.

Taken together, these results indicate that the mechanisms mediating translational enhancement by the *lat*52 5'-UTR in pollen are not simply the result of its length, optimal translational context and lack of predicted secondary structure. A further possibility is that the rate of translational elongation is increased by the *lat*52 5'-UTR. Although this cannot be ruled out, the efficiency of the enhancement mediated by the *lat*52 5'-UTR was similar for different coding regions, and it is difficult to conceive of long range interactions between the assembled 80S ribosomes and the 5'-UTR which could affect the rate of elongation.

5.5.3 Viral and native plant 5'-UTRs act as translational enhancers of gene expression in pollen

Data presented in figure 5.13 demonstrated that viral 5'-UTRs enhanced reporter gene activity in pollen and in leaves. Previous analyses clearly demonstrated that the viral 5'-UTRs enhanced reporter gene expression in leaves by increasing the translational yield of chimeric transcripts (reviewed by Turner and Foster, 1995). The viral 5'-UTRs originate from single-stranded positive-sense RNA viruses (Turner and Foster, 1995), which replicate using viral encoded RNA-dependent RNA polymerases in the cytoplasm (Allison *et al.*, 1986). It is therefore difficult to conceive that the viral 5'-UTRs contain pollen-specific transcriptional *cis*-regulatory elements which are recognised by DNA dependent transcription factors. Therefore, the enhancement of reporter gene activity in pollen displayed by the viral 5'-UTRs (fig 5.13) is most likely a result of increased translational yield and not due to an increase in transcription rate.

Previous analysis demonstrated that the translational enhancement mediated by the TMV 5'-UTR was dependent upon a 25 bp poly (CAA) region (Gallie and Walbot, 1992). This sequence motif is not present within either the *lat*52 5'-UTR (fig 5.15) or the viral 5'-UTRs (fig. 5.13). Even though single CAA motifs are present within the *lat*52 5'-UTR (fig. 5.15) and the viral 5'-UTRs (fig. 5.13), there appears to be no direct

correlation between the number of CAA repeats within a 5'-UTR and its ability to enhance reporter gene activity (compare AMV to VTE, fig. 5.13). Further sequence analysis revealed no conserved motifs within all the 5'-UTRs, but indicated that sequence composition alone was sufficient to regulate the level of gene expression.

The 5'-UTRs containing a high percentage of A and T residues ($\geq$ 68 %) enhanced reporter gene activity in pollen by at least 5-fold (fig.s 5.6 and 5.13). Increasing the percentage of C and G residues within a 5'-UTR progressively decreased the level of enhancement seen in pollen (compare PVS, VTE and SYN fig. 5.13). The decrease in reporter gene activity in pollen observed for 5'-UTRs with an increased percentage of C and G residues could arise from more stable 5'-UTR/coding region interactions. This is unlikely because the *luc* coding region contains a greater proportion of A and T residues, this would therefore favour stable interactions between the A and T rich 5'-UTRs and the coding sequence. In addition, no sequences were found within the luciferase coding region which could potentially base pair with sequences within each 5'-UTR. The *lat*52 and viral 5'-UTRs contained a similar percentage of C residues, but the 5'-UTRs that enhanced reporter gene activity in pollen by at least 5-fold (fig.s 5.4 and 5.13) contained $\leq$ 11 % G residues. This implies that G residues within a 5'-UTR could inhibit the rate of translation in pollen.

Results presented in figure 5.15 showed that the native plant 5'-UTRs *NTP6*, *G10*, *NTP303* and *Npg1* enhanced reporter gene activity in pollen and leaves. The level at which the 5'-UTRs enhance reporter gene activity cannot be established from the bombardment data alone. Enhancement could arise from either an increase in the rate of transcription or translation. The presence of positive transcriptional *cis*-regulatory elements within each 5'-UTR cannot be ruled out, but a combination of leaf-specific and pollen-specific, or constitutive elements would be required in order to explain the observed results.

Vice versa, these results could equally be due to an enhanced level of translation. The above analysis of the *lat*52 and viral 5'-UTRs suggested that 5'-UTRs which contain a high percentage of A and T residues are translated more efficiently in pollen than 5'-UTRs which contain a high percentage of C and G residues, and also that G residues could have a negative effect on translation rate. If these sequence constraints are applied to the native plant 5'-UTRs then it would further suggest that the enhancement seen for each 5'-UTR occurred at the level of translation. The *NTP6*, *G10*, *NTP303* and *Npg1* 5'-UTRs all contained a high percentage of A and T residues (61-66 %) and enhanced reporter gene activity in pollen. In contrast, the *Zm13* 5'-UTR contained a high percentage of C and G residues (59 %) and was translated inefficiently in pollen.

The percentage of C and G residues within the *Zm13* 5'-UTR (59 %) was lower than that within the SYN 5'-UTR (69 %), this would suggest that the *Zm13* 5'-UTR

would be translated more efficiently than the SYN 5'-UTR. This was not the case, the SYN 5'-UTR was translated 5 times more efficiently than the *Zm13* 5'-UTR in pollen (fig. 5.15). This result may be explained on the basis of total G residues. The SYN 5'-UTR would be translated more efficiently because it contains 12 G residues, with no more than 2 consecutive Gs, compared to the *Zm13* 5'-UTR which contains more G residues (28), of which 10 are contained within an 11 bp stretch. This indicates that not only the number but the proximity of G residues to each other within a 5'-UTR can be detrimental to the rate of translation.

In summary, the enhancement mediated by the viral 5'-UTRs is most likely due to an increase in translation rate and suggests a common mechanism of translational enhancement in pollen. No direct evidence is available to determine the level at which native plant 5'-UTRs enhance reporter gene activity, but based upon their sequence composition it is proposed that they act at the level of translation. This can only be clarified by determining the translation yield of chimeric 5'-UTR transcripts *in planta.*

5.5.4 Preferential translational enhancement in pollen displayed by the viral and native plant 5'-UTRs could result from their reduced requirement for the cap-binding protein eIF-4E

Translational enhancement mediated by viral and plant 5'-UTRs in pollen appears to be regulated by a common mechanism (section 5.5.3). Inherent within this mechanism is the ability to discriminate between 5'-UTRs based solely upon sequence composition, such that 5'-UTRs with a high content of A and T residues are translated more efficiently. How this is achieved is unclear but it could involve the use of specific RNA-binding proteins which are able to increase the efficiency of translational initiation of mRNAs containing A and T rich 5'-UTRs. Tanguay and Gallie, (1996) isolated a 102 KDa RNA-binding protein that binds to the poly (CAA) region within the TMV 5'-UTR and to the upstream pseudoknot domain within the TMV 3'-UTR. The authors also showed that this protein was widely conserved in plant species and demonstrated that depletion of this protein from an *in vitro* translation system inhibited the translation of all mRNAs tested to a similar extent. Even though this analysis would argue against the presence of sequence specific RNA-binding proteins which increase the efficiency of translational initiation of mRNAs containing A and T rich 5'-UTRs it does not rule out the existence of such proteins *in planta.*

In the absence of sequence specific RNA-binding proteins it is possible that A and T rich 5'-UTRs preferentially utilise general translation initiation factors. Previous analysis in animal systems demonstrated that mRNAs with extensive secondary structure in their 5'-UTRs were inefficiently translated (Koromilas *et al.*, 1992; Kozak, 1988;

1989). Overexpression of the cap-binding protein eIF-4E in an animal system stimulated the translation of mRNAs with extensive secondary structure in their 5'-UTRs (Koromilas *et al.*, 1992). Results presented in figure 5.16 showed that as secondary structure within the thymidine kinase 5'-UTR was increased, gene expression levels decreased dramatically in pollen but not in leaves. This could indicate a higher concentration of the cap-binding protein eIF-4E in leaves than in pollen. Consistent with this hypothesis is the fact that the expression of an eIF-4E-related gene is down-regulated after pollen mitosis I (Combe and Twell, unpublished). The down regulation of the eIF-4E gene is not a general characteristic of all translation initiation factors, since an eIF-4A-related gene has been isolated which is expressed specifically in pollen and is strongly up-regulated after pollen mitosis I (Brander and Kuhlemeier, 1995). Furthermore, previous analysis demonstrated that the TMV 5'-UTR showed a reduced requirement for the cap-binding protein eIF-4E, but was still dependent on eIF-4A (Altman *et al.*, 1990). If this is a common attribute of all A and T rich 5'-UTRs, then the preferential translational enhancement in pollen displayed by the viral and native plant 5'-UTRs (fig.s 5.13 and 5.15) could result from their reduced requirement for the cap-binding protein eIF-4E.

5.5.5 The translational yield of *52-luc* transcripts may be developmentally regulated by controlling the access of *52-luc* transcripts to the general translation machinery

The data presented in figure 5.12 demonstrated that transcripts containing the *lat*52 5'-UTR were translated differentially during pollen development. During the early stages of pollen development (bud length $\leq$ 45 mm) the translational yield of *52-luc* transcripts was 2- to 3.5-fold lower than the translational yield of *SYN-luc* transcripts. The translational advantage of *52-luc* transcripts became apparent specifically during the final stages of pollen development (bud length >50 mm). This analysis does not preclude that the translational advantage displayed by the *52-luc* transcripts during the final stages of pollen development resulted from a lower translation yield of *SYN-luc* transcripts. Data presented in figure 5.11 showed that during early to mid pollen development (20-40 mm bud lengths) LUC accumulation profiles were similar for both *lat52-SYN-luc-C3'* plants and *lat52-52-luc-C3'* plants. However, during the final stages of pollen maturation *lat52-52-luc-C3'* plants showed a dramatic acceleration in LUC accumulation, while LUC levels increased steadily for *lat52-SYN-luc-C3'* plants. Taken together, these data strongly suggest that differences seen in translational yield between *52-luc* and *SYN-luc* transcripts arise from the differential translation of *52-luc* transcripts and not *SYN-luc* transcripts. Therefore, developmental mechanisms exist to positively and negatively

regulate the translation of transcripts containing the *lat*52 5'-UTR.

During pollen development the major period of mRNA accumulation occurs after pollen-mitosis I which rapidly increases up to anthesis (Tupy, 1983). A significant fraction of mRNAs synthesised during this period are stored until required during early pollen germination (reviewed by Mascarenhas, 1990). Perhaps the most simple explanation regarding the mechanisms which regulate the translation of transcripts containing the *lat*52 5'-UTR is based upon access of transcripts to the translation machinery. During early pollen development transcripts containing the *lat*52 5'-UTR could be sequestered and stored as cytoplasmic ribonucleoprotein particles. When LAT52 protein is required stored transcripts would be made available to the translation machinery. The accessibility of transcripts containing the *lat*52 5'-UTR to the translation machinery would provide a fine tuning mechanism to ensure that LAT52 protein only accumulated at the correct (final) stages of pollen maturation. Indeed, it may be detrimental for pollen to allow transcripts such as *lat*52, which contain translational enhancers access to the translation machinery at early developmental stages. For example, since *lat*52 contains a strong translational enhancer, protein synthesis would be diverted into the synthesis of LAT52 at the expense of other important proteins, such as transcription factors.

Chapter 6

General discussion

6.1 The tomato *lat*52 gene is developmentally regulated in a tri-phasic manner in tobacco

At the onset of this thesis the regulation of genes expressed during pollen development was understood in terms of a rather simplistic model. Developmental northern blot analyses of the pollen-specific genes *lat*52 (Twell *et al.*, 1989b), *lat*59 (Wing *et al.*, 1989), *NTP303* (Weterings *et al.*, 1992), *G10* (Rogers *et al.*, 1992) and *Npg1* (Tebbutt *et al.*, 1994) showed that transcript was first detectable at or around PMI, with a continuous increase in transcript levels throughout pollen development up to anthesis. From these analyses it was proposed that transcript levels were controlled by positive transcriptional *cis*-regulatory elements, activated at PMI and utilised throughout pollen development.

Data presented in chapter 3 showed that in transgenic tobacco 93 % of *lat*52 promoter activity in pollen was accounted for by the region -492 to -101 (domains A+B). Analysis of the accumulation of GUS activity during anther development demonstrated that the region -492 to -101 strongly enhanced *lat*52 promoter activity during the early stages of pollen maturation (chapter 3). These results highlight the significant contribution of domains A+B together in achieving maximal levels of *lat*52 promoter activity during pollen development. Dissection of the region -492 to -52 using a gain of function approach demonstrated that domain A or B alone could activate a heterologous core promoter in a pollen-specific manner (chapter 4). Domain B contains the PBII motif. Previous analysis demonstrated that a fragment (-194 to -176) containing PBII activated a heterologous core promoter (Twell *et al.*, 1991). Results presented in chapter 4 suggest that in the presence of downstream sequences (-176 to -101) PBII (sub-domain B1) is redundant in germinating pollen since mutation or 5' deletion of PBII did not affect the pollen-specific activity of domain B.

Analysis of the accumulation of GUS activity in transgenic tobacco during anther development demonstrated that the region -100 to +110 strongly enhanced *lat*52 promoter activity during the mid to late stages of pollen maturation (chapter 3). During the early stages of pollen maturation this region was virtually inactive (chapter 3). Deletion analysis of the *lat*52 promoter in the transient assay demonstrated that domains B2, B3 and C together (-176 to +110) accounted for 84 % of *lat*52 promoter activity in germinating pollen (chapter 4).

The data discussed above shows that transcription of the *lat*52 gene is developmentally regulated during pollen maturation. Analyses presented in chapter 5 demonstrated that transcripts containing the *lat*52 5'-UTR were translated differentially during pollen development. During the early stages of pollen development (bud length $\leq$ 45 mm) the translational yield of *52-luc* transcripts was 2- to 3.5-fold lower than the

translational yield of *SYN-luc* transcripts. The translational advantage of *52-luc* transcripts became apparent specifically during the final stages of pollen development (bud length >50 mm).

Taken together, these data suggest that *lat*52 expression is developmentally regulated in a tri-phasic manner in tobacco. In the first phase, it is proposed that the *lat*52 promoter utilises domains A+B1 together to activate transcription at or around PMI, with a subsequent stimulation in activity up to mid-pollen development. During the second phase of development (mid-pollen to germination) domains B2+B3+C together are activated while domains A+B1 together become redundant. Both the first and second phases ensure the rapid and continuous accumulation of *lat*52 transcript from PMI up to and including pollen germination. The third phase of developmental control occurs at the level of translation. This ensures that *lat*52 transcript is efficiently translated only at the very latter stages of pollen development and during germination/pollen tube growth.

The combination of transcriptional and post-transcriptional control of gene activity is not restricted to pollen-specific genes such as *lat*52. Previous analysis demonstrated that the accumulation of *ß-phaseolin* transcript during seed development is regulated by positive *cis*-regulatory elements and at least one temporal element (Burow *et al.*, 1992). Post-transcriptional regulation of *ß-phaseolin* transcript stability has been shown to affect the level of PHASEOLIN accumulation (Chapell and Chrispeels, 1986). Other examples include the barley *α-amylase* gene which is transcribed specifically in aleurone layers during seed development (Gallie and Young, 1994, references therein). RNA transfection studies have demonstrated that the *α-amyalse* 5' and 3'-UTRs together, enhanced reporter gene activity specifically in maize aleurone and endosperm protoplasts (Gallie and Young, 1994). During hypoxia alcohol dehydrogenase-1 (*adh1*) and glyceraldehyde-3-phosphate dehydrogenase (*gapc3*) are transcriptionally enhanced and their transcripts show increased stability and enhanced translation (Fennoy and Bailey-Serres, 1995, references therein).

At present, *lat*52 represents the only documented example of a pollen-specific gene developmentally regulated in a tri-phasic manner using a combination of transcriptional and translational mechanisms. Since no developmental 5' promoter deletion analyses in transgenic plants have been reported for the *lat59*, *NTP303*, *G10* and *Npg1* promoters it is difficult to determine if the same transcriptional *cis*-regulatory sequences within a given promoter are utilised throughout pollen development, or like *lat*52 two distinct sets are used. The developmentally regulated transcription of the *lat*52 promoter was first observed from the differential in the activities of 5' promoter deletion mutants in transgenic plants and the transient assay (chapter 3). If this same criteria is applied to the pollen-specific genes where appropriate data is available, then it could indicate whether other pollen-specific genes are transcriptionally regulated during pollen

development and germination.

Previous analysis of *lat*59 5' promoter deletion mutants (Twell *et al.*, 1991) demonstrated that deletion from -418 to -115 did not affect promoter activity in transgenic plants, but this deletion was sufficient to decrease promoter activity by 4-fold in the transient assay. In addition, deletion from -115 to -45 decreased promoter activity in transgenic plants and the transient assay to a similar level. Therefore, these data suggest that *lat*59 promoter sequences within the region -115 to -46 are sufficient to support similar levels of promoter activity during pollen maturation and germination. With a subsequent increase in *lat*59 promoter activity in germinating pollen sustained by sequences within the region -418 to -116. Previous analysis of *NTP303* 5' promoter deletion mutants (Weterings *et al.*, 1995) showed that in the transient assay deletion of the promoter from -147 to -103 decreased promoter activity by 30 %, deletion from -103 to -86 decreased promoter activity by a further 60 %. In contrast, analysis of these deletion mutants in transgenic plants demonstrated that deletion from -147 to -103 had no effect, but further deletion to -86 decreased promoter activity by 3-fold. These data indicate that sequences within the *NTP303* promoter region -147 to -103 are preferentially active in germinating pollen.

The above data suggests that the *lat*59 and *NTP303* promoters are developmentally regulated during pollen maturation and germination. Both promoters contain a combination of constitutive pollen-specific transcriptional *cis*-regulatory sequences which are recognised during pollen development and germination. Also present are additional *cis*-regulatory sequences which may be specifically recognised during pollen germination. From these analyses it is clear that detailed developmental analysis of 5' promoter deletion mutants in transgenic plants is required to establish the developmental regulatory patterns of the late pollen genes.

The third phase of developmental control exhibited by *lat*52 occurred at the level of translation. Analysis presented in chapter 5 demonstrated that the 5'-UTRs from pollen-specific genes enhanced reporter gene activity preferentially in germinating pollen independent of native promoter sequences. For these pollen-specific genes it remains to be established at which level this preferential enhancement occurs, but analysis of the viral 5'-UTRs in germinating pollen (chapter 5) would suggest a mechanism of translational enhancement.

In conclusion, late pollen gene expression during pollen development and germination clearly involves complex combinations of transcriptional and translational mechanisms designed to produce controlled amounts of protein at the exact time of requirement.

6.2 Tissue-specific regulation of the *lat*52 gene

Previous analysis of endogenous *lat*52 transcript in tomato showed that the steady state level of *lat*52 mRNA in mature anthers was 20- to 50-fold higher than in petals (Twell *et al.*, 1989b). In addition, *in situ* localisation demonstrated that transcript was detectable in the anther wall (Ursin *et al.*, 1989). The 5' flanking region of the *lat*52 gene (-492 to +110) when fused to the reporter gene *gus* was sufficient to direct GUS activity preferentially in mature pollen of transgenic tobacco (chapter 3; Twell *et al.*, 1990). Data presented in chapter 3 also highlighted a low but detectable level of GUS activity in other sporophytic tissues. These results closely reflected the expression pattern of the above transgene in tomato and *Arabidopsis* (Twell *et al.*, 1990). Previous analysis of a further 2.4 kb of *lat*52 5' flanking sequence (-3000 to +110-*gus-nos3'*) in transgenic tomato did not show enhanced GUS activities above those seen for the -492 to +110 region in petals and anther wall (Twell *et al.*, 1991). These data indicate that regulatory sequences required for enhanced *lat*52 expression in petals and anther wall are absent from the region -3000 to +110. Regulatory sequences controlling tissue-specific gene expression have been localised to the 3'-UTR of the *Arabidopsis GLABROUS1* gene (Larkin *et al.*, 1993), to the intragenic region of *AGAMOUS* (Sieburth and Meyerowitz. 1997) and the transcribed region of *psaDb* (Yamamoto *et al.*, 1997) . Therefore, regulatory sequences required for enhanced *lat*52 expression in petals and anther wall could be contained within either the intragenic region or the 3'-UTR.

