Sex Plant Reprod (1995) 8:242–246

ORIGINAL PAPER

S.J. Tebbutt · D.M. Lonsdale

Deletion analysis of a tobacco pollen-specific polygalacturonase promoter

Received: 13 September 1994 / Revision accepted: 17 October 1994

Abstract Deletion analysis of the promoter sequence of the tobacco gene encoding a pollen-specific polygalacturonase (Npg1) revealed several motifs related to the LAT52/56 and LAT56/59 boxes (Twell et al. 1991) and a 14 bp element with homology to the maize polygalacturonase promoter (Allen and Lonsdale 1993), termed the PG box. Deletion analysis also revealed a complex pattern of domains which increase or reduce promoter activity. In particular, deleton of the LAT52/56 and LAT56/59 box motifs resulted in significant reductions in expression giving weight to the idea that such sequences function as transcriptional positive regulatory elements. Deletion of sequences immediately 3′ to the LAT52/56 and LAT56/59 motifs fully restore promoter function, suggesting the presence of suppressor elements associated with these positive regulatory elements. Analysis of the deletion series allowed the promoter to be divided into four domains: a modulation domain (–744 to –362), a basic promoter (–362), a core promoter (–267) and a minimal promoter (–182). Two sequence elements, Eh-1 (22 bp) and Eh-2 (28 bp), were identified between –362 and –267. Deletion of these two unrelated sequence elements reduces the activity of the basic promoter to that of the core promoter, a level of activity which is retained by the minimal promoter. A background level of activity is reached when the promoter is reduced to –86. Therefore, the sequence between –182 and –86, which contains two LAT52/56 motifs and the PG box, contains sufficient information to direct efficient gene expression.

Key words Tobacco · Polygalacturonase · Pollen · Promoter analysis · Biolistic · Particle gun

S.J. Tebutt
Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK

D.M. Lonsdale (✉)
Department of Molecular Genetics, Cambridge Laboratory, John Innes Centre, Colney Lane, Norwich NR4 7UJ, UK;
FAX no: (0)603 505725; E-mail: Lonsdale@bbsrc.ac.uk

Introduction

The identification of specific DNA sequences and proteins required to establish the spatial and temporal patterns of gene expression is important for our understanding of gametophytic development within the anthers of flowering plants. The promoter sequences of genes can be readily fused to reporter genes such as β-glucuronidase (GUS), chloramphenicol acetyl transferase (CAT) or luciferase (LUC) and transferred into the genome of plants to confirm and study gene expression patterns. Analysis of the expression patterns of promoter deletions provides information on domains which determine tissue, cell specific, and temporal expression.

The most comprehensive promoter analysis of pollen-specific genes has involved the tomato LAT52, LAT56 and LAT59 genes (Twell et al. 1990, 1991). Various cis-regulatory elements involved in promoter strength and specificity have been identified using transgenic methods. These include transgenic tobacco plant experiments as well as transient expression systems in which the chimeric DNA constructs were delivered into tobacco pollen by microprojectile bombardment (Twell et al. 1989). These experiments have identified two particularly important regulatory elements required for late pollen gene expression: the 52/56 box (TGTGGTTATATA) and the 56/59 box (GAAATTGTGA), which are shared among two of the three LAT gene promoters. The 52/56 box seems to have a quantitative role in which basal LAT52 and LAT56 promoter activity in pollen is coordinately enhanced by a transcription factor or factors that recognise this sequence. The 52/56 box resembles the *RbcS* GT-1 box which binds the nuclear factor GT-1 (Green et al. 1988) and is characteristic of many promoters, and thus probably does not serve as a pollen-specific sequence. The conserved 56/59 box may also activate basal promoter activity and/or it may have a more determinative role in turning on expression of the LAT56 and LAT59 genes in pollen (Twell et al. 1991). Similar motifs have been found in the upstream sequences of other anther-specific genes, including Bp4, Bp19 and Bp10

from *Brassica* (Albani et al. 1992; Fabijanski et al. 1992). Zm13 from maize (Hanson et al. 1989) and CHI-P_{A2} from petunia (van Tunen et al. 1989). In addition, other short sequence homologies are discernable. These may represent the binding domains of regulatory proteins, though such a role is questionable until a controlling function is demonstrated.

We have previously established that the −744 bp to +85 bp sequence region of the tobacco pollen-specific polygalacturonase is sufficient to drive the spatial and temporal expression of the Npg1 gene in tobacco microgametophytic tissue (Tebbutt et al. 1994). In this paper we describe the analysis of this region more closely in an attempt to define a minimal promoter and to investigate the role of the LAT52/56 box and the LAT56/59 box in the regulation of expression of the tobacco pollen polygalacturonase gene.

Materials and methods

Plasmids and construction of the promoter deletion series

The plasmid pG27.11, which contains 744 bp of the polygalacturonase promoter, the 81 bp 5′-UTL fused in frame at +4 to GUS::NosTer (Tebbutt et al. 1994), was used to derive a 5′-deletion series using the exonuclease III/SI protocol of Henikoff (1987). Suitable recombinants were selected and sequenced.

In addition to the deletion series plasmids, two reference plasmids were utilized: (1) *pTac25-Nco::LUC* contains 594 bp of the tobacco actin promoter, a 130 bp 5′-untranslated leader (UTL), which includes the 490 bp leader intron, fused to the firefly luciferase (LUC) coding region and terminated by NosTer (Thangavelu et al. 1993; Lonsdale, unpublished); (2) *pPL-A173::GUS* contains 937 bp of the tobacco pectate lyase promoter (Rogers et al. 1992) with a 212 bp 5′-UTL fused to the GUS coding sequence in pBI201 (Jefferson 1987).

Plasmid DNA was isolated from 0.4 l overnight cultures according to the maxi-prep procedure described by Lonsdale et al. (1988). DNA was purified either by Qiagen column chromatography or by banding twice in CsCl-ethidium bromide gradients. Ethidium bromide was removed by extracting 4 times with CsCl-saturated isopropanol following which the DNA was precipitated and resuspended in water. Based on E^{260}, the GUS and LUC plasmids were mixed in a 7:3 ratio to a concentration of 1 µg/µl.

Microprojectile bombardment

The method of Twell et al. (1989) was used as described by Lonsdale et al. (1990). The particle accelerator was made by Shearline Precision Engineering, Ely, Cambridge, UK. Stopper plates and macroprojectiles were supplied by Shearline Precision Engineering Ltd. Eley, Birmingham, UK, manufactured 0.22 cartridges to our specification.

Tobacco pollen

Nicotiana tabacum (var. Samsun) pollen was collected and stored at −70° C. Pooled pollen samples were thawed and dispersed to 20 mg/ml in MS medium (Murashige and Skooge 1962), 3% w/v sucrose, 200 µg/ml carbenicillin and 0.5 µg/ml