Previous analysis of endogenous *lat*52 transcript in tomato showed that the steady state level of *lat*52 mRNA was at least 200-fold higher in pollen than in leaves, where no signal was detected (Twell *et al.*, 1989b). Results presented in chapters 3 and 5 highlighted low but detectable levels of GUS and LUC reporter gene activity respectively in leaves from transgenic plants harbouring the *lat*52 -492 to +110 promoter region. Deriving a ratio of the mean reporter gene activities detected in pollen and leaves indicated that the -492 to +110 promoter region was preferentially expressed in pollen by ~ 40,000-fold (GUS-pollen/GUS-leaves) and ~ 46,000-fold (LUC-pollen/LUC-leaves). In addition, ratios of mean LUC activity between pollen and leaves derived for plants harbouring either the *lat52-52 5'-UTR-luc-C3'* or *lat52-SYN 5'-UTR-luc-C3'* gene fusions (chapter 5) were 46,000 and 4,000 respectively. This difference can be completely accounted for by the 13-fold pollen-specific translational enhancement mediated by the *lat*52 5'-UTR (chapter 5). Therefore, the *lat*52 promoter is transcriptionally activated 4,000-fold more in pollen than in leaves which explains why no endogenous *lat*52 transcript was detectable in tomato leaves by northern blot analyses.

In summary, *cis*-regulatory sequences required for enhanced gene expression in petal and anther wall are absent from the *lat*52 promoter. The *lat*52 promoter is

transcribed at a very low level in sporophytic tissues, but appears pollen-specific as a result of a combination between strong positive pollen-specific transcriptional regulatory sequences and the pollen-specific translational enhancement mediated by the *lat*52 5'-UTR.

Investigation of the spatial and temporal expression patterns of the pollen-specific genes *NTP303* (Weterings *et al.*, 1992), *TUA1* (Ludwig *et al.*, 1988) and *G10* (Rogers *et al.*, 1992) using northern blot analyses showed that endogenous transcript was restricted to the developing pollen grain. Promoter fusion analyses in transgenic plants revealed a low but detectable level of reporter gene activity in tissues where endogenous transcript was shown to be absent for *NTP303* (Weterings, 1994) *TUA1* (Carpenter *et al.*, 1992) and *G10* (Sullivan and Twell, unpublished). These analyses suggest that genes termed pollen-specific may well be expressed at very low levels in other tissue types but transcripts are not detected due to the limited sensitivity of the northern blot analyses used. Since ectopic activation of transgenes in transgenic tobacco by endogenous enhancers has been demonstrated for the *PR-1* promoter (Beilmann *et al.*, 1992). It cannot be ruled out that low level reporter gene activity detected in sporophytic tissues could have arisen from the integration of transgene fusions into or close to sporophytic enhancer/activator elements.

6.3 Shared *cis*-regulatory elements within the late pollen promoters

The detailed mutational analysis performed on the *lat*52 promoter presented in this thesis and by Twell *et al.*, (1991) has led to the identification of a set of distinct *cis*-regulatory elements which act together to coordinate *lat*52 expression during pollen development and germination. Previous analysis showed that the *cis*-regulatory elements PBI (-100 to -86) and PBII (-194 to -176) were essential for maximal expression levels of *lat*52 in pollen (Twell *et al.*, 1991). Gain of function analysis (chapter 4; Twell *et al.*, 1991) clearly demonstrated that the *cis*-regulatory elements PBI and PBII were functionally distinct from each other. Further analysis of the *lat*52 promoter presented in chapter 4 identified two new novel functional *cis*-regulatory elements, GAAA (-71 to -68) and TCCACCATA (-60 to -52) required for *lat*52 expression in pollen. In addition, this analysis also suggested the putative action of at least three other *cis*-regulatory sequences, ACGTAGAAAAATAATACCATT (-332 to -312), AAACATAAAAAATAAAAAATAAAAAAAATAA (-176 to -146) and ACTTCATTATTAATTTTAATTAAT (-124 to -101) involved in the regulation of *lat*52 pollen-specific expression. It is clear that further mutational analysis of these putative *cis*-regulatory sequences is required to establish what role they play in the regulation of *lat*52

pollen-specific expression.

The functional identification of the above and or other *cis*-regulatory sequences present within the late pollen promoters would serve to provide essential information regarding the array of different transcription factors utilised during pollen development and germination. To date only a small number of late pollen promoters have been functionally characterised. Previous analysis by Twell *et al.*, (1991) functionally identified two shared regulatory motifs, the 52/56 box (TGTGGTTATATA) and the 56/59 box (GAATTTGTGA) shared between the *lat*52, *lat*56 and *lat*59 promoters. Targeted mutagenesis throughout the 52/56 box and the 56/59 box demonstrated that the essential nucleotides required for maximum expression levels were restricted to the core sequences GTGG and GTGA respectively (chapter 4; Twell *et al.*, 1991).

5' deletion analysis of the *lat*52 and *lat*59 promoters using the transient assay demonstrated that the 52/56 and 56/59 boxes increased the level of gene expression in pollen but were not essential for pollen-specific expression (chapter 3; Twell *et al.*, 1991). Analysis of the accumulation of GUS activity in transgenic tobacco during anther development demonstrated that the *lat*52 promoter region -100 to +110, and the *lat*59 promoter region -115 to +91 strongly enhanced GUS activity specifically during the mid to late stages of pollen maturation (chapter 3; Patel and Twell, unpublished). Addition of the *lat*52 promoter region -100 to -72, containing the 52/56 box or the *lat*59 promoter region -118 to 101, containing the 56/59 box onto a minimal *CaMV35S* promoter failed to enhance reporter gene activity in pollen (chapter 4; Davies and Twell, unpublished). Taken together, these analyses suggest a common functional role for the 52/56 and 56/59 boxes in the transcriptional activation of late pollen genes. Since the core sequences are similar, GTGG (52/56 box) and GTGA (56/59 box), it is conceivable that a single transcription factor may recognise both target sequences. Transcription factor (s) which bind the aforementioned core sequences have not been isolated, therefore no direct evidence is currently available to substantiate this hypothesis. However, the direct substitution of one box for the other within the *lat*52 and *lat*59 promoters, without decreasing promoter activity could provide indirect evidence that both core sequences are potentially recognised by the same transcription factor.

Gain of function analysis demonstrated that the 52/56 box (PBI) only activated the minimal *CaMV35S* promoter in pollen in the presence of the *cis*-regulatory elements GAAA (-71 to -68) and TCCACCATA (-60 to -52) (chapter 4). This analysis also showed that the *cis*-regulatory elements GAAA (-71 to -68) and TCCACCATA (-60 to -52) together activated minimal *CaMV35S*, but to a lesser extent than when PBI was also present. In addition, the substitution mutation of the *cis*-regulatory element TCCACCATA (-60 to -52) within the *lat*52 promoter region -100 to +110 decreased but did not abolish promoter activity (chapter 4). Taken together, these data suggest that the

cis-regulatory elements PBI+GAAA+TCCACCATA together form the strongest activator complex, followed by GAAA and TCCACCATA together, with the weakest activator complex formed by PBI and GAAA together.

Recently a tobacco genomic clone containing the homolog (*NTP52*) of the tomato *lat*52 gene has been isolated (Bate and Twell, unpublished). The 5' flanking region of the *NTP52* gene (~ 0.5 Kb) when fused to the reporter gene *luc* was sufficient to direct high level LUC activity specifically in germinating pollen using the transient assay (Bate and Twell, unpublished). Comparison of the tomato *lat*52 promoter with the promoter of its tobacco homolog *NTP52* has revealed blocks of sequence identity within the proximal promoter regions (fig. 6.1). The three *cis*-regulatory elements PBI, GAAA and TCCACCATA, which have been shown to be essential for maximum levels of *lat*52 proximal promoter activity (chapter 4) can be identified within the *NTP52* promoter region. The minimal *lat*52 promoter pollen-specific unit -71 to -52 (GAAA+TCCACCATA) shares 80 % homology (16/20) with its respective region within the *NTP52* promoter. Within this region of the *NTP52* promoter the sequence TTTACCATT shows 6/9 bp homology to the *lat*52 *cis*-regulatory element TCCACCATA. Interestingly, the PBII core sequence CCAC is not completely conserved between the two promoters, this may explain the different properties of the two *cis*-regulatory elements as discussed in chapter 4. Therefore, the putative functional sequence within TCCACCATA is ACCAT. The GAAA motif and its position relative to the ACCAT motif is completely conserved within the *NTP52* promoter. Furthermore, the position of the region containing the motifs GAAA and ACCAT relative to the TATA box differs by only 2 bp. Although no extensive sequence identity to the 52/56 box is apparent within the *NTP52* proximal promoter, the sequence GTGG occurs upstream of the region containing the GAAA and ACCAT motifs. The spacing of the GTGG motif relative to the GAAA and ACCAT motifs differs between the two promoters by 18 bp. Since only the central core sequence GTGG of the PBI motif is essential for full activity of the *lat*52 promoter region -100 to +110 (chapter 4) this sequence could well be functional within the *NTP52* promoter.

Previous analysis showed that the *Zm13* promoter region -314 to +61 was sufficient to direct pollen-specific transcription in tobacco (Guerrero *et al.*, 1990). Moreover, analysis of the *Zm13* promoter using the transient expression assay in *Tradescantia* pollen demonstrated that the promoter region -260 to -101 was required for high level activity (Hamilton *et al.*, 1992). Contained within this region are the core motifs GTGG, GAAA, ACCA (fig 6.2 A). Other motifs are also present and these include a 9 bp identical match to the PBII core sequence and a reverse copy of the GTGA motif. Additional sequences essential for pollen-specificity were retained within the region -101 to -54 which contains GTGG in reverse orientation (Hamilton *et al.*, 1992).

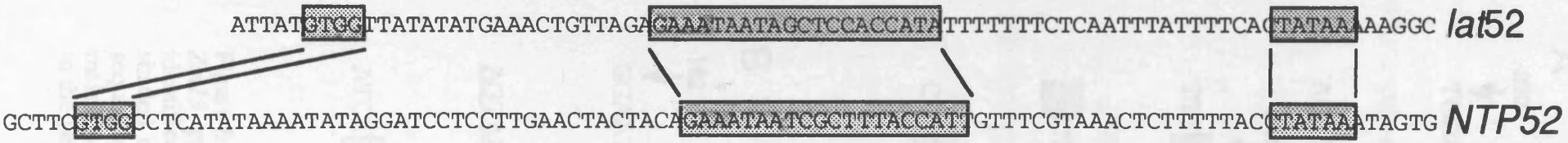


Figure 6.1. Comparison of the *lat*52 and *NTP52* proximal promoter sequences. Conserved pollen specific *cis*-regulatory elements functionally identified within the *lat*52 promoter (chapter 4) present within the *NTP52* promoter are enclosed in shaded boxes.

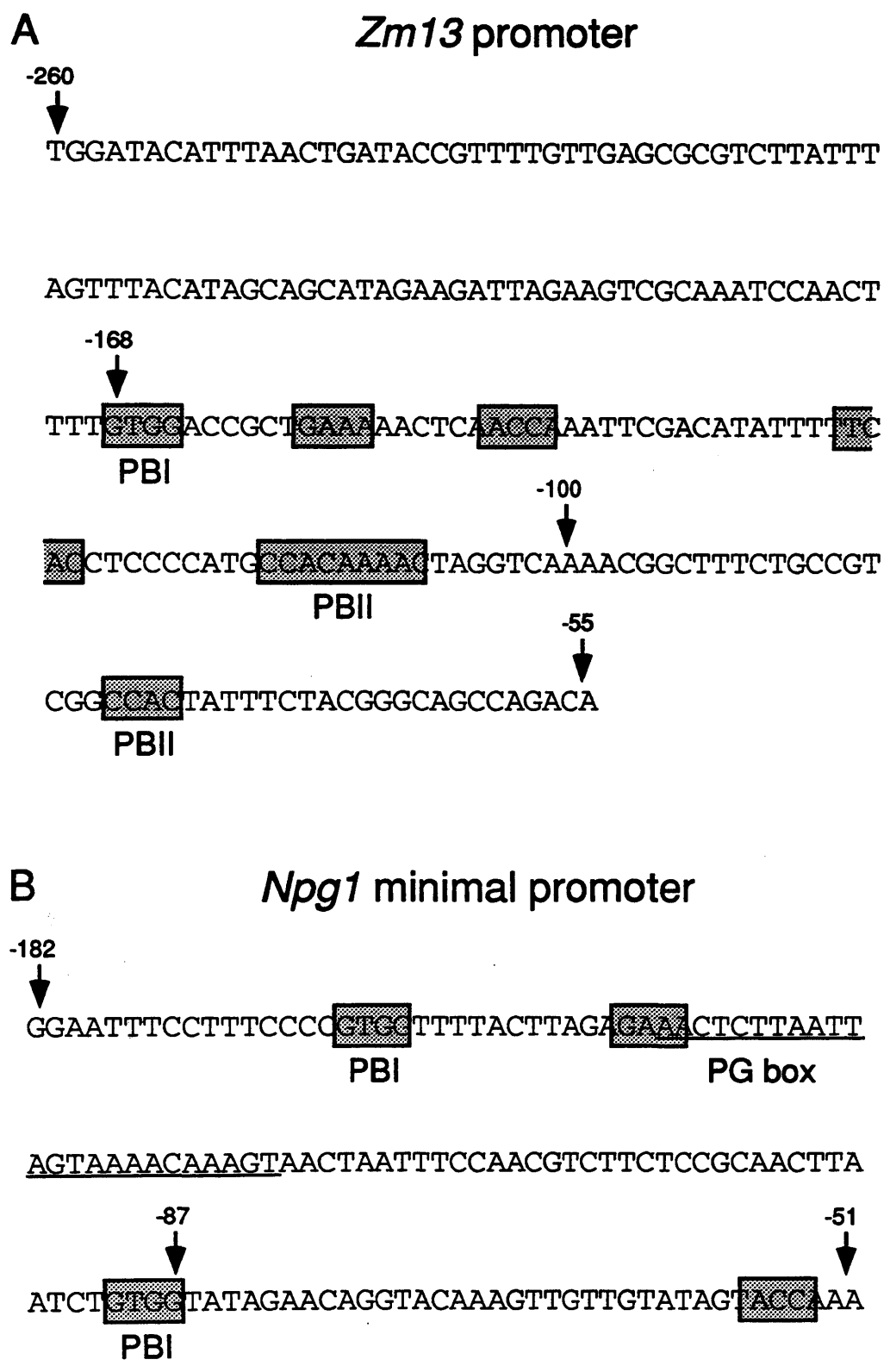


Figure 6.2. Identification of putative pollen-specific *cis*-regulatory motifs within the *Zm13* promoter A) and *Npg1* promoter B). *Cis*-regulatory motifs functionally identified within the *lat*52 promoter (chapter 4; Twell *et al.*, 1991) are enclosed in shaded boxes. Also shown is the PG box (underlined) which was identified from sequence comparisons between the tobacco *Npg1* promoter and the promoter of its maize homolog (Allen and Lonsdale, 1993). Nucleotide positions listed are relative to the beginning of the 5'-UTR of each gene.

Recent analysis of *Npg1* 5' promoter deletion mutants in tobacco pollen using the transient expression assay functionally defined the minimal promoter region required for pollen-specificity as -182 to +85 (Tebbutt and Lonsdale, 1995). Further 5' deletion analysis of the region -182 to +85 demonstrated that deletion from -182 to -86 was sufficient to abolish promoter activity in pollen. These analyses strongly indicated that sequences required for pollen-specific expression of *Npg1* reside within the region -182 to -87. Contained within this region are two copies of the GTGG motif and a GAAA motif (fig 6.2 B). Also present is the PG box which was identified from sequence comparisons between the tobacco *Npg1* promoter and the promoter of its maize homolog (Allen and Lonsdale, 1993). Present within the region -86 to +85 is the ACCA motif (fig 6.2 B).

In summary, the GTGG, GAAA and ACCA motifs, which together form a functional pollen-specific activator unit within the *lat*52 proximal promoter (chapter 4) are also present within the promoters of the pollen-specific genes *NTP52*, *Zm13* and *Npg1*. The order but not the spacing between the three motifs is conserved within all four promoters. Further mutational analysis is required to establish the function of these and other motifs identified within the aforementioned promoters.

Comparison of the tomato *lat*59 promoter with the promoter of its tobacco homolog (*Nt59*) also revealed blocks of sequence identity within the proximal promoter regions (Kulikauskas and McCormick, 1997) (fig 6.3). Previous analysis of the *lat*59 promoter defined two functional regions. The 56/59 box which increased the level of gene expression in pollen but was not essential for pollen-specific expression (Twell *et al.*, 1991) and a minimal 20 bp region within the proximal promoter which is essential for pollen-specificity (Eyal *et al.*, 1995). Within the *Nt*59 promoter the 56/59 box is conserved except for 1 bp (GAACTTGTGA). In addition, the 20 bp sequence identified as essential for pollen-specific expression of *lat*59 is well conserved (16/20 bp) within the *Nt*59 promoter. The distance between the 56/59 box and the 20 bp minimal region required for pollen-specificity is conserved between the two promoters, and both regions are located at similar positions with respect to the TATA box. Previous analysis demonstrated that both GTGA motifs within the region -115 to +91 of the *lat*59 promoter are required for maximal gene expression in pollen (Twell *et al.*, 1991; Patel and Twell, unpublished). The most proximal of the two GTGA motifs present within the *lat*59 promoter is absent from the *Nt*59 promoter, but a GAAA motif is located just upstream of its expected position. Interestingly, a second GTGA motif is present upstream of the 56/59 box within the *Nt59* promoter. Since only the central core sequence GTGA, of the 56/59 box is essential for full activity of the *lat*59 promoter region -115 to +91 (Twell *et al.*, 1991) this sequence could well be functional within the *Nt59* promoter.

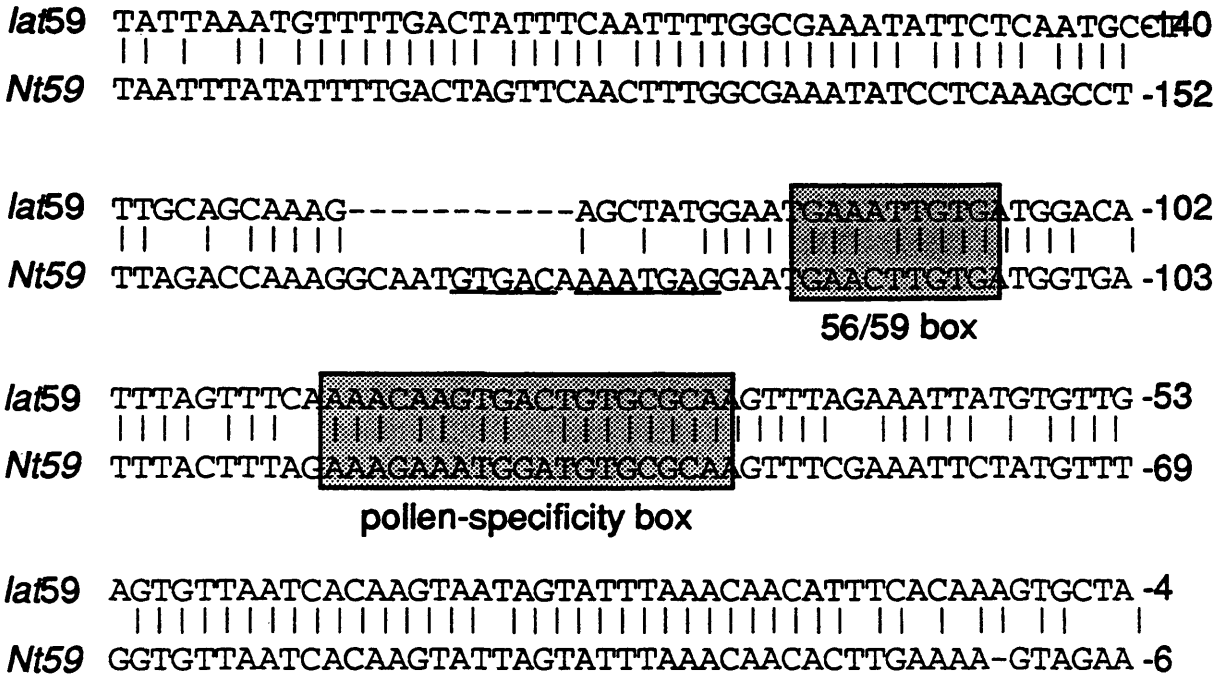


Figure 6.3. Comparison of *lat59* and *Nt59* promoter sequences. Conserved pollen specific *cis*-regulatory elements functionally identified within the *lat59* promoter (Twell *et al.*, 1991; Eyal *et al.*, 1995) present within the *Nt59* promoter are enclosed in shaded boxes. Additional putative *cis*-regulatory elements present within the *Nt59* promoter are underlined. Nucleotide positions listed are relative to the beginning of the 5'-UTR of each gene.

Deletion analysis of the *Arabidopsis TUA1* promoter in transgenic *Arabidopsis* identified two upstream pollen-specific activator regions -1500 to -534 and -271 to -218 (Carpenter *et al.*, 1992). This analysis also showed that the minimal core promoter -97 to +56 was sufficient to direct gene expression specifically in pollen. Further deletion from -97 to -39 abolished promoter activity in pollen which indicated the presence of essential sequences within the region -97 to -40. The regions -271 to -218 and -97 to -40 both contain a single GTGA motif. Interestingly, downstream of the GTGA motif within the region -271 to -218 is the sequence AGTTACTG which shares 7/8 bp homology (underlined) to a sequence present within the *lat*59 promoter defined as essential for the activity of the upstream 56/59 box (Eyal *et al.*, 1995).

Detailed analysis of the *NTP303* promoter identified the novel *cis*-regulatory element AAATGA, of which the TGA triplet was shown to constitute an active part (Weterings *et al.*, 1995). The sequence AAATGA and its position relative to the TATA box is completely conserved within the promoters of *NTP303* and its *Brassica napus* homolog *Bp10* (Albani *et al.*, 1992). In addition, the AAATGA motif is present within the *Nt59* promoter.

In conclusion, the detailed functional analyses of late pollen promoters has led to the identification of a small group of unique *cis*-regulatory elements. These elements can be identified in the promoters of other pollen-specific genes, but whether they are functional remains to be established by site directed mutagenesis.

6.4 Towards the isolation of pollen-specific transcription factors

The identity or properties of the transcription factors which regulate the expression of late pollen genes during pollen development and germination are currently unknown. Genes which encode putative transcription factors/nucleic acid binding proteins have been isolated which are expressed specifically during pollen development (chapter 1). But, the target *cis*-regulatory elements with which these putative factors interact with have not been defined.

Detailed analysis of the *lat*52 promoter (chapter 4; Twell *et al.*, 1991) identified *cis*-regulatory elements which could potentially be recognised by the transcription factors GT-1a (Gilmartin *et al.*, 1992; Perisic and Lam, 1992) and 3AF1 (Lam *et al.*, 1990). In addition, nuclear GT-1- and GT-2-like factors have been shown to interact with the *NTP303* promoter (Hochstenbach *et al.*, 1996). The cDNA clones encoding the transcription factors GT-1a (Gilmartin *et al.*, 1992; Perisic and Lam, 1992), 3AF1 (Lam *et al.*, 1990) and GT-2 (Dehesh *et al.*, 1992) have been isolated. Perhaps the simplest line of investigation would be to first determine if transcripts encoding these transcription

factors are present in mature pollen. Previous analysis using RT-PCR showed that GT-1-related transcript products were detectable in tobacco pollen (Eyal *et al.*, 1995). Sequence analysis of the RT-PCR products identified two different but homologous genes, one of which was proposed to be GT-1a (Gilmartin *et al.*, 1992). If these transcription factors are expressed in pollen then the ultimate goal would be to down-regulate their expression during pollen development using either the antisense or co-suppression approach in transgenic tobacco. Quantification of late pollen gene transcript levels in these plants would indicate if they are regulated by these transcription factors. Recently, the down-regulation of a *myb*-related transcription factor in transgenic *Arabidopsis* was shown to specifically reduce the phytochrome induction of the endogenous *Lhcb1* gene (Wang *et al.*, 1997).

Genes which encode transcription factors such as TGA1a (Katagiri *et al.*, 1989), EPF1 (Takatsuji *et al.*, 1991) and TAF-1 (Oeda *et al.*, 1991) have been isolated by screening lambda-based cDNA expression libraries with multermerised *cis*-regulatory element probes. Prior to expression library screening electrophoretic mobility shift assays were utilised to demonstrate specific protein/DNA interactions. Functional *cis*-regulatory elements have been identified within the promoters of the late pollen genes (chapter 4; Twell *et al.*, 1991; Eyal *et al.*, 1995; Weterings *et al.*, 1995). The application of electrophoretic mobility shift assays to determine pollen-specific protein/DNA interactions could prove to be extremely difficult. The LAT52 homologous, olive tree pollen-allergen, *Ole e* I, has been reported to represent about 1 % of the total dry mass of the mature pollen grain (Villalba *et al.*, 1993). This would imply that nuclear proteins contained within a whole cell protein extract prepared from mature pollen would be present at a very low concentration. This could be overcome by the preparation of nuclear extracts, which would involve the isolation vegetative nuclei (de Paepe *et al.*, 1990) and the utilization of a small scale nuclear protein miniprep procedure (Sablowski *et al.*, 1994). Failure to demonstrate pollen-specific protein/DNA interactions would not prevent the screening of cDNA expression libraries with multermerised *cis*-regulatory element probes. Perhaps the best approach would be to screen with a heterologous population consisting of several known pollen-specific *cis*-regulatory elements.

Alternative methods to clone the genes encoding transcription factors that interact with pollen-specific *cis*-regulatory elements are based on genetic selection systems. One such method based on genetic selection in yeast has been used to clone cDNAs encoding the olfactory neuronal transcription factor Olf-1 (Wang and Reed, 1993) and bZip transcription factors which interact with the *Dc3* promoter (Kim *et al.*, 1997). Briefly, a cDNA library prepared from mature pollen is fused to the DNA sequence encoding the Gal4 activator domain and transfected into yeast cells containing pollen-specific *cis*-regulatory elements fused upstream of the minimal Gal4 promoter. When a protein

encoded by a pollen cDNA binds to the pollen-specific *cis*-regulatory element, it brings the Gal4 activator domain in contact with the Gal4 minimal promoter which activates the transcription of a selectable marker gene. The second system would involve the mutagenesis of transgenic *Arabidopsis*, homozygous for hybrid *lat52/CaMV35S* promoter fusions (chapter 4) fused to the reporter gene green fluorescent protein (GFP). Mutation of a gene (s) encoding a transcription factor (s) which regulates *lat52* expression would be identified from the loss of GFP activity in 50 % of mature pollen grains from an individual M1 plant.

In conclusion, multiple pollen-specific *cis*-regulatory elements within the late pollen promoters which are likely to be recognised by transcription factors have been defined (this thesis; chapter 1). These elements will provide excellent target sequences to identify and clone transcription factors expressed during pollen development. Furthermore, these elements will provide good candidate sequences with which to identify the binding sites of putative transcription factors already isolated from pollen.

Bibliography

Albani, D., Robert, L. S., Donaldson, P. A., Altosaar, I., Arnison, P. G. and Fabijanski, S. F. (1990). Characterization of a pollen-specific gene family from *Brassica napus* which is activated during early microspore development. Plant Mol. Biol. 15, 605-622.

Albani, D., Altosaar, I., Arnison, P. G. and Fabijanski, S. F. (1991). A gene showing sequence similarity to pectin esterase is specifically expressed in developing pollen of *Brassica napus*. Sequences in its 5' flanking region are conserved in other pollen-specific promoters. Plant Mol. Biol. 16, 501-513.

Albani, D., Sardana, R., Robert, L. S., Altosaar, I., Arnison, P. G. and Fabijanski, S. F. (1992). A *Brassica napus* gene family which shows sequence similarity to ascorbate oxidase is expressed in developing pollen - molecular characterization and analysis of promoter activity in transgenic plants. Plant J. 2, 331-342.

Allen, R. L. and Lonsdale, D. M. (1993). Molecular characterization of one of the maize polygalacturonase gene family members which are expressed during late pollen development. Plant J. 3, 261-271.

Allison, R., Johnston, R. E. and Dougherty, W. G. (1986). The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. Virology 154, 9-20.

Altman, M., Blum, S., Wilson, T. M. A. and Trachsel, H. (1990). The 5'-leader sequence of tobacco mosaic virus RNA mediates initiation-factor-4E-independent, but still initiation-factor-4A-dependent translation in yeast extracts. Gene 91, 127-129.

Baerson, S. R. and Lamppa, G. K. (1993). Developmental regulation of an acyl carrier protein gene promoter in vegetative and reproductive tissues. Plant Mol. Biol. 22, 255-267.

Baerson, S. R., Vander Heiden, M. G. and Lamppa, G. K. (1995). Identification of domains in an *Arabidopsis* acyl carrier protein gene promoter required for maximal organ-specific expression. Plant Mol. Biol. 26, 1947-1959.

Bailey-Serres, J. and Dawe, R. K. (1996). Both 5' and 3' sequences of maize *adh1* mRNA are required for enhanced translation under low-oxygen conditions. Plant Physiol. 112, 685-695.

Baltz, R., Domon, C. and Steinmetz, A. (1992). Characterization of a pollen-specific cDNA from sunflower encoding a zinc finger protein. Plant J. 2, 713-721.

Basso, J., Dallaire, P., Charest, P. J., Devantier, Y. and Laliberte, J-F. (1994). Evidence for an internal ribosome entry site within the 5' non-translated region of turnip mosaic potyvirus RNA. J. Gen. Virol. 75, 3157-3165.

Batanero, E., Villalba, M., Lopez-Otin, C. and Rodriguez, R. (1994). Isolation and characterisation of an olive allergen-like protein from lilac pollen. Sequence analysis of three cDNA encoding protein isoforms. Eur. J. Biochem. 221 , 187-193.

Bedinger, P. (1992). The remarkable biology of pollen. Plant Cell 4, 879-887.

Belostotsky, D. A. and Meagher, R. B. (1993). Differential organ-specific expression of three poly (A) binding protein genes from *Arabidopsis*. Proc. Natl. Acad. Sci. USA 90, 6686-6690.

Belostotsky, D. A. and Meagher, R. B. (1996). A pollen-, ovule-, and early embryo-specific poly (A) binding protein from Arabidopsis complements essential functions in yeast. Plant Cell 8, 1261-1275.

Belsham, G. J. and Sonenberg, N. (1996). RNA-protein interactions in regulation of picornavirus RNA translation. Microbiological Reviews 60, 499-511.

Benfey, P. N., Ren, L. And Chua, N-H. (1989). The CaMV S-35 enhancer contains at least 2 domains which can confer different developmental and tissue-specific expression patterns. EMBO J. 8, 2195-2202.

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

Bhattacharyya, M. K., Paiva, N. L., Dixon, R., Korth, K. L. and Stermer, B. A. (1995). Features of the *hmg* 1 subfamily of genes encoding HMG-CoA reductase in potato. Plant Mol. Biol. 28, 1-15.

Blomstedt, C. K., Knox, R. B. and Singh, M. B. (1996). Generative cells of *lilium longiflorum* possess translatable messenger-RNA and functional protein-synthesis machinery. Plant Mol. Biol. 31, 1083-1086.

Bolle, C., Sopory, S., Lubberstedt, T., Herrmann, R. G. and Oelmüller, R. (1994). Segments encoding 5'-untranslated leaders of genes for thylakoid proteins contain *cis*-elements essential for transcription. Plant J. 6, 513-523.

Brander, K. A. and Kuhlemeier, C. (1995). A pollen-specific DEAD-box protein related to translation initiation factor eIF-4A from tobacco. Plant Mol. Biol. 27, 637-649.

Brown, S. M. and Crouch, M. L. (1990). Characterization of a gene family abundantly expressed in *Oenothera organenis* pollen that shows sequence similarity to polygalacturonase. Plant Cell 2, 263-274.

Buchel, A. S., Molenkamp, R., Bol, J. F. and Linthorst, H. J. M. (1996). The PR-1a promoter contains a number of elements that bind GT-1 like nuclear factors with different affinity. Plant Mol. Biol. 30, 493-504.

Bucher, M., Brander, K. A., Sbicego, S., Mandel, T. and Kuhlemeier, C. (1995). Aerobic fermentation in tobacco pollen. Plant Mol. Biol. 28, 739-750.

Bullock, W. O., Fernandez, J. M. and Short, J. M. (1987). XL1-blue- a high-effiviency plasmid transforming *RecA Escherichia coli* strain with ß-galactosidase selection. Biotechniques 5,376.

Buratowski, S. (1994). The basics of basal transcription by RNA polymerase II. Cell 77, 1-3.

Burrow, M. D., Sen, P., Chlan, C. A. and Murai, N. (1992). Developmental control of the ß-phaseolin gene requires positive, negative, and temporal seed-specific transcriptional regulatory elements and a negative element for stem and root expression. Plant J. 2, 537-548.

Carpenter, J. L., Ploense, S. E., Snustad, D. P. and Silflow, C. D. (1992). Preferential expression of an α-tubulin gene of Arabidopsis in pollen. Plant Cell 4, 557-571.

Carrington, J. C. and Freed, D. D. (1990). Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region. J. Virology 64, 1590-1597.

Casper, T. and Quail, P. H. (1993). Promoter and leader regions involved in the expression of the *Arabidopsis* ferredoxin A gene. Plant J. 3, 161-174.

Chamot, D. and Kuhlemeier, C. (1992). Differential expression of genes encoding the hypusine-containing translation initiation factor, eIF-5A, in tobacco. Nucleic Acids Res. 20, 665-669.

Chappell, J. and Chrispeels, M. J. (1986). Transcriptional and posttranscriptional control of phaseolin and phytohemagglutinin gene expression in developing cotyledons of *Phaseolus vulgaris*. Plant Physiol. 81, 50-54.

Chung, Y-Y., Kim, S-R., Kang, H-G., Noh, Y-S., Park, M. C., Finkel, D. and An, G. (1995). Characterization of two rice MADS box genes homologous to *GLOBOSA*. Plant Sci. 109, 45-56.

Coupe, S. A. and Deikman. J. (1997). Characterization of a DNA-binding protein that interacts with 5' flanking regions of two fruit-ripening genes. Plant J. 11, 1207-1218.

Dawe, R. K., Lachmansingh, A. R. and Freeling, M. (1993). Transposon-mediated mutations in the untranslated leader of maize *Adh1* that increase and decrease pollen-specific gene expression. Plant Cell 5, 311-319.

Deboo, G. B., Albertsen, M. C. and Taylor, L. P. (1995). Flavanone 3-hydroxylase transcripts and flavanol accumulation are temporally coordinate in maize anthers. Plant J. 7, 703-713.

Dehesh, K., Bruce, W. B. and Quail, P. H. (1990). A trans-acting factor that binds to a GT-motif in a phytochrome gene promoter. Science 250, 1397-1399

Dehesh, K., Hung, H., Tepperman, J. M. and Quail, P. H. (1992). GT-2: a transcription factor with twin autonomous DNA-binding domains of closely related but different target sequence specificity. EMBO J. 11, 4131-4144.

de Paepe, R., Koulou, A., Pham, J. L. Brown, S. C. (1990). Nuclear DNA content and separation of *Nicotiana sylvestris* vegetative and generatative nuclei at various stages of male gametogenesis. Plant Sci. 70, 255-265.

Eady, C. Lindsey, K. and Twell, D. (1994). Differential activation and conserved vegetative cell-specific activity of a late pollen promoter in species with bi- and tricellular pollen. Plant J. 5, 543-550.

Eady, C. Lindsey, K. and Twell, D. (1995). The significance of microspore division and division symmetry for vegetative cell-specific transcription and generative cell differentiation. Plant Cell 7 , 65-74.

Eloranta, J. J. and Goodbourn, S. (1996). Positive and negative regulation of RNA polymerase II transcription. In Eukaryotic gene transcription (ed, Goodbourn, S.). New York: Oxford University Press inc., pp.1-33.

Enjuto, M., Lumbreras, V., Marin, C. and Boronat, A. (1995). Expression of the *Arabidopsis HMG2* gene, encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase, is restricted to meristematic and floral tissues. Plant Cell 7, 517-527.

Eyal, Y., Curie, C. and McCormick, S. (1995). Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen-expressed genes. Plant Cell 7, 373-384.

Fang, R-X., Nagy, F., Sivasubramaniam, S. and Chua, N-H. (1989). Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic 35S promoter in transgenic plants. Plant Cell 1, 141-150.

Fennoy, S. L. and Bailey-Serres, J. (1995). Post-transcriptional regulation of gene expression in oxygen-deprived roots of maize. Plant J. 7, 287-295.

Foster, G. D., Robinson, S. W., Blundell, R. P., Roberts, M. R., Hodge, R., Draper, J. and Scott, R. J. (1992). A *Brassica napus* mRNA encoding a protein homologous to phospholipid transfer proteins, is expressed specifically in the tapetum and developing microspores. Plant Sci. 84, 187-192.

Fu, H., Kim, S. Y. and Park, W. D. (1995), A potato *Sus3* sucrose synthase gene contains a context-dependent 3' element and a leader intron with both positive and negative tissue-specific effects. Plant Cell 7, 1395-1403.

Gaillard, A., Vergne, P. and Beckert, M. (1991). Optimization of maize microspore isolation and culture conditions for reliable plant regeneration. Plant Cell Rep. 10, 55-58.

Gallie, D. R., Sleat, D. E., Watts, J. W., Turner, P. C. and Wilson, T. M. (1987). The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. Nucleic Acids Res. 15, 8693-8711.

Gallie, D. R., Feder, J. N., Schimke, R. T. and Walbot, V. (1991). Post-transcriptional regulation in higher eukaryotes: The role of the reporter gene in controlling expression. Mol. Gen. Genet. 228, 258-264.

Gallie, D. R. and Walbot, V. (1992). Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation. Nucleic Acids Res. 20, 4631-4638.

Gallie, D. R. (1993). Posttranscriptional regulation of gene expression in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 44, 77-105.

Gallie, D. R. and Young, T. E. (1994). The regulation of gene expression in transformed maize and aleurone and endosperm protoplasts. Plant Physiol. 106, 929-939.

Gallusci, P., Salamini, F. and Thompson, R. D. (1994). Differences in cell type-specific expression of the gene *Opaque-2* in maize and transgenic tobacco. Mol. Gen. Genet. 244, 391-400.

Gasch, A., Hoffmann, A., Horikoshi, M., Roeder, R. G. and Chua, N-H. (1990). *Arabidopsis thaliana* contains two genes for TFIID. Nature 346, 390-394.

Gilmartin, P. M., Memelink, J., Hiratsuka, K., Kay, S. A. and Chua, N.-H. (1992). Characterisation of a gene encoding a DNA binding protein with specificity for a light-responsive element. Plant Cell 4, 839-849.

Green, P. J., Yong, M. H., Cuozzo, M., Kano-Murakami, Y., Silverstein, P. and Chua, N.-H. (1988). Binding site requirements for pea nuclear protein factor GT-1 correlate with sequences required for light-dependent transcriptional activation of the *rbcS-3A* gene. EMBO J. 7, 4035-4044.

Greenblatt, J. (1997). RNA polymerase II holoenzyme and transcriptional regulation. Curr. Opin. Cell Biol. 9, 310-319.

Gu, W. and Hecht, N. B. (1996). Translation of a testis-specific Cu/Zn superoxide dismutase (SOD-1) mRNA is regulated by a 65-Kilodalton protein which binds to its 5' untranslated region. Mol. Cell. Biol. 16, 4535-4543.

Guerrero, F. D., Crossland, L., Smutzer, G. S., Hamilton, D. A. and Mascarenhas, J. P. (1990). Promoter sequences from a maize pollen-specific gene direct transcription in tobacco. Mol. Gen. Genet. 224, 161-168.

Haass, M. M. and Feix, G. (1992). Two different cDNAs encoding TFIID proteins of maize. Fed. Eur. Biochem. Soc. 3, 294-298.

Haass, M., Griess, E., Goddemeier, M., Egly, J. M. and Feix, G. (1994). The TATA box-binding protein-1 (*Tbp1*) of maize displays promoter-specific DNA-binding affinities. Plant Sci. 100, 187-194.

Hamilton, D. A., Roy, M., Rueda, J., Sindhu, R. K., Sanford, J. and Mascarenhas, J. P. (1992). Dissection of a pollen-specific promoter from maize by transient transformation assays. Plant Mol. Biol. 18, 211-218.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Biol. Chem. 166, 557-580.

Hanson, D. D., Hamilton, D. A., Travis, J. L., Bashe, D. M. and Mascharenhas, J. P. (1989). Characterization of a pollen-specific cDNA clone from *Zea mays* and its expression. Plant Cell 1, 173-179.

Hartings, H., Maddaloni, M., Lazzaroni, N., Di Fonzo, N., Motto, M., Salamini, F. and Thompson, R. (1989). The *O2* gene which regulates zein deposition in maize endosperm encodes a protein with structural homologies to transcriptional activators. EMBO J. 8, 2795-2801.

Heslop-Harrison, J. (1987). Pollen germination and pollen tube growth. International reveiw of cytology (Giles, K. L. and Prakash, J. eds). Academic press inc. Orlando 107, 1-78.

Hochstenbach, R., de Groot, P., Jacobs, J., Schrauwen, J. and Wullens, G. (1996). The promoter of a gene that is expressed only in pollen interacts with ubiquitous transcription factors. Sex. Plant Reprod. 9, 197-202.

Hofgen, R. and Willmitzer, L. (1988). storage of competent cells for *Agrobacterium* transformation. Nucleic Acids Res. 16, 9877.

Holdsworth, M. J., Grierson, C., Schuch, W. and Bevan, M. (1992). DNA-binding properties of cloned TATA-binding protein from potato tubers. Plant Mol. Biol. 19, 455-464.

Horner, Jr. H. T. (1977). A comparative light- and electron-microscopic study of microsporogenesis in male-fertile and cytoplasmic male-sterile sunflower (*Helianthus annus*). Am. J. Bot. 64, 745-759.

Iwataki, N., Hoya, A. and Yamazaki, K. (1997). Restoration of TATA-dependent transcription in a heat inactivated extract of tobacco nuclei by recombinant TATA-binding protein (TBP) from tobacco. Plant Mol. Biol. 34, 69-79.

Jacobsen, K., Laursen, N. B., Jensen, E. O., Marcker, A., Poulsen, C. and Marcker, K. A. (1990). HMG I-like proteins from leaf and nodule nuclei interact with different AT motifs in soybean nodulin promoters. Plant Cell 2, 85-94.

Jackson, R. J., Hunt, S. L., Reynolds, J. E. and Kaminski, A. (1996). Cap-dependent and cap-independent translation: operational distinctions and mechanistic interpretations. In Current topics in microbiology and immunology (ed Sarnow, P.). Springer-Verlag, Berlin Heidelberg, pp 1-29.

Jefferson, R. A., Kavanagh, T. A and Bevan, M. W. (1987). GUS fusions: ß-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. 6, 3901-3907.

Jobling, S. E. and Gehrke, L. (1987). Enhanced translation of chimaeric messenger RNAs containing a plant viral untranslated leader sequence. Nature 325, 622-625.

Joshi, C. P. (1987). An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucleic Acids Res. 15, 6643-6653.

Kamalay, J. C. and Goldberg, R. B. (1980). Regulation of structural gene expression in tobacco. Cell 19, 935-946.

Kamalay, J. C. and Goldberg, R. B. (1984). Organ-specific nuclear RNAs tobacco. Proc. Natl. Acad. Sci. USA 81, 2801-2805.

Katagiri, F., Lam, E. and Chua, N-H. (1989). Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. Nature 340, 727-730.

Kawata, T., Minami, M., Tamura, T-A., Sumita, K. and Iwabuchi, M. (1992). Isolation and characterization of a cDNA clone encoding the TATA box-binding protein (TFIID) from wheat. Plant Mol. Biol. 19, 867-872.

Kim, S. Y., Chung, H-J. and Thomas, T. L. (1997). Isolation of a novel class of bZip transcription factors that interact with ABA-responsive and embryo-specification elements in the *Dc3* promoter using a modified yeast one-hybrid system. Plant J. 11, 1237-1251.

Klausner, R. D., Rouault, T. A. and Harford, J. B. (1993). Regulating the fate of mRNA: The control of cellular iron metabolism. Cell 72, 19-28.

Kloeckener-Gruissem, B., Vogel, J. M. and Freeling, M. (1992). The TATA box promoter region of maize *Adh1* affects its organ-specific expression. EMBO J. 11, 157-166.

Kodrzycki, R., Boston, R. S. and Larkins, B. A. (1989). The *opaque-2* mutation of maize differentially reduces zein gene transcription. Plant Cell 1, 105-114.

Koromilas, A. E., Lazaris-Karatzas, A. and Sonenberg, N. (1992). mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E.EMBO J. 11, 4153-4158.

Kozak, M. (1986). Point muations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44, 283-292.

Kozak, M. (1988). Leader length and secondary structure modulate mRNA function under conditions of stress. Mol. Cell. Biol. 8, 2737-2744.

Kozak, M. (1989). Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. Mol. Cell. Biol. 9, 5134-5142.

Kuhlemeier, C., Cuozzo, M., Green, P. J., Goyvaerts, E., Ward, K. and Chua, N.-H. (1988). Localisation and conditional redundancy of regulatory elements in *rbcS-3A*, a pea gene encoding the small subunit of ribulose-bisphosphate carboxylase. Proc. Natl. Acad. Sci. USA 85, 4662-4666.

Kuhlemeier, C. (1992). Transcriptional and post-transcriptional regulation of gene expression in plants. Plant Mol. Biol. 19, 1-14.

Kulikauskas, R. and McCormick, S. (1997). Identification of the tobacco and *Arabidopsis* homologues of the pollen-expressed LAT59 gene of tomato. Plant Mol. Biol. 34, 809-814.

Lam, E., Benfey, P. N., Gilmartin, P. M., Fang, R-X. and Chua, N-H. (1989). Site-specific mutations alter *in vitro* factor binding and change promoter expression pattern in transgenic plants. Proc. Natl. Acad. Sci. USA 86, 7890-7894.

Lam, E. and Chua, N-H. (1989). ASF-2: a factor that binds to cauliflower mosaic virus 35S promoter and a conserved GATA motif in *Cab* promoters. Plant Cell 1, 1147-1156.

Lam, E. and Chua, N-H. (1990). GT-1 binding site confers light responsive expression in transgenic tobacco. Science 248, 471-474.

Lam, E., Kano-Murakami, Y., Gilmartin, P., Niner, B. and Chua, N-H. (1990). A metal-dependent DNA-binding protein interacts with a constitutive element of a light-responsive promoter. Plant Cell 2, 857-866.

Larkin, J. C., Oppenheimer, D. G., Pollock, S. and Marks, M. D. (1993). *Arabidopsis GLABROUS1* gene requires downstream sequences for function. Plant Cell 5, 1739-1748.

Lohmer, S., Maddaloni, M., Motto, M., Di Fonzo, N., Hartings, H., Salamini, F. and Thompson, R. D. (1991). The maize regulatory locus *Opaque-2* encodes a DNA-binding protein which activates the transcription of the b-32 gene. EMBO J. 10, 617-624.

Lombardero, M., Barbas, J. A., Moscoso del Prado, J. and Carreira, J. (1994). cDNA sequence analysis of the main olive allergen, *Ole e I*. Clin. Exp. Allergy 24, 765-770.

Lopez, I., Anthony, R. G., Maciver, S. K., Jiang, C-J., Khan, S., Weeds, A. G. and Hussey, P. J. (1996). Pollen specific expression of maize genes encoding actin depolymerizing factor-like proteins. Proc. Natl. Acad. Sci. USA 93, 7415-7420.

Lubberstedt, Th., Oelmuller, R., Wanner, G. and Herrmann, R. G. (1994). Interacting *cis* elements in the plastocyanin promoter from spinach ensure regulated high-level expression. Mol. Gen. Genet. 242, 602-613.

Ludwig, S. R., Oppenheimer, D. G., Silflow, C. D. and Snustad, D. P. (1988). The α1-tubulin gene of *Arabidopsis thaliana*: Primary structure and prefferential expression in flowers. Plant Mol. Biol. 10, 311-321.

Lutcke, H. A., chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F., Scheele, G. A. (1987). Selection of AUG initiation codons differs in plants and animals. EMBO J. 6, 43-48.

Marcotte, W. R., Russell, S. H. and Quatrano, R. S. (1989). Abscisic acid-responsive sequences from the *Em* gene of wheat. Plant Cell 1, 969-976.

Marsolier, M-C., Carrayol, E. and Hirel, B. (1993). Multiple functions of promoter sequences involved in organ-specific expression and ammonia regulation of a cytosolic soybean glutamine synthetase gene in transgenic *Lotus corniculatus*. Plant J. 3, 405-414.

Mascarenhas, J. P. (1989). The male gametophyte of flowering plants. Plant cell 1, 657-664.

Mascarenhas, J. P. (1990). Gene activity during pollen development. Ann. Rev. Plant Phys. Mol. Biol. 41, 317-338.

McCormick, S., Smith, A., Gasser, C., Sachs, K., Hinchee, M., Horsch, R., and Fraley, R. (1987). The identification of genes specifically expressed in reproductive organs of tomato. In Tomato biotechnology (ed. D. Nevins and R. Jones), pp. 255-265. A.R. Liss, New York.

McCormick, S. (1993). Male gametophyte development. Plant Cell 5, 1265-1275.

McCormick, S., Twell, D., Vancanneyt, G. and Yamaguchi, J. (1991). Molecular analysis of gene regulation and function during male gametophyte development. In: G. Jenkins & W. Schuch (eds.) Molecular Biology of Plant Development. Proceedings of SEB symposium No. 45. Company of Biologists, Cambridge.

Michel, D., Salamini, F., Bartels, D., Dale, P., Baga, M. and Szalay, A. (1993). Analysis of a dessication and ABA-responsive promoter isolated from the ressurection plant *Caterostigma plantagineum*. Plant J. 4, 29-40.

Michelet, B., Lukaszewicz, M., Dupriez, V. and Boutry, M. (1994). A plant plasma membrane proton-ATPase gene is regulated by development and enviroment and shows signs of a translational regulation. Plant Cell 6, 1375-1389.

Mo, Y., Nagel, C. and Taylor, L. P. (1992). Biochemical complementation of chalcone synthase mutants defines a role for flavanols in functional pollen. Proc. Natl. Acad. Sci. USA 89, 7213-7217.

Moore, I., Diefenthal, T., Zarsky, V., Schell, J. and Palme, K. (1997). A homolog of the mammalian GTPase *Rab2* is present in *Arabidopsis* and is expressed predominantly in pollen grains and seedlings. Proc. Natl. Acad. Sci. USA 94, 762-767.

Motto, M., Maddaloni, M., Ponziani, G., Brembilla, M., Maratta, R., Di Fonzo, N., Soave, C., Thompson, R. and Salamini, F. (1988). Molecular cloning of the *o2-m5* allele of *Zea mays* using transposon marking. Mol. Gen. Genet. 212, 488-494.

Mu, J-H., Stains, J. and Kao, T-H. (1994). Characterization of a pollen-expressed gene encoding a putative pectin esterase of *Petunia inflata*. Plant Mol. Biol. 25, 539-544.

Murgia, M., Charzynska, M., Rougier, M. and Cresti, M. (1991). Secretory tapetum of *Brassica oleracea* L.: polarity and ultrastructural features. Sex. Plant Reprod. 4, 28-35.

Muschietti, J., Dircks, L., Vancanneyt, G. and McCormick, S. (1994). LAT52 protein is essential for tomato pollen development: pollen expressing antisense LAT52 RNA hydrates and germinates abnormally and cannot achieve fertilization. Plant J. 6, 321-338.

Muth, J. R., Muller, M., Lohmer, S., Salamini, F. and Thompson, R. D. (1996). The role of multiple binding sites in the activation of zein gene expression by *opaque-2*. Mol. Gen. Genet. 252, 723-732.

Neuhaus, G., Neuhaus-Uri, G., Katagiri, F., Seipel, K. and Chua, N-H. (1994). Tissue-specific expression of *as-1* in transgenic tobacco. Plant Cell 6, 827-834.

Ni, M., Dehesh, K., Tepperman, J. M. and Quail, P. H. (1996). GT-2: in vivo transcriptional activation activity and definition of novel twin DNA binding domains with reciprocal target sequence selectivity. Plant Cell 8, 1041-1059.

Niogret, M. F., Dubald, M., Mandaron, P. and Mache, R. (1991). Characterization of pollen polygalacturonase encoded by several cDNA clones in maize. Plant Mol. Biol. 17, 1155-1164.

Nishihara, M., Ito, M., Tanaka, I., Kyo, M., Ono, K., Irifune, K. and Morikawa, H. (1993). Expression of the ß-glucuronidase gene in pollen of lily (*Lilium longiflorum*), Tobacco (*Nicotiana tabacum*), *Nicotiana rustica*, and peony (*Paeonia lactiflora*) by particle bombardment. Plant Physiol. 102, 357-361.

Oeda, K., Salinas, J. and Chua, N-H. (1991). A tobacco bZip transcriptional activator (TAF-1) binds to a G-box-like motif conserved in plant genes. EMBO J. 10, 1793-1802.

Oh, S-K., Scott, M. P. and Sarnow, P. (1992). Homeotic gene *Antennapedia* mRNA contains 5'-noncoding sequences that confer translational initiation by internal ribosome binding. Genes Dev. 6, 1643-1653.

Oldenhof, M. T., de Groot, P. F. M., Visser, J. H., Schrauwen, J. A. M. and Wullems, G. J. (1996). Isolation and characterization of a microspore-specific gene from tobacco. Plant Mol. Biol. 31, 213-225.

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

Ow, D. W., Wood, K. V., DeLuca, M., DeWet, J. R., Helinski, D. R. and Howell, S. H. (1986). Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. Science 234, 856-859.

Ow, D. W., Jacobs, J, D. and Howell, S. H. (1987). Functional regions of the cauliflower mosaic virus 35S RNA promoter determined by use of the firefly luciferase gene as a reporter of promoter activity. Proc. Natl. Acad. Sci. USA 84, 4870-4874.

Owen, H. A. and Makaroff, C. A. (1995). Ultrastructure of microsporogenesis and microgametogenesis in *Arabidopsis thaliana* (L.) Heynh. ecotype Wassilewskija (Brassicaceae). Protoplasma 185, 7-21.

Owttrim, G. W., Hofman, S. and Kuhlemeier, C. (1991). Divergent genes for translation initiation factor eIF-4A are coordinately expressed in tobacco. Nucleic Acids Res. 19, 5491-5496.

Pacini, E. (1990). Tapetum and microspore function. In Microspores: Evolution and Ontogeny, ed. Blackmore, S. and Knox, R. B. pp. 213-237. London: Academic Press.

Pacini, E., Franchi, G. G., Lisci, M. and Nepi, M. (1997). Pollen viability related to type of pollination in six angiosperm species. Annals of Botany 80, 83-87.

Paul, W., Hodge, R., Smartt, S., Draper, J. and Scott, R. (1992). The isolation and characterization of the tapetum-specific *Arabidopsis thaliana* A9 gene. Plant Mol. Biol. 19, 611-622.

Pe, M. E., Frova, C., Colombo, L., Binelli, G., Mastroianni, N., Tarchini, R. and Visioli, G. (1994). Molecular cloning of genes expressed in pollen of *Sorghum bicolor*. Maydica 39, 107-113.

Perisic, O. and Lam, E. (1992). A tobacco DNA binding protein that interacts with a light-responsive box II element. Plant Cell 4, 831-838.

Prat, S., Willmitzer, L. and Sanchez-Serrano, J. J. (1989). Nuclear proteins binding to a cauliflower mosaic virus 35S truncated promoter. Mol. Gen. Genet. 217, 209-214.

Puente, P., Wei, N. and Deng, X. W. (1996). Combinatorial interplay of promoter elements constitutes the minimal determinants for light and developmental control of gene expression in *Arabidopsis*. EMBO J. 15, 3732-3743.

Rafnar, T., Rogers, B. L., Bond, J. E., Kuo, M. and Klapper, D. G. (1991). Isolation of cDNA clones coding for the major short ragweed allergen *Amb a* (antigen E.). J. Biol. Chem. 266, 1229-1236.

Roberts, M. R., Robson, F., Foster, G. D., Draper, J. and Scott, R. J. (1991). A *Brassica napus* mRNA expressed specifically in developing microspores. Plant Mol. Biol. 17, 295-299.

Roberts, M. R., Foster, G. D., Blundell, R. P., Robinson, S. W., Kumar, A., Draper, J. and Scott, R. J. (1993a).Gametophytic and sporophytic expression of an anther-specific *Arabidopsis thaliana* gene. Plant J. 3, 111-120.

Roberts, M. R., Hodge, R., Ross, J. H. E., Sorensen, A., Murphy, D. J., Draper, J. and Scott, R. J. (1993b). Characterization of a new class of oleosins suggests a male gametophyte-specific lipid storage pathway. Plant J. 3, 629-636.

Rodriguez-Garcia, M. I., Fernandez, M. C. and Alche, J. D. (1995). Immunocytochemical localisation of allergenic protein (*Ole e I*) in the endoplasmic reticulum of the developing pollen grain of olive (*Olea europaea L.*). Planta 196, 558-563.

Rogers, H. J., Harvey, A. and Lonsdale, D. M. (1992). Isolation and characterisation of a tobacco gene with homology to pectate lyase which is specifically expressed during microsporogenesis. Plant Mol. Biol. 20, 493-502.

Sablowski, R. W. M., Moyano, E., Culianez-Macia, A., Schuch, W., Martin, C. Bevan, M. (1994). A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. EMBO J. 13, 128-137.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular cloning - a laboratory manual. Cold Spring Harbour Press., second edition.

Sawadogo, M. And Van Dyke, M. W. (1991). A rapid method for the purification of deprotected oligodeoxynucleotides. Nucleic Acids Res. 19, 674.

Schmidt, R. J., Burr, F. A. and Burr, B. (1987). Transposon tagging and molecular analysis of the maize regulatory locus *opaque-2*. Science 238, 960-963.

Schmidt, R. J., Ketudat, M., Aukerman, M. J. and Hoschek, G. (1992). *Opaque-2* is a transcriptional activator that recognises a specific target site in 22-KD zein genes. Plant Cell 4, 689-700.

Schmidt-Puchta, W., Dominguez, D., Lewetag, D. and Hohn, T. (1997). Plant ribosome shunting in vitro. Nucleic Acids Res. 25, 2854-2860.

Schrauwen, J. A. M., de Groot, P. F. M., van Herpen, M. M. A., van der Lee, T., Reynen, W. H., Weterings, K. A. P. and Wullems, G. J. (1990). Stage-related expression of mRNAs during pollen development in lily and tobacco. Planta 182, 298-304.

Scott, R., Dagless, E., Hodge, R., Paul, W., Soufleri, I. and Draper, J. (1991a). Patterns of gene expression in developing anthers of *Brassica napus*. Plant Mol. Biol. 17, 195-207.

Scott, R., Hodge, R., Paul, W. and Draper, J. (1991b). The molecular biology of anther differentiation. Plant Sci. 80, 167-191.

Shen, J. B. and Hsu, F. C. (1992). *Brassica* anther-specific genes: characterization and *in situ* localization of expression. Mol. Gen. Genet. 234, 379-389.

Sieburth, L. E. and Meyerowitz, E. M. (1997). Molecular dissection of the *AGAMOUS* control region shows that *cis* elements for spatial regulation are located intragenically. Plant Cell 9, 355-365.

Silverstone, A. L., Chang, C-W., Krol, E. and Sun, T-P. (1997). Developmental regulation of the gibberellin biosynthetic gene *GA1* in *Arabidopsis thaliana*. Plant J. 12, 9-19.

Smith, A. G., Gasser, C. S., Budelier, K. A. and Fraley, R. T. (1990). Identification and characterization of stamen- and tapetum-specific genes from tomato. Mol. Gen. Genet. 222, 9-16.

Smyth, D. R. (1997). Plant development: Attractive ovules. Current Biology 7, R64-R66.

Spielman, M., Preuss, D., Feng-Lan, L., Browne, W. E., Scott, R. and Dickinson, H. G. (1997). *TETRASPORE* is required for male meiotic cytokinesis in *Arabidopsis thaliana*. Development 124, 2645-2657.

Stinson, J. R., Eisenberg, A. J., Willing, R. P., Pe, M. E., Hanson, D. D. and Mascarenhas, J. P. (1987). Genes expressed in the male gametophyte of flowering plants and their isolation. Plant Physiol. 83, 442-447.

Sullivan, M. L. and Green, P. J. (1993). Post-transcriptional regulation of nuclear-encoded genes in higher plants: the roles of mRNA stability and translation. Plant Mol. Biol. 23, 1091-1104.

Sweetman, J. P. (1996). Characterization and functional analysis of pollen-specific cDNAs encoding putative transcriptional regulators from Nicotiana tabacum. Ph.D Thesis, University of Leicester, U.K.

Takatsuji, H., Mori, M., Benfey, P. N., Ren, L. and Chua, N-H. (1991). Characterization of a zinc finger DNA-binding protein expressed specifically in *Petunia* petals and seedlings. EMBO J. 11, 241-249.

Tanguay, R. L. and Gallie, D. R. (1996). Isolation and characterization of the 102-Kilodalton RNA-binding protein that binds to the 5' and 3' translational enhancers of tobacco mosaic virus RNA. J. Biol. Chem. 271, 14316-14322.

Tansey, W. P. and Herr, W. (1997). TAFs: Guilt by association ? Cell 88, 729-732.

Tebbutt, S. J., Rogers, H.J. and Lonsdale, D. M. (1994). Characterisation of a tobacco gene encoding a pollen-specific polygalacturonase. Plant Mol. Biol. 25, 283-297.

Tebbutt, S. J. and Lonsdale, D. M. (1995). Deletion analysis of a tobacco pollen-specific polygalacturonase promoter. Sex. Plant Reprod. 8, 242-246.

Thangavelu, M., Belostotsky, D., Bevan, M. W., Flavell, R. B., Rogers, H. J. and Lonsdale, D. M. (1993). Partial characterization of the *Nicotiana tabacum* actin gene family: evidence for pollen-specific expression of one of the gene family members. Mol. Gen. Genet. 240, 290-295.

Theerakulpisut, P., Xu, H., Singh, M. B., Pettitt, J. M. and Knox, R. B. (1991). Isolation and developmental expression of Bcp1, an anther specific cDNA clone in *Brassica campestris*. Plant Cell 3, 1073-1084.

Tijan, R. and Maniatis, T. (1994). Transcriptional activation: A complex puzzle with few easy pieces. Cell 77, 5-8.

Timmer, R. T., Benkowski, L. A., Schodin, D., Lax, S. R., Metz, A. M., Ravel, J. M. and Browning, K. S. (1993). The 5' and 3' untranslated regions of satellite tobacco necrosis virus RNA affect translational efficiency and dependence on a 5' cap structure. J. Biol. Chem. 268, 9504-9510.

Topping, J., Wei, W. and Lindsey, K. (1991). Functional tagging of regulatory elements in the plant genome. Development 112, 1009-1019.

Treacy, B. K., Hattori, J., Prud'homme, I., Barbour, E., Boutilier, K., Baszcynski, C. L., Huang, B., Johnson, D. A and Miki, B. L. (1997). *Bnm1*, a *Brassica* pollen-specific gene. Plant Mol. Biol. 34, 603-611.

Tsuchiya, T., Toriyama, K., Nasrallah, M. E. and Ejiri, S-I. (1992). Isolation of genes abundantly expressed in rice anthers at the microspore stage. Plant Mol. Biol. 20, 1189-1193.

Tupy, J., Suss, J., Hrabetova, E. and Rihova, L. (1983). Developmental changes in gene expression during pollen differentiation an maturation in *Nicotiana tabacum* L. Biologia. Plantarum 25, 231-237.

Tupy, J., Ribhova, L. and Zarsky, V. (1991). Production of fertile tobacco pollen from microspores in suspension culture and its storage for *in situ* pollination. Sex. Plant Reprod. 4, 284-287.

Turner, R. (1995). Sequences involved in the control of translation in potato virus S. Ph.D Thesis, University of Leicester, U.K.

Turner, R., Bate, N. Twell, D. and Foster. G. (1994). In vivo characterisation of a translational enhancer upstream from the coat protein open reading frame of potato virus S. Arch. Virol. 137, 123-132.

Turner, R. and Foster, G. D. (1995). The potential exploitation of plant viral translational enhancers in biotechnology for increased gene-expression. Molecular Biotechnology 3, 225-236.

Twell, D. (1992). Use of a nuclear-targeted ß-glucuronidase fusion protein to demonstrate vegetative cell-specific gene expression in developing pollen. Plant J. 2, 887-892.

Twell, D. (1994). The diversity and regulation of gene expression in the pathway of male gametophyte development. In Molecular and Cellular Aspects of Plant Reproduction (Scott, R, J. and Stead, A. D., eds). Cambridge: Cambridge University Press, pp. 83-115.

Twell, D. (1995). Diptheria toxin-mediated cell ablation in developing pollen: vegetative cell ablation blocks generative cell migration. Protoplasma 187, 144-154.

Twell, D., Klein, T. M., Fromm, M. E. and McCormick, S. (1989a). Transient expression of chimeric genes delivered into pollen by microprojectile bombardment. Plant Physiol. 91, 1270-1274.

Twell, D., Wing, R., Yamaguchi, J. and McCormick, S. (1989b). Isolation and expression of an anther-specific gene from tomato. Mol. Gen. Genet. 217, 240-245.

Twell, D., Yamaguchi, J. and McCormick, S. (1990). Pollen-specific gene expression in transgenic plants: Coordinate regulation of two different tomato gene promoters during microsporogenesis. Development 109, 705-713.

Twell, D., Yamaguchi, J., Wing, R. A., Ushiba, J. and McCormick, S. (1991). Promoter analysis of three genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. Genes Dev. 5, 496-507.

Twell, D., Patel, S., Sorenson, A., Roberts, M., Scott, R., Draper, J. and Foster, G. (1993). Activation and developmental regulation of an *Arabidopsis* anther-specific promoter in microspores and pollen of *Nicotiana tabacum*. Sex. Plant Reprod. 6, 217-224.

Ueda, T., Waverczak, W., Ward, K., Sher, N., Ketudat, M., Schmidt, R. J. and Messing, J. (1992). Mutations of the 22- and 27-KD zein promoters affect transactivation by the Opaque-2 protein. Plant Cell 4, 701-709.

Ursin, V. M., Yamaguchi, J. and McCormick, S. (1989). Gametophytic and sporophytic expression of anther-specific genes in developing tomato anthers. Plant Cell 1, 727-736.

Valenta, R., Duchene, M., Pettenburger, K., Sillaber, C., Valent, P., Bettelheim, P., Breitenbach, M., Rumpold, H., Kraft, D. and Scheiner, O. (1991). Identification of profillin as a novel pollen allergen: IgE autoreativity in sensitised individuals. Science 253, 557-560.

van der Meer, I. M., Brouwer, M., Spelt, C. E., Mol, J. N. M. and Stuitje, A. R. (1993). The TACPyAT repeats in the chalcone synthase promoter of *Petunia hybrida* act as a dominant negative *cis*-acting module in the control of organ-specific expression. Plant J. 2, 525-535.

Villalba, M., Batanero, E., Lopez-Otin, C., Sanchez, L. M., Monsalve, R. I., Gonzalez de la Pena, M. A., Lahoz, C. and Rodriguez, R. (1993). The amino acid sequence of Ole e I, the major allergen from olive tree (Olea europaea) pollen. Eur. J. Biochem. 216, 863-869.

Villemur, R., Haas, N. A., Joyce, C. M. Snustad, D. P. and Silflow, C. D. (1994). Characterization of four new ß-tubulin genes and their expression during male flower development in maize (*Zea mays* L.) Plant Mol. Biol. 24, 295-315.

Vicente-Carbajosa, J., Moose, S. P., Parsons, R. L. and Schmidt, R. J. (1997). A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator *opaque-2*. Proc. Natl. Acad. Sci. USA 94, 7685-7690.

Vincentz, M., Leite, A., Neshich, G., Vriend, G., Mattar, C., Barros, L., Weinberg, D., de Almeida, E. R., Paes de Carvalho, M., Aragao, F. and Gander, E. S. (1997). ACGT and vicilin core sequences in a promoter domain required for seed-specific expression of a *2S* storage protein gene are recognised by the *opaque-2* regulatory protein. Plant Mol. Biol. 34, 879-889.

Vogel, J. M., Roth, B., Cigan, M. and Freeling, M. (1993). Expression of the two maize TATA binding protein genes and function of the encoded TBP proteins by complementation in yeast. Plant Cell 5, 1627-1638.

Wade, P. A. and Wolffe, A. P. (1997). Chromatin: Histone acetyltransferases in control. Current Biol. 7, R82-R84.

Wagner, V. T., Cresti, M., Salvatici, P. and Tiezzi, A. (1990). Changes in volume, surface area, and frequency of nuclear pores on the vegetative nucleus of tobacco pollen in fresh, hydrated and activated conditions. Planta 181, 304-309.

Wang, M. and Reed, R. R. (1993). Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. Nature 364, 121-126.

Wang, S. and Miller, W. A. (1995). a sequence located 4.5 to 5 Kilobases from the 5' end of the barley yellow dwarf virus (PAV) genome strongly stimulates translation of uncapped mRNA. J. Biol. Chem. 270, 13446-13452.

Wang, Z-Y., Kenigsbuch, D., Sun, L., Harel, E., Ong, E. H. and Tobin, E. M. (1997). A myb-related transcription factor is involved in the phytochrome regulation of an *Arabidopsis Lhcb* gene. Plant Cell 9, 491-507.

Weisshaar, B., Armstrong, G. A., Block, A., da Costa e Silva, O. And Hahlbrock, K. (1991). Light-inducible and constitutively expressed DNA-binding proteins recognizing a plant promoter element with functional relevance in light responsiveness. EMBO J. 10, 1777-1786.

Weterings, K. (1994). Pollen gene regulation. Ph.D Thesis, University of Nijmegen, Nijmegen, The Netherlands.

Weterings, K., Reijnen, W., Aarssen, R., Kortstee, A., Spijkers, J., Harpen, M., Schrauwen, S. and Wullems, G. (1992). Characterisation of a pollen-specific cDNA clone from Nicotiana Tabacum expressed during microgametogenesis and germination. Plant Mol. Biol. 18, 1101-1111.

Weterings, K., Schrauwen, S., Wullems, G. and Twell, D. (1995). Functional dissection of the promoter of the pollen-specific gene *NTP303* reveals a novel pollen-specific, and conserved *cis*-regulatory element. Plant J. 8, 55-63.

Willing, R. P. and Mascharenhas, J. P. (1984). Analysis of the complexity and diversity of mRNAs from pollen and shoots of *Tradescantia.* Plant Physiol. 75, 865-868.

Willing, R. P., Bashe, D. and Mascharenhas, J. P. (1988). An analysis of the quantity and diversity of messenger RNAs from pollen and shoots of *Zea mays* . Theor. Appl. Genet. 75, 751-753.

Wing, R. A., Yamaguchi, J., Larabell, S. K., Ursin, V. M. and McCormick, S. (1989). molecular and genetic characterization of two pollen-expressed genes that have sequence similarity to pectate lyases of the plant pathogen *Erwinia.* Plant Mol. Biol. 14, 17-28.

Wong, E. Y., Hironaka, C. M. and Fischloff, D. A. (1992). *Arabidopsis thaliana* small subunit leader and transit peptide enhance the expression of *Bacillus thuringiensis* proteins in transgenic plants. Plant Mol. Biol. 20, 81-93.

Worrall, D., Hird, D. L., Hodge, R., Paul, W., Draper, J. and Scott, R. (1992). Premature dissolution of the microsporocyte callose wall causes male sterility in transgenic tobacco. Plant Cell 4, 759-771.

Wu, Y., Qiu, X., Du, S. and Erickson, L. (1996). PO149, a new member of pollen pectate lyase-like gene family from alfalfa. Plant Mol. Biol. 32, 1037-1042.

Xiang, C., Miao, Z. and Lam, E. (1997). DNA-binding properties, genomic organization and expression pattern of *TGA6*, a new member of bZIP transcription factors in *Arabidopsis thaliana.* Plant Mol. Biol. 34, 403-415.

Xu, H., Knox, R. B., Taylor, P. E. and Singh,M. B. (1995). *Bcp1*, a gene required for male fertility in arabidopsis. Proc. Natl. Acad. Sci. USA 92, 2106-2110.

Xu, R., Goldman, S., Coupe, S. and Deikman, J. (1996). Ethylene control of E4 transcription during tomato fruit ripening involves two cooperative *cis* elements. Plant Mol. Biol. 31, 1117-1127.

Yamamoto, Y. Y., Tsuji, H. and Obokata, J. (1995). 5'-Leader of a photosystem I gene in *Nicotiana sylvestris*, *psaDb*, contains a translational enhancer. J. Biol. Chem. 270, 12466-12470.

Yamamoto, Y. Y., Kondo, Y., Kato, A., Tsuji, H. and Obokata, J. (1997). Light-responsive elements of the tobacco *PSI-D* gene are located both upstream and within the transcribed region. Plant J. 12, 255-265.

Yamazaki, K-I., Katagiri, F., Imaseki, H. and Chua, N-H. (1990). TGA1a, a tobacco DNA-binding protein, increases the rate of initiation in a plant *in vitro* transcription system. Proc. Natl. Acad. Sci. USA 87, 7035-7039.

Yanagisawa, S. and Izui, K. (1992). MNF1, a leaf tissue-specific DNA-binding protein of maize, interacts with the cauliflower mosaic virus 35S promoter as well as the C_4 photosynthetic phosphoenolpyruvate carboxylase gene promoter. Plant Mol. Biol. 19, 545-553.

Yang, Q. and Sarnow, P. (1997). Location of the internal ribosome entry site in the 5' non-coding region of the immunoglobin heavy-chain binding protein (BiP) mRNA: evidence for specific RNA-protein interactions. Nucleic Acids Res. 25, 2800-2807.

Yunes, J. A., Cord Neto, G., da silva, M., Leite, A., Ottoboni, L. M. M. and Arruda, P. (1994). The transcriptional activator *opaque-2* recognises two different target sequences in the 22-KD-like α-prolamin genes. Plant Cell 6, 237-249.

Zachgo, S., Saedler, H. and Schwarz-Sommer, Z. (1997). Pollen-specific expression of *DEFH125*, a MADS-box transcription factor in *Antirrhinum* with unusual features. Plant J. 11, 1043-1050.

Zappavigna, V., Falciola, L., Citterich, M. H., Mavilio, F. and Bianchi, M. E. (1996). HMG1 interacts with HOX proteins and enhances their binding and transcriptional activation. Embo J. 15, 4981-4991.

Zhu, Q., Doerner, P. W. and Lamb, C. J. (1993). Stress induction and developmental regulation of a rice chitinase promoter in transgenic tobacco. Plant J. 3, 203-212.

Zhu, Q., Dabi, T. and Lamb, C. (1995). TATA box and initiator functions in the accurate transcription of a plant minimal promoter *in vitro*. Plant Cell 7, 1681-1689.

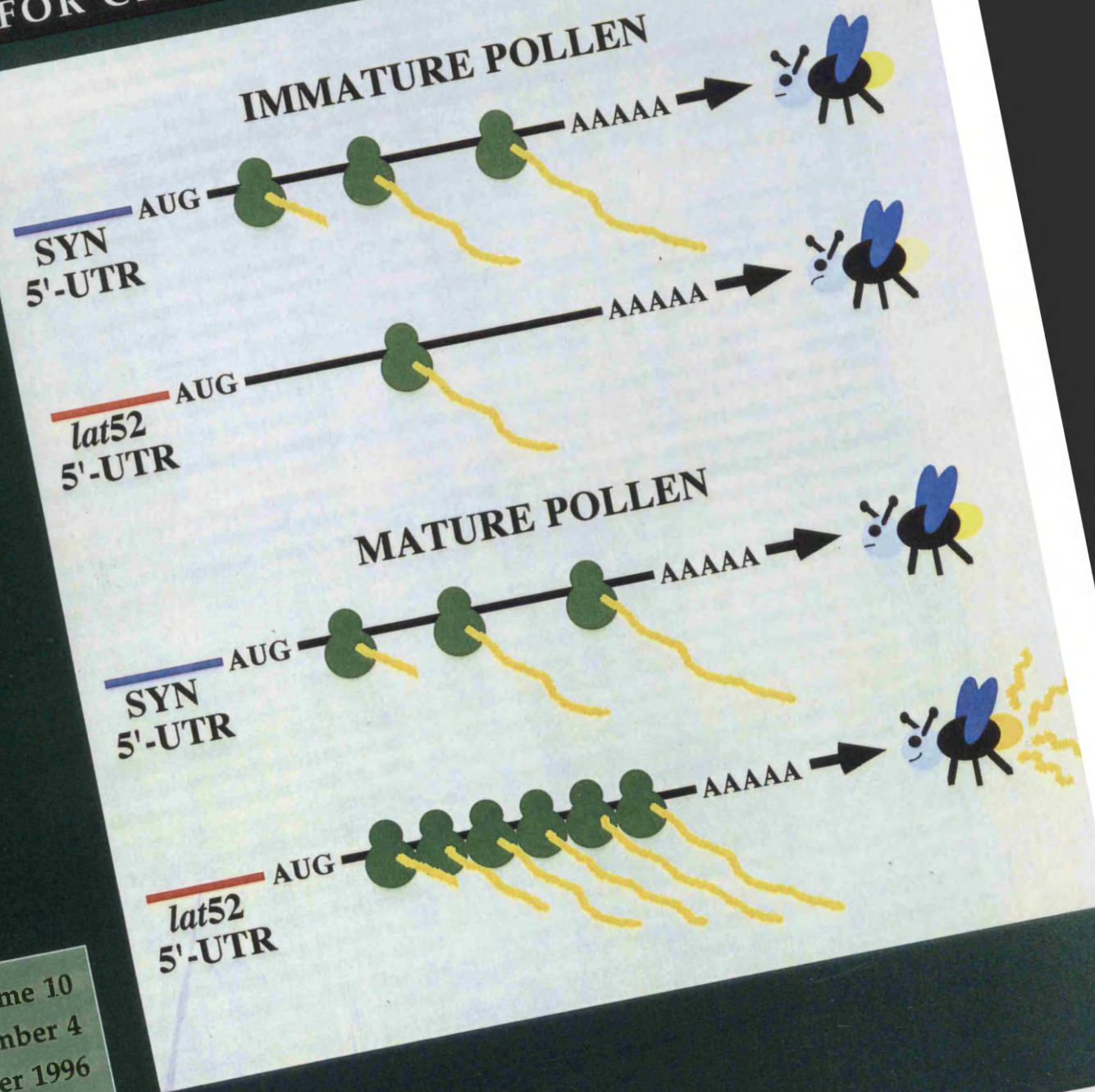
Zou, J.-T., Zhan, X.-Y., Wu, H.-M., Wang, H. and Cheung, A. Y. (1994). Characterization of a rice pollen-specific gene and its expression. American Journal of Botany 81, 552-561.

Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. Science 244, 48-52.

CODEN PLJUED ISSN 0960-7412

THE plant JOURNAL

FOR CELL AND MOLECULAR BIOLOGY



Volume 10
Number 4
October 1996

The Plant Journal (1996) **10**(4), 613–623

Maturation-specific translational enhancement mediated by the 5′-UTR of a late pollen transcript

N. Bate, C. Spurr, G.D. Foster and D. Twell*
Department of Botany, University of Leicester, Leicester LE1 7RH, UK

Summary

The tomato *lat*52 gene encodes an abundant protein specifically expressed in the vegetative cell of the pollen grain during pollen maturation which is essential for normal pollen tube growth. Multiple upstream *cis*-regulatory elements controlling the level and specificity of *lat*52 transcription have previously been identified. This research investigated the role of the 5′-untranslated region (5′-UTR) in controlling *lat*52 expression. In transient expression assays, gene fusion constructs containing the *lat*52 5′-UTR were expressed in pollen at levels 13- to 60-fold above those in which synthetic polylinker sequences replaced the *lat*52 5′-UTR. This enhancement was shown to be independent of both the promoter sequences, the linked reporter gene and the 3′-UTR. Analysis of RNA and protein levels in transgenic plants containing such gene fusions demonstrated that the *lat*52 5′-UTR conferred a dramatically increased translational yield to heterologous transcripts in a pollen-specific and strictly developmentally regulated manner during the final stages of pollen maturation. These results represent a novel example of translational enhancement in plants in that translational yield is regulated developmentally in a cell-specific manner via sequences located within the 5′-UTR.

Introduction

The importance of transcriptional controls in regulating both the level and specificity of expression in higher eukaryotes is well established, commonly being determined by the distribution and combinatorial activities of sequence-specific DNA-binding proteins. In contrast, the contribution of post-transcriptional processes in controlling these processes is far less well understood. However, a few well-documented examples exist in the literature of gene regulation at the level of translation, which demonstrate that sequences within the 5′ untranslated regions (5′-UTRs) of certain viral and native cellular RNAs are used to modulate the rate of translational initiation (for recent reviews see Hershey, 1991; Sonenberg, 1994). First, the translational repression of ferritin mRNA is mediated through the binding of a repressor protein to the iron responsive element in the 5′-UTR (Klausner *et al.*, 1993). Second, specific regions of the 5′-UTRs of several animal viruses from the picornavirus group, are recognized as internal ribosome entry sites (Meerovitch and Sonenberg, 1993; Oh and Sarnow, 1993).

In plants several examples of translational control have been documented (Kuhlemeier, 1992). For example, it is well established that the 5′-UTRs present on several plant virus RNAs possess the ability to increase the translational yield (reviewed in Sleat and Wilson, 1992). The mechanism of action of these efficient 5′-UTRs is not clear, although they appear to act at the level of translational initiation rather than elongation, RNA processing or mRNA stability. Whether such sequences enhance translation differentially in different cell types and/or are modulated during development has not been investigated.

Translational enhancers have only recently been identified within native plant genes. These include sequences within the 5′-UTRs of the photosystem I gene *psaDb* (Yamamoto *et al.*, 1995) and the plasma membrane protein-ATPase *pma1* (Michelet *et al.*, 1994). In these examples evidence was provided that reporter gene activity was enhanced at the level of translation either in transgenic plants (*psaDb*) or from RNA transfection studies in protoplasts (*pma1*). However, enhancement was not shown to be independent of the linked coding sequence or the 3′-UTR and no evidence for possible developmental or tissue-specific translational control was presented. One example of a translational enhancer which does show some degree of tissue specificity is associated with a barley α-amylase gene (Gallie and Young, 1994). From RNA transfection studies the α-amylase 5′- and 3′-UTR regions when fused to the reporter gene luciferase (*luc*) enhanced activity in maize aleurone and endosperm protoplasts but not in suspension cells. Other examples of 5′-UTR-mediated enhancement of reporter gene activity in transgenic plants have been described (Casper and Quail, 1993; Kuhlemeier, 1992; Sullivan and Green, 1993). These remain to be rigorously investigated since in at least one case enhancement has been shown to result from the presence of transcriptional regulatory elements within the 5′-UTR (Bolle *et al.*, 1994). A further example of translational control is one modulated by an environmental signal in which preferential translation of *adh1* (Fennoy and Bailey-Serres, 1995) occurs during hypoxia. Thus, to our knowledge there have been no reports in which it has been established that the 5′-UTR of a native gene transcript acts as a translational enhancer in a precisely defined developmental and tissue-

Received 7 March 1996; revised 31 May 1996; accepted 28 June 1996.
*For correspondence (fax +44 116 2522791; e-mail twe@le.ac.uk).

specific manner. The identification and characterization of such examples would provide significant new insight into mechanisms which regulate the expression of gene products during development.

We are investigating the control of cell-specific and coordinate gene expression during pollen development. A number of highly abundant 'late-pollen'-specific transcripts appear in the developing pollen grain after pollen mitosis I and dramatically accumulate during maturation (Eady *et al.*, 1994, 1995; Twell *et al.*, 1989a; Ursin *et al.*, 1989). It is well established that the level and specificity of expression of several late pollen genes is controlled by both shared and unique transcriptional *cis*-regulatory elements (Eyal *et al.*, 1995; Twell *et al.*, 1990, 1991). One of these genes, *lat*52, encodes a cysteine-rich kunitz trypsin inhibitor-related protein, which is abundant in the mature pollen grain (Muschietti *et al.*, 1994). *lat*52 transcripts persist for at least 16 h following pollen germination and tube growth (Ursin *et al.*, 1989). Thus, *lat*52 transcripts are translated during pollen development, but are also stored in the mature pollen grain, together with LAT52 protein, for their utilization during pollen tube growth. With regard to its putative role in the growing pollen tube, there appears to be a particular requirement for high levels of LAT52 protein during pollen maturation and early pollen tube growth, since antisense-mediated down regulation of LAT52 protein levels prevents normal pollen hydration and tube growth (Muschietti *et al.*, 1994). Furthermore, the LAT52 homologous, olive tree pollen-allergen, *Ole e* I, has been reported to represent about 1% of the dry mass of mature pollen (Villalba *et al.*, 1993).

Previous studies have established the minimal upstream region of the *lat*52 gene capable of directing pollen-specific expression as the region −71 to +110 bp (Twell *et al.*, 1991). This region includes the TATA box at position −28 bp and the entire 110 nt 5′-UTR. Because of the requirement for high levels of LAT52 protein specifically in mature pollen, we have investigated the role of the *lat*52 5′-UTR in contributing to both the level and specificity of *lat*52 expression. Our results clearly show that the *lat*52 5′-UTR contains sequences which dramatically enhance the translational yield of chimeric transcripts in a pollen-specific manner strictly during the final stages of pollen maturation.

Results

The presence of the lat*52 5′-UTR dramatically enhances transient expression in pollen*

To investigate the role of sequences comprising the *lat*52 5′-UTR, the expression of gene fusion constructs was analysed using a microprojectile-bombardment-mediated transient expression assay in tobacco pollen (Twell *et al.*, 1989b). In this assay, plasmids containing promoter–5′-UTR–luciferase (*luc*) reporter gene (Ow *et al.*, 1986) fusions (test plasmids), were co-precipitated with a reference plasmid expressing the β-glucuronidase (*gus*) reporter gene (Jefferson *et al.*, 1987) and bombarded into plant cells. The reference plasmid serves as an internal control correcting for the variability between independent bombardments since both plasmids are targeted to the same cell with similar efficiencies. Luciferase (LUC) activities were corrected by deriving a ratio of LUC:GUS and relative activity of each test plasmid was then calculated. Initially, native *lat*52 sequences upstream of the initiator ATG (−492 to +110 bp), including the entire 110 bp 5′-UTR, were linked to the *luc* reporter gene by fusion at the initiator ATG with an *Nco*I restriction site to create pNBL52-5 (Figure 1a). An otherwise identical plasmid was built (pNBL52-42) in which the *lat*52-5′-UTR region was replaced by a synthetic polylinker (Figure 1a). The activities of these plasmids were assayed following their introduction into mature tobacco pollen which was then cultured *in vitro*. The results showed that bombardment of pollen with pNBL52-5 resulted in normalized luciferase (LUC) activities that were 24-fold higher than those observed with pNBL52-42 (Figure 2a). Thus, sequences within the *lat*52 5′-UTR play an important role in contributing to the high level of *lat*52 expression.

Enhancement is independent of promoter and linked reporter gene

To investigate whether the enhancement mediated by the *lat*52 5′-UTR was dependent upon upstream sequences within the *lat*52 promoter, a gene fusion was built in which the cauliflower mosaic virus 35S (*CaMV35S*) promoter replaced *lat*52 promoter sequences, in plasmid pNBL52-26 (Figure 1b). The *CaMV35S* promoter is active at high levels in most sporophytic cell types in tobacco, but only minimal levels of activity are detectable in pollen (Twell *et al.*, 1989b). A similar control construct containing synthetic polylinker sequences replacing the *lat*52 5′-UTR region was built (Figure 1b), and their activities assayed in pollen, leaves and in cells grown in suspension using the bombardment assay. The results showed that the presence of the *lat*52 5′-UTR dramatically enhanced the expression of constructs containing the *CaMV35S* promoter in pollen compared with those containing the SYN 5′-UTR (Figure 2a). Levels of enhancement in pollen (approximately 14-fold) were less than two-fold lower than those observed when the *lat*52 5′-UTR was present in its native context downstream of the *lat*52 promoter (Figure 2a). Therefore, enhancement was mainly independent of the upstream transcriptional regulatory sequences. Interestingly, the absolute levels of LUC activity in pollen directed by otherwise identical constructs driven by the *lat*52 promoter were at least 50-fold higher (data not shown) than those driven by the *CaMV35S* promoter, which further

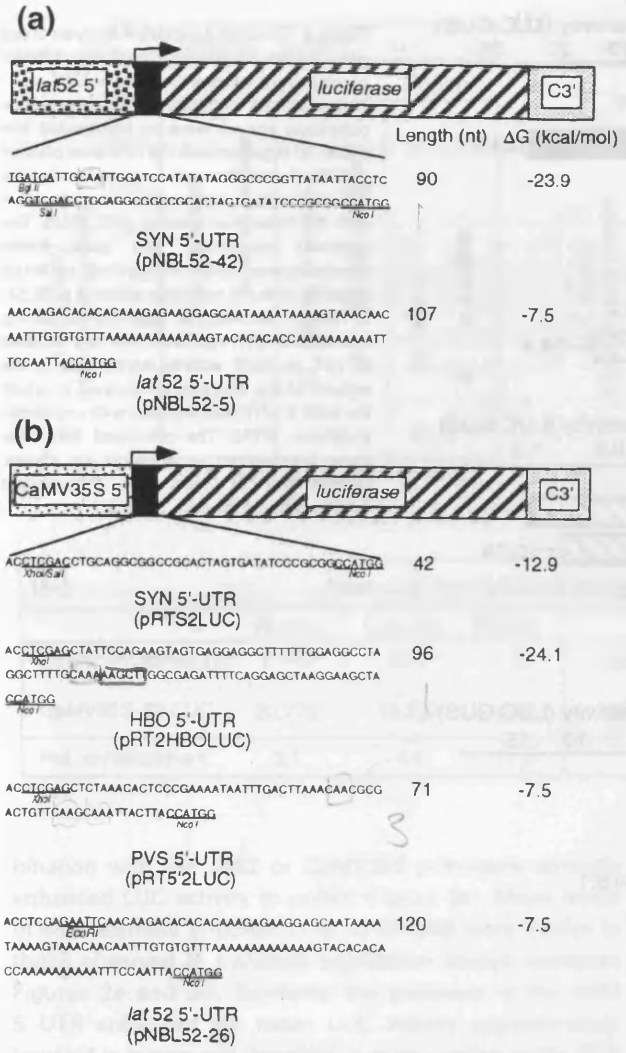


Figure 1. Test gene fusion constructs containing different 5′-UTRs used in transient expression analysis.
Constructs present in test plasmids containing the *lat*52 promoter (a) or *CaMV35S* promoter (b) are shown with the nucleotide sequences beginning at +1 of the different 5′-UTRs. Promoter sequences were linked to the *luc* reporter gene and the 3′-UTR (C3′) from *CaMV35S*. The length, in nucleotides, and the calculated stability (ΔG in kcal mol^{-1}) of secondary structure within the 5′-UTRs are shown on the right. Restriction sites used in plasmid construction and initiator ATGs are underlined.

suggested that the enhancement observed was independent of the intrinsic strength of transcriptional regulatory sequences and therefore, of the transcript abundance. These results suggested that enhanced LUC expression was the result of increased translational yield and that the *lat*52 5′-UTR contains a translational enhancer.

Since translational enhancement mediated by certain plant viral sequences has been reported to be strongly dependent on the linked protein coding sequences (Sleat and Wilson, 1992), the ability of the *lat*52 5′-UTR to enhance the expression of the *gus* reporter gene was tested. The activities of two plasmids were compared, in which the *LUC–CaMV35S3′* sequences in plasmids pNBL52-5 and pNBL52-42 were replaced with a *GUS–NOS3′* cassette. The results showed that the *lat*52 5′-UTR enhanced expression approximately 60-fold in pollen compared with the SYN 5′-UTR construct (Figure 2d). This dismisses the possibility of either 5′-UTR/coding region or 5′-UTR/3′-UTR interactions as the sole cause of the enhanced expression of constructs containing the *lat*52 5′-UTR.

To exclude the possibility that the *lat*52 5′-UTR-mediated enhancement in pollen resulted from the inefficient translation of the particular synthetic leader sequence used for comparison, two additional plasmids containing either positive or negative control 5′-UTRs were built. Both of these 5′-UTRs were constructed to contain an identical *Nco*I site at the initiator ATG, and in addition possessed optimal translational context at the –3 and +4 positions (Figure 1b). The negative control comprised a longer control 5′-UTR (96 nt), present in plasmid pRT2HBOLUC (Figure 1b). This 5′-UTR has previously been shown to be translated efficiently in mammalian cells, relative to modified versions in which increasingly stable secondary structures have been introduced (Koromilas *et al.*, 1992). To verify this *in planta*, the construct pRT2HBOLUC and versions with increasing secondary structure were transiently expressed in pollen (Figure 2a). These plasmids showed a similar pattern of relative activity to that observed in mammalian cells. Increasing the potential for secondary structure of HB0 from a ΔG of –24.1 to –33.1 (HB1) had little effect on LUC activity, but an increase in ΔG to – 59.6 (HB3) resulted in an eightfold reduction of LUC activity. This suggests that 5′-UTRs with ΔG values less than –33 are translated relatively efficiently in pollen. The positive control comprised the 5′-UTR of the potato virus S 5′-UTR, in plasmid pRT5′2LUC, which we have shown to enhance translational efficiency both *in vitro*, and *in vivo* (Turner *et al.*, unpublished). In the bombardment assay in pollen, the negative control, HBO 5′-UTR construct, was expressed at similar levels to the SYN 5′-UTR construct (Figure 2a). In contrast, the positive control, PVS 5′-UTR construct, enhanced expression approximately fivefold in pollen (Figure 2a). Thus, enhancement mediated by the *lat*52 5′-UTR was significantly above that of a known viral 5′-UTR, and independent of the length of the synthetic negative control 5′-UTR.

The lat*52 5′-UTR preferentially enhances gene expression in pollen*

To investigate the ability of the *lat*52 5′-UTR to enhance gene expression in sporophytic cell types, the activities of the *luc*-based constructs were assayed after their introduction into leaves and suspension-cultured cells by bombardment. In striking contrast to the results obtained in pollen, the presence of the *lat*52 5′-UTR led only to a modest

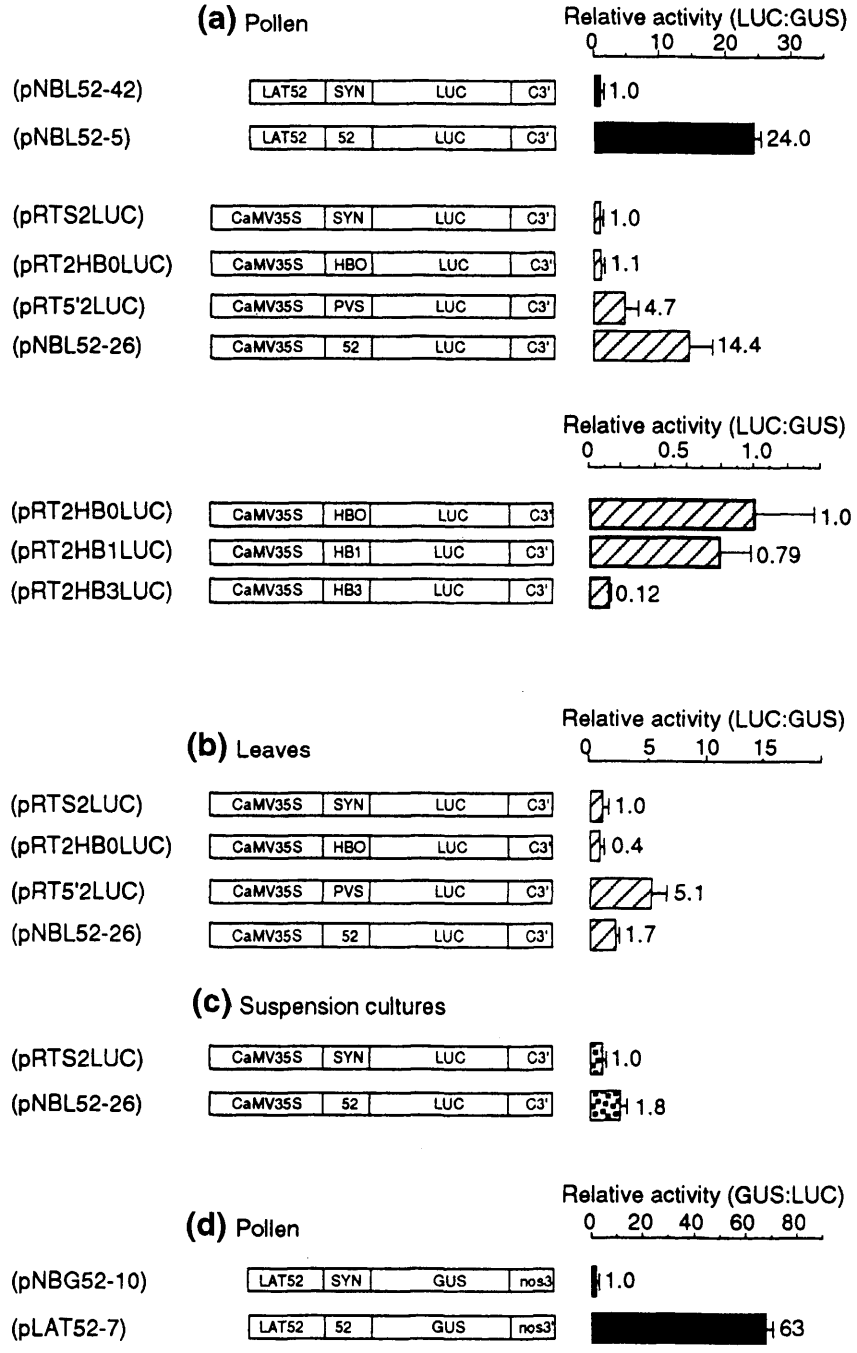


Figure 2. Transient expression analysis of test gene fusion constructs containing different combinations of promoter and 5'-UTR.
Test plasmids containing *luc* gene fusion constructs shown were co-bombarded into pollen (a) together with the reference plasmid pLAT52–7. Test plasmids were co-bombarded into leaves (b) or suspension cells (c) together with the reference plasmid pRTL2GUS. Test plasmids containing *gus* gene fusion constructs were co-bombarded into pollen (d) together with the reference plasmid pNBL52–5. Relative activities of plasmids containing the *lat*52 5'-UTR represent the fold increase in LUC or GUS activity normalized to the activity of the equivalent construct in which the *lat*52 5'-UTR was replaced with a synthetic polylinker (SYN). The combined data from three independent experiments are shown, representing at least six independent bombardments for each plasmid.

(approximately twofold) enhancement of expression compared with the SYN 5'-UTR control (Figure 2b, c). The low levels of enhancement were reproducible and similar in leaves and suspension cells. These results show that the activity of the *lat*52 5'-UTR is dramatically increased specifically in pollen. In contrast the PVS and HBO 5'-UTRs did not show substantial differences in their abilities to enhance gene expression in pollen or leaves (Figure 2a and b). These data demonstrate a strong preferential activity of the *lat*52 5'-UTR in pollen.

To determine whether the *lat*52 5'-UTR-mediated enhancement observed in transient expression assays also occurred *in planta*, the four gene fusion constructs, *52–SYN–LUC*, *52–52–LUC*, *35S–SYN–LUC* and *35S–52–LUC* (Figure 1) were introduced into tobacco by *Agrobacterium*-mediated transformation. Thirty to thirty-five kanamycin resistant-primary transformants containing each construct were regenerated and grown to maturity. LUC activity was assayed in extracts from roots, leaves and pollen of these plants. The results showed that the *lat*52 5'-UTR in com-

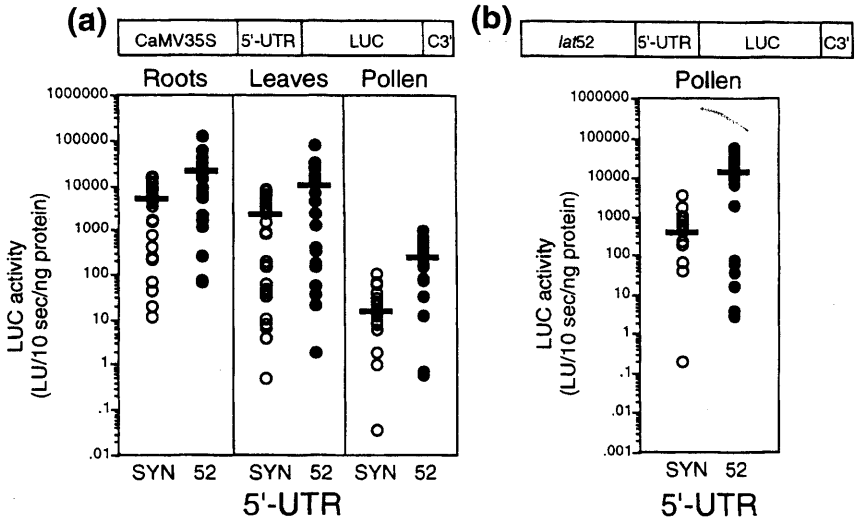


Figure 3. Analysis of *lat*52- 5'-UTR activity in transgenic plants.
The scatter histograms show LUC activities determined in different tissues of transgenic plants containing gene fusion constructs driven by the *CaMV35S* (a) or *lat*52 (b) promoters containing either the *lat*52 5'-UTR or a synthetic polylinker 5'-UTR. Each point represents the value derived from the analysis of a single independent transgenic plant, with the mean values indicated by a black bar. The data summarized below show the mean values and the relative enhancement mediated by the *lat*52- 5'-UTR in each tissue.

(c) Mean LUC Activity (LU/10 sec/ng protein)

	Roots	Leaves	Pollen		Pollen
CaMV35S-**SYN**-LUC	6,780	2,664	16	*lat*52-**SYN**-LUC	16,007
CaMV35S-**52**-LUC	20,772	12,232	286	*lat*52-**52**-LUC	449,718
Rel. enhancement	3.1	4.6	17.8		28.1

bination with the *lat*52 or *CaMV35S* promoters strongly enhanced LUC activity in pollen (Figure 3a). Mean levels of enhancement in pollen (~18- to 28-fold) were similar to those observed in transient expression assays (compare Figures 2a and 3c). Similarly, the presence of the *lat*52 5'-UTR enhanced the mean LUC activity approximately fourfold in leaves and threefold in roots relative to the SYN 5'-UTR control (Figure 3c). These data confirm the ability of the *lat*52 5'-UTR to act as an enhancer of gene expression when stably integrated into the plant genome. These data further substantiate the results of the transient expression experiments above, in that the *lat*52 5'-UTR was highly preferentially active in developing pollen.

The lat*52 5'-UTR increases translational yield in pollen but not in leaves*

To directly investigate whether the *lat*52 5'-UTR acted at the transcriptional or post-transcriptional level, steady-state levels of *luc* transcript and LUC activity were measured in leaves of *CaMV35S–SYN–LUC* and *CaMV35S–52–LUC* plants and in pollen of *lat52–SYN–LUC* and *lat52–52–LUC* plants. The relative translational yield of *luc* mRNA in individual plants was calculated after quantification of *luc* mRNA abundance corrected for differences in RNA loading on Northern blots. The results of the RNA and LUC analysis carried out on three to five independent transformed plants containing each construct are presented in Figure 4. In leaves the mean translational yield of plants synthesising *52–LUC* transcripts was slightly above (1.3-fold) that of plants synthesizing *SYN–LUC* transcripts (Figure 4a). In contrast, plants synthesizing *52–LUC* transcripts showed 7- to 40-fold higher translational yields compared with plants synthesizing *SYN–LUC* transcripts (Figure 4b). Thus, the mean translational yield from *52–LUC* transcripts, relative to *SYN–LUC* transcripts, was approximately 13-fold in pollen and 1.3-fold in leaves. These data demonstrate that the presence of the *lat*52 5'-UTR dramatically increases translational yield specifically in pollen.

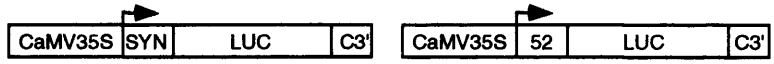
Developmental regulation of the translational activity of the lat*52 5'-UTR*

Given the specific activity of the *lat*52 5'-UTR in the mature male gametophyte we investigated whether translational enhancement was uniform during pollen development or became apparent at specific developmental stages. LUC activity was determined in anthers of two groups of three transgenic lines expressing *52–LUC* or *SYN–LUC* transcripts under the control of the *lat*52 promoter at 15–20 developmental stages spanning microspore release to mature dehisced pollen (bud lengths 10 to more than or equal to 52 mm). Figure 5(a) illustrates that the developmental profiles of LUC accumulation for *52–SYN–LUC* and

52–52–LUC plants are strikingly different. This is most clearly illustrated for all three individuals within both groups of plants when LUC accumulation profiles are plotted as a percentage of maximum (Figure 5b). While LUC levels increased in an approximately linear manner during pollen development for *52–SYN–LUC* plants, *52–52–LUC* plants showed a dramatic acceleration in LUC accumulation during the final stages of pollen maturation in buds greater than 50 mm. *52–52–LUC* plants consistently accumulated greater than 80% of their final LUC activity

(a)

LEAVES



LUC activity	1,350	1,510	1,670	2,580	371	4,320	4,370	14,100
Rel. LUC activity	3.63	4.07	4.5	6.95	1.0	11.64	11.78	38.0
luc mRNA counts	64,492	61,428	150,157	57,973	32,922	172,473	99,257	355,674
rRNA counts	7,931	5,936	10,437	8,741	8,443	8,959	9,987	8,156
Rel. *luc* mRNA	8.1	10.3	14.4	6.6	3.9	19.3	9.9	43.6
Rel. translational yield	0.5	0.4	0.3	1.1	0.3	0.6	1.2	0.9

Mean TL Yield	0.6 ± 0.4	0.8 ± 0.4
Rel. Mean TL Enhancement	1.0	1.3

(b)

POLLEN

LUC activity	170	600	610	4,000	6,300	19,000	57,500	59,000
Rel. LUC activity	1.0	3.52	3.59	23.5	37.1	118.8	338.2	347.1
luc mRNA counts	6,840	5,070	6,169	4,408	7,456	21,571	83,245	81,032
rRNA counts	69,014	46,999	68,990	57,307	73,711	79,811	72,936	86,370
Rel. *luc* mRNA	0.10	0.11	0.09	0.08	0.10	0.27	1.14	0.94
Rel. translational yield	10	32	40	294	371	440	297	369

Mean TL Yield	27.3 ±15.5	354 ± 61
Rel. Mean TL Enhancement	1.0	13.0

during the final few millimetres of development, a period of 24–36 h (Twell, 1994). In contrast, during early to mid-pollen development (20–40 mm bud lengths) LUC accumulation profiles were similar for both groups (Figure 5a inset) and both were initiated at approximately 20 mm bud length.

To establish whether the final burst of LUC accumulation was the result of increased translation of *52–LUC* transcripts, the translational yield was determined at different developmental stages in the *52–SYN–LUC* and *52–52–LUC* lines analysed in Figure 5(a). Translational yield was determined initially at three discrete stages, 40–45 mm, greater than or equal to 52 mm and in dehisced mature pollen. The results showed that *SYN–LUC* transcripts had a similar translational yield at each stage but *52–LUC* transcripts showed an eightfold increase in translational yield in buds larger than or equal to 52 mm and a further sixfold increase in mature pollen (Figure 6a). It was also notable that *52–LUC* transcripts showed a lower (2.5-fold) translational efficiency than *SYN–LUC* transcripts in 40–45 mm buds. To confirm and extend these data two further analyses of translational yield were performed in which two earlier developmental stages were included. The results of these experiments showed that the translational yield of *SYN–LUC* transcripts was similar throughout pollen development and that *52–LUC* transcripts consistently showed reduced translational yield (2.5- to threefold) in buds less than 45 mm (Figure 6b). The translational yield of *52–LUC* transcripts subsequently increased in buds of 50 mm reaching a maximum in mature dehisced pollen. Throughout pollen development, (bud length 20 mm to mature pollen) the translational yield of *SYN–LUC* transcripts was similar with *52–LUC* transcripts showing an overall increase in translational yield of at least 40-fold (Figure 6b). These data strongly suggest that developmentally regulated mechanisms exist to preferentially repress or enhance the translatability of *lat*52 5′-UTR-containing transcripts during pollen development.

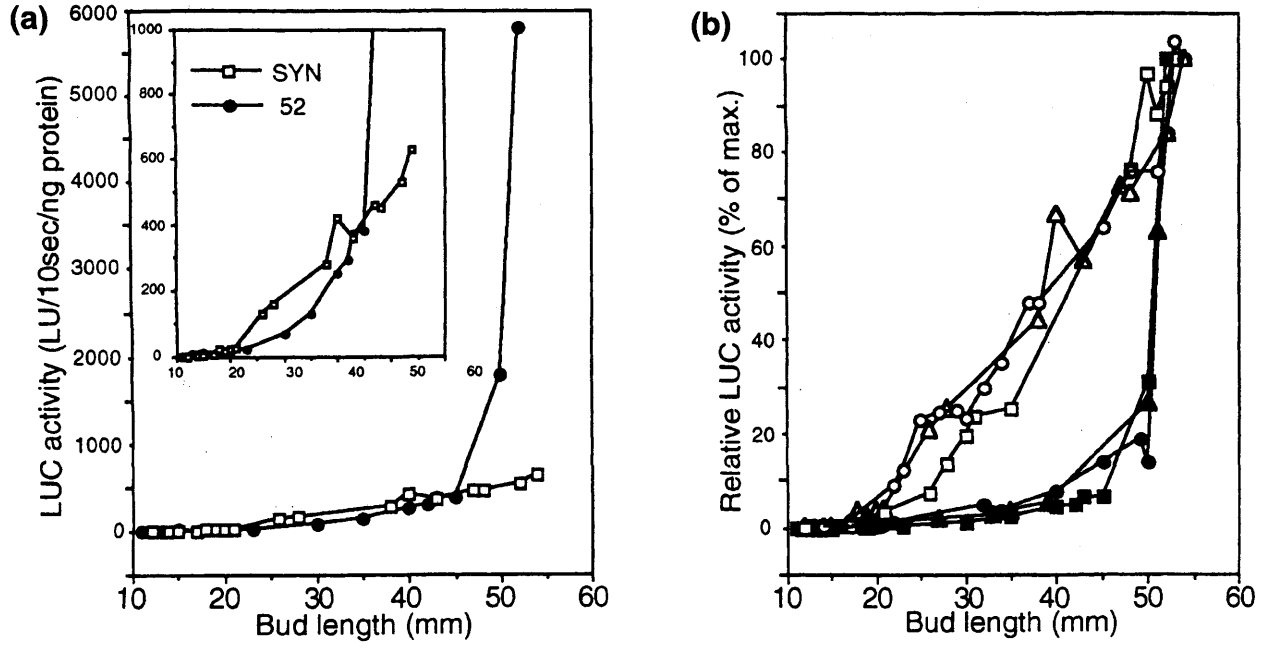


Figure 5. Analysis of LUC accumulation profiles during anther development in transgenic tobacco plants synthesizing *SYN–LUC* and *52–LUC* transcripts driven by *lat*52 promoter sequences.
(a) LUC activity accumulation profiles for one representative *SYN–LUC* (open boxes) and one *52–LUC* (closed circles) plant. The inset shows the same data with an extended Y-axis to illustrate similar point of initiation of LUC synthesis and similar initial rate of LUC accumulation in both plants.
(b) LUC accumulation profiles for three *SYN–LUC* (open symbols) and three *52–LUC* (closed symbols) plants plotted as the percentage of maximum LUC activity.

Figure 4. Analysis of the translational yield from chimeric transcripts synthesized in mature pollen and leaves of transgenic plants.
Levels of LUC activity (light units per 10 sec per mg protein $\times 10^6$) were determined for leaf (a) or pollen (b) tissue sample extracts of several independent transformants containing the gene fusion constructs *CaMV35S–SYN–LUC, CaMV35S–52–LUC* or *lat52–SYN–LUC, lat52–52–LUC*, respectively. LUC activity is also shown as Relative LUC activity after normalizing to the value of the lowest individual. RNA was isolated from the remainder of the same tissue samples and Northern blot analysis used to determine *luc* transcript and rRNA (after re-hybridization) abundance using phosphorimagery. Arbitrary phosphorimagery counts are shown for *luc* mRNA and rRNA abundance. To account for any differences arising from RNA transfer, *luc* transcript abundance from each primary transformant was normalized to its corresponding rRNA value to give relative *luc* mRNA abundance. The relative translational yield for each individual was then calculated as the ratio of Rel. LUC activity:Rel. *luc* mRNA abundance.

(a)

5'-UTR		SYN			LAT52	
Stage	40-45	>52	MP	40-45	>52	MP
Rel. *luc* activity	1.00	1.65	2.63	1.51	14.03	35.09
Rel. *luc* mRNA	1.00	1.79	2.24	3.95	4.66	2.08
Rel. TL yield	1.00	0.92	1.17	0.38	3.01	16.87

(b)

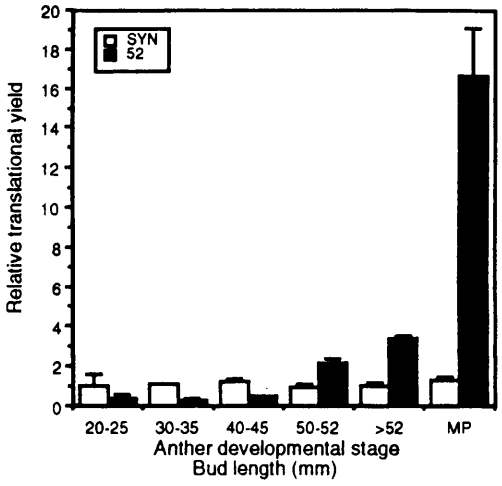


Figure 6. Analysis of translational yield of *SYN–LUC* and *52–LUC* transcripts at different stages during pollen development.
(a) Relative LUC activity, relative *luc* mRNA abundance and the relative translational yield (Rel. TL Yield) were determined as described in Figure 4 in anthers harvested from buds of 40–45 mm, greater than or equal to 52 mm and from mature pollen (MP).
(b) Histogram showing the relative translational yield of *SYN–LUC* and *52–LUC* transcripts including earlier developmental stages. Each bar represents data from two to four determinations at each stage with error bars representing the standard error.

Discussion

Previous analyses have demonstrated that multiple *cis*-regulatory sequences upstream of the transcription start site play an important role in controlling the level of transcriptional activity of the *lat*52 promoter in pollen (Eyal *et al.*, 1995; Twell *et al.*, 1990, 1991). Here, we have demonstrated that the *lat*52 5'-UTR is an important positive *cis*-regulatory element determining the level of *lat*52 gene expression in pollen. Based upon the analysis of the translational yield of transcripts in leaves and pollen of transgenic plants, we conclude that the *lat*52 5'-UTR can enhance translational yield in pollen approximately 13-fold. Consistent with the conclusion that the enhancement effect attributable to the *lat*52 5'-UTR results from an increase in the efficiency of translation, is the ability of the *lat*52 5'-UTR to enhance expression independent of transcriptional regulatory sequences, linked coding regions and the 3'-UTR. Furthermore, levels of enhancement were shown to be similar whether a strong or weak promoter was used to drive expression, which suggests that transcripts containing the *lat*52 5'-UTR, are translated more efficiently, regardless of their abundance. The transient expression experiments utilizing pollen germinated *in vitro* demonstrated the ability of the *lat*52 5'-UTR to enhance gene expression during pollen tube growth, and the transgenic analyses demonstrated enhancement during pollen maturation. Thus, the mechanisms which mediate such enhancement appear to operate both prior to, and following pollen germination. Regardless of the mechanisms involved these data demonstrate that the *lat*52 5'-UTR is an important component in controlling the high level of LAT52 protein synthesized during pollen development and pollen tube growth (Muschietti *et al.*, 1994; Villalba *et al.*, 1993).

With regard to the mechanisms which increase the translational yield of transcripts containing the *lat*52 5'-UTR in pollen several possibilities exist. Perhaps the most simple explanation is a structural one in which the low stability of secondary structure present within the *lat*52 5'-UTR would allow efficient progress of the scanning ribosomes to the initiator AUG. However, calculation of the stability of secondary structure within each of the 5'-UTRs tested showed that none had predicted ΔG values above –33 kcal mol^{-1} (see Figure 1), a value shown to have little influence on the efficiency of translation in animal systems (Koromilas *et al.*, 1992). Indeed, we have also shown that in pollen and leaves, structures with predicted ΔG values of up to –33 kcal mol^{-1} have little or no effect on the level of expression. It was not until values greater than –59 kcal mol^{-1} were reached that a significant (eight-fold reduction) effect on gene expression was observed. It is therefore unlikely that differences in the intrinsic stability of secondary structure within the *lat*52 5'-UTR, account for its ability to increase translational yield. Since the optimal initiation codon context as defined by Kozak (1986) was present at the –3 (A/G) and +4 (G) positions within all of the constructs tested, and variations in the length of the synthetic 5'-UTRs had little effect on the level of expression, it is unlikely that these factors can account for the significant increase in translational yield of messages containing the *lat*52 5'-UTR. One other possibility which could falsely indicate the presence of a translational enhancer within the *lat*52 5'-UTR would be if the synthetic controls (SYN and HB0) were translated with an artificially lower efficiency. This is highly unlikely since if this was the case one would expect that the *lat*52 5'-UTR would be translated with significantly greater efficiency in all tissue types and at all stages in pollen development, which was not observed. Taken together, our results indicate that the mechanisms mediating translational enhancement by the *lat*52 5'-UTR in pollen are not simply the result of its length, optimal translational context and lack of predicted secondary structure. A further possibility is that the rate of translational elongation is increased by the *lat*52 5'-UTR. Although we cannot rule out this possibility, the efficiency of the enhancement mediated by the *lat*52 5'-UTR was similar

for different coding regions, and it is difficult to conceive of long-range interactions between the assembled 80S ribosomes and the 5′-UTR which could affect the rate of elongation.

In the absence of obvious structural explanations our results lead us to speculate that the *lat*52 5′-UTR enhances translational yield by increasing the rate or efficiency of translational initiation. This conclusion is consistent with the generally accepted view that the rate-limiting step of translation usually occurs during initiation (Hershey, 1991). Since *lat*52 5′-UTR-containing transcripts are preferentially translated in pollen, compared with sporophytic tissues, we propose that pollen-specific factors may interact with the *lat*52 5′-UTR to increase the efficiency or access of translational initiation factors. Inherent in this proposal is the prediction that such factors would be synthesized or become activated only during the final stages of pollen maturation, since the translational advantage of *52–LUC* transcripts was apparent specifically in the final 24–36 h of development. Furthermore, the reduced translational yield of *52–LUC* transcripts at earlier stages of pollen development suggests that specific mechanisms exist to repress translation of transcripts containing the *lat*52 5′-UTR. Again, *lat*52 5′-UTR-specific factors may be involved. The modulation of translational efficiency by such positive and negative factors would provide a fine tuning mechanism to ensure that LAT52 protein only accumulated at the correct (final) stages of pollen maturation. Indeed, it may be detrimental for pollen to accumulate significant amounts of late gene products at earlier developmental stages. Although there is no direct evidence for protein factors which are able to increase the efficiency of translational initiation of specific mRNAs by the conventional scanning model, recent work indicates that specific protein factors bind to the tobacco mosaic virus translational enhancer, Ω (Gallie and Walbot, 1992). Since Ω shows a reduced requirement for the cap-binding protein eIF-4E, but is dependent on eIF-4A, such proteins may facilitate the association of eIF-4A with Ω, enabling internal initiation in a cap-independent manner (Altman *et al.*, 1990).

Current efforts to characterize cellular factors which are involved in translational initiation in plants have led to the isolation of gene families from tobacco encoding proteins closely related to eIF-4A and eIF-5A (Kuhlemeier, 1992). Both families contain members which appear to be differentially expressed among vegetative organs or in photosynthetic tissues (Chamot and Kuhlemeier, 1992; Owttrim *et al.*, 1991). Recently, an eIF-4A-related gene has been isolated which is expressed specifically in pollen and is strongly upregulated after pollen mitosis I, during which late pollen genes are activated (Brander and Kuhlemeier, 1995). Whether the expression of a pollen-specific eIF-4A is required to establish the high translational capacity required during pollen development and tube growth, or is recruited to increase the translation of specific groups of late pollen genes including *lat*52 remains to be established. We are currently investigating this possibility and whether other pollen-specific proteins specifically interact with the *lat*52 5′-UTR to mediate enhancement.

Experimental procedures

Plasmid construction

The plasmid pRTS2LUC has been described previously (Turner *et al.*, 1994). The 5′-UTR sequences HB0, HB1, HB3, and *lat*52 were derived from the plasmid templates pSV2CAT (Koromilas *et al.*, 1992) and pLAT52-7 (Twell *et al.*, 1989b), respectively. PCR was used in conjunction with specifically designed oligonucleotide primers to introduce an *Xho*I (HB0–HB3) or *Eco*RI (*lat*52) restriction site 5′ of these UTRs, and an *Nco*I site at their initiator ATGs. The HB series of PCR fragments were cloned into *Xho*I/*Nco*I-digested pRTL2GUS (Carrington and Freed, 1990) to give the pRT2HB0LUC series. The 5′-UTR from the genomic RNA of potato virus S was amplified using PCR and cloned in the same manner as for HB0 to create the plasmid pRT5′2LUC. The *lat*52 PCR fragment was digested with *Eco*RI/*Nco*I and ligated to a *luc* gene fragment digested with *Nco*I/*Bam*HI and cloned into pRTL2GUS digested with *Eco*RI/*Bam*HI to give pNBL52–26. The 5′-UTR's manipulated using PCR were sequenced and no errors were found. The construct pNBL52–5 was made by subcloning the *lat*52 promoter–5′-UTR region from pLAT52–7 as a *Kpn*I/*Nco*I fragment into a plasmid containing a *LUC–CaMV35S3′* cassette. pNBL52–42 was constructed by subcloning a *Sal*I/*Nco*I fragment from pGEM 5Zf (–) (Promega) into pRT252LUC, the fragment was then excised as an *Xho*I/*Nco*I fragment and cloned into pSL301 (Invitrogen). This synthetic polylinker 5′-UTR was excised as a *Bgl*II/*Nco*I fragment and cloned into pNBLT-1 (*lat52 promoter–TEV 5′-UTR–LUC–CaMV35S3′*), to replace the TEV 5′-UTR. pNBG52–10 was made by replacing the *LUC–CaMV35S3′* cassette of pNBL52–42 with the *GUS–NOS3′* cassette from pLAT52–7 as an *Nco*I/*Eco*RI fragment. Binary vectors were made by subcloning the *CaMV35S promoter–5′-UTR–LUC–CaMV35S3′* cassettes from pNBL52–26 and pRTS2LUC as *Hin*dIII fragments into pBIN19 (Bevan, 1984) to give pNBL52–43 and pNBL52–44 respectively, and the *lat52 promoter–5′-UTR–LUC–CaMV35S3′* cassettes from pNBL52–5 and pNBL52–42 as *Kpn*I/*Hin*dIII fragments into pBIN19, to give the plasmids pNBL52–45 and pNBL52–46, respectively.

Microprojectile bombardment

Equal amounts of supercoiled test plasmid with an appropriate reference plasmid were co-precipitated on to tungsten particles and bombarded into *Nicotiana tabacum* cv. Samsun pollen and leaves essentially as described previously (Twell *et al.*, 1989b). However, pollen was incubated for 16 h on an optimized Mes-buffered agar-solidified pollen germination medium (Tupy *et al.*, 1991) and leaves on agar-solidified MS30 medium (Twell *et al.*, 1989b) for 24 h prior to protein extraction. The reference plasmids used for *gus* and *luc* containing test plasmids bombarded into pollen were pNBL52–5 (Turner *et al.*, 1994) and pLAT52–7 (Twell *et al.*, 1990), respectively, in which the *lat*52 promoter directed LUC or GUS expression. The reference plasmids used in leaf bombardments were pRTL2GUS (Carrington and Freed, 1990) and pRTL2LUC (Turner *et al.*, 1994), in which the *CaMV35S* promoter directed GUS or LUC expression, respectively.

Plant transformation

Binary vectors were transferred directly into the disarmed *Agrobacterium tumefaciens* strain LBA4404 (pAL4404) as described previously (Twell *et al.*, 1990). Leaf discs from 6-week-old sterile shoot cultures of *Nicotiana tabacum* cv. Samsun were transformed with *A. tumefaciens* as described previously (Twell *et al.*, 1990). Kanamycin-resistant shoots were rooted on solid MS30 medium containing 50 mg l^{-1} kanamycin, transferred to soil and grown to maturity under greenhouse conditions.

Reporter gene assays

GUS and LUC activities in plant extracts were determined as described previously (Twell *et al.*, 1991). Protein concentrations in plant extracts was determined using the Bradford reagent (Bio-Rad) according to the manufacturer's instructions.

RNA analysis

Total RNA was isolated from leaves and pollen using a mini-scale preparation method essentially according to Wadsworth *et al.* (1988) starting with 50–500 mg tissue fresh weight. Ten micrograms of total RNA were separated by agarose-formaldehyde gel electrophoresis and capillary blotted on to Zetaprobe membrane using standard methods (Maniatis *et al.*, 1982). Northern blots were probed with DNA probes labelled to high specific activity with ^{32}P by oligolabelling (Feinberg and Vogelstein, 1984). The *luc* probe corresponded to a 1.1 kb fragment of the *luc* gene coding DNA. The rRNA probe was a 9.8 kb DNA fragment containing a complete rRNA repeat unit isolated from *Brassica rapa* (Da Rocha and Bertrand, 1995). Washing of filters after hybridization was carried out under stringent conditions at 65°C in 20 mM sodium phosphate buffer (pH 7.4) containing 5% sodium dodecyl sulphate (SDS), with the final wash containing 1% SDS (Church and Gilbert, 1984). Hybridization signals were quantified using a Molecular Dynamics Phosphorimager using the ImageQuant software provided by the manufacturers. Predictions of the stability of secondary structure within 5′-UTRs were calculated using the software Mulfold 2.0 (Jaeger *et al.*, 1989a, 1989b; Zuker, 1989) adapted for the Macintosh by Don Gilbert.

Acknowledgements

We are particularly grateful to Roisin Turner for providing pRT5′2LUC, Kim Cheung for PVS leaf bombardments, Gemma Allnut for help with the transgenic analysis and Ian and Neil for help with Phosphorimaging. We also thank Rouric Fuerst for recovery of Mulfold, Nahum Sonenberg for providing the HB0 control sequence, Pedro Rocha for providing the rRNA probe and Charly Brander and Cris Kuhlemeier for communicating their unpublished results and for stimulating discussion. We gratefully acknowledge The Royal Society and BBSRC for support to D.T. and G.F. as research fellows and BBSRC for financial support for this project.

References

Altman, M., Blum, S., Wilson, T.M.A. and Trachsel, H. (1990) The 5′-leader sequence of tobacco mosaic virus RNA mediates initiation-factor-4E-independent, but still initiation-factor-4A-dependent translation in yeast extracts. *Gene*, **91**, 127–129.

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

Bolle, C., Sopory, S., Lübberstedt, T., Herrmann, R.G. and Oelmüller, R. (1994) Segments encoding 5′-untranslated leaders of genes for thylakoid proteins contain *cis*-elements essential for transcription. *Plant J.* **6**, 513–523.

Brander, K.A. and Kuhlemeier, C. (1995) A pollen-specific DEAD-box protein related to translation initiation factor eIF-4A from tobacco. *Plant Mol. Biol.* **27**, 637–649.

Carrington, J.C. and Freed, D.D. (1990) Cap-independent enhancement of translation by a plant potyvirus 5′ nontranslated region. *J. Virol.* **64**, 1590–1597.

Casper, T. and Quail, P.H. (1993) Promoter and leader regions involved in the expression of the *Arabidopsis* ferredoxin A gene. *Plant J.* **3**, 161–174.

Chamot, D. and Kuhlemeier, C. (1992) Differential expression of genes encoding the hypusine-containing translation initiation factor, eIF-5A, in tobacco. *Nucl. Acids Res.* **20**, 665–669.

Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl Acad. Sci. USA*, **81**, 1991–1995.

Da Rocha, P.S.C.F. and Bertrand, H. (1995) Structure and comparative analysis of the rDNA intergenic spacer of *Brassica rapa*. *Eur. J. Biochem.* **229**, 550–557.

Eady, C., Lindsey, K. and Twell, D. (1994) Differential activation and conserved vegetative cell-specific activity of a late pollen promoter in species with bi- and tricellular pollen. *Plant J.* **5**, 543–550.

Eady, C., Lindsey, K. and Twell, D. (1995) The significance of microspore division and division symmetry for vegetative cell-specific transcription and generative cell differentiation. *Plant Cell*, **7**, 65–74.

Eyal, Y., Curie, C. and McCormick, S. (1995) Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen-expressed genes. *Plant Cell*, **7**, 373–384.

Feinberg, A.P. and Vogelstein, B. (1984) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**, 266–267.

Fennoy, S.L. and Bailey-Serres, J. (1995) Post-transcriptional regulation of gene expression in oxygen-deprived roots of maize. *Plant J.*, **7**, 287–295.

Gallie, D.R. and Walbot, V. (1992) Identification of the motifs within the tobacco mosaic virus 5′-leader responsible for enhancing translation. *Nucl. Acids Res.* **20**, 4631–4638.

Gallie, D.R. and Young, T.E. (1994) The regulation of gene expression in transformed maize aleurone and endosperm protoplasts. *Plant Physiol.* **106**, 929–939.

Hershey, J.W.B. (1991) Translational control in mammalian cells. *Ann. Rev. Biochem.* **61**, 717–755.

Jaeger, J.A., Turner, D.H. and Zuker, M. (1989a) Improved predictions of secondary structures for RNA. *Proc. Natl Acad. Sci. USA*, **86**, 7706–7710.

Jaeger, J.A., Turner, D.H. and Zuker, M. (1989b) Predicting optimal and suboptimal secondary structure for RNA. *Methods Enzymol.* **183**, 281–306.

Jefferson, R.A., Kavanagh, T.A and Bevan, M.W. (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907.

Klausner, R.D., Rouault, T.A. and Harford, J.B. (1993) Regulating the fate of mRNA: The control of cellular iron metabolism. *Cell*, **72**, 19–28.

Koromilas, A.E., Lazaris-Karatzas, A. and Sonenberg, N. (1992) mRNAs containing extensive secondary structure in their 5′ non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. *EMBO J.* **11**, 4153–4158.

Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**, 283–292.

Kuhlemeier, C. (1992) Transcriptional and post-transcriptional regulation of gene expression in plants. *Plant Mol. Biol.* **19**, 1–14.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Meerovitch, K. and Sonenberg, N. (1993) Internal initiation of picornavirus RNA translation. *Sem. Virol.* **4**, 217–227.

Michelet, B., Lukaszewicz, M., Dupriez, V. and Boutry, M. (1994) A plant plasma membrane proton-ATPase gene is regulated by development and environment and shows signs of a translation regulation. *Plant Cell*, **6**, 1375–1389.

Muschietti, J., Dircks, L., Vancanneyt, G. and McCormick, S. (1994) LAT52 protein is essential for tomato pollen development: pollen expressing antisense LAT52 RNA hydrates and germinates abnormally and cannot achieve fertilization. *Plant J.* **6**, 321–338.

Oh, S.K. and Sarnow, P. (1993) Translational initiation by internal ribosome binding in gene regulation. *Curr. Opinions Genet. Devel.* **3**, 295–300.

Ow, D.W., Wood, K.V., DeLuca, M., DeWet, J.R., Helinski, D.R. and Howell, S.H. (1986) Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. *Science*, **234**, 856–859.

Owttrim, G.W., Hofman, S. and Kuhlemeier, C. (1991) Divergent genes for translation initiation factor eIF-4A are coordinately expressed in tobacco. *Nucl. Acids Res.* **19**, 5491–5496.

Sleat, D.E. and Wilson, T.M.A. (1992) Plant virus genomes as sources of novel functions for genetic manipulations. In *Genetic Engineering with Plant Viruses* (Wilson, T.M.A. and Davies, J.W., eds). Boca Raton, FL: CRC Press Inc., pp. 55–113.

Sonenberg, N. (1994) mRNA translation: influence of the 5′ and 3′ untranslated regions. *Curr. Opinions Genet. Devel.* **4**, 310–315.

Sullivan, M.L. and Green, P.J. (1993) Post-transcriptional regulation of nuclear-encoded genes in higher plants: the roles of mRNA stability and translation. *Plant Mol. Biol.* **23**, 1091–1104.

Tupy, J., Ribhova, L. and Zarsky, V. (1991) Production of fertile tobacco pollen from microspores in suspension culture and its storage for in situ pollination. *Sex Plant Reprod.* **4**, 284–287.

Turner, R., Bate, N., Twell, D. and Foster. G. (1994) Analysis of a translational enhancer upstream from the coat protein open reading frame of potato virus S. *Arch. Virol.* **134**, 321–333.

Twell, D. (1994) The diversity and regulation of gene expression in the pathway of male gametophyte development. In *Molecular and Cellular Aspects of Plant Reproduction* (Scott, R.J. and Stead, A.D., eds). Cambridge: Cambridge University Press, pp. 83–115.

Twell, D., Wing, R., Yamaguchi, J. and McCormick, S. (1989a) Isolation and expression of an anther-specific gene from tomato. *Mol. Gen. Genet.* **217**, 240–245.

Twell, D., Klein, T.M., Fromm, M.E. and McCormick, S. (1989b) Transient expression of chimeric genes delivered into pollen by microprojectile bombardment. *Plant Physiol.* **91**, 1270–1274.

Twell, D., Yamaguchi, J. and McCormick, S. (1990) Pollen-specific gene expression in transgenic plants: Coordinate regulation of two different tomato gene promoters during microsporogenesis. *Development*, **109**, 705–713.

Twell, D., Yamaguchi, J., Wing, R.A., Ushiba, J. and McCormick, S. (1991) Promoter analysis of three genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. *Genes Devel.* **5**, 496–507.

Ursin, V.M., Yamaguchi, J. and McCormick, S. (1989) Gametophytic and sporophytic expression of anther-specific genes in developing tomato anthers. *Plant Cell*, **1**, 727–736.

Villalba, M., Batanero, E., Lopez-Otin, C., Sanchez, L.M., Monsalve, R.I., Gonzalez de la Pena, M.A., Lahoz, C. and Rodriguez, R. (1993) The amino acid sequence of Ole e I, the major allergen from olive tree (*Olea europaea*) pollen. *Eur. J. Biochem.* **216**, 863–869.

Wadsworth, G.J., Redinbaugh, M.G. and Scandalios, J.G. (1988) A procedure for the small-scale isolation of plant RNA suitable for RNA blot analysis. *Anal. Biochem.* **172**, 279–283.

Yamamoto, Y.Y., Tsuji, H. and Obokata, J. (1995) 5′-leader of a photosystem I gene in *Nicotiana sylvestris, psaDb*, contains a translational enhancer. *J. Biol. Chem.* **270**, 12 466–12 470.

Zuker, M. (1989) On finding all suboptimal foldings of an RNA molecule. *Science*, **244**, 48–52.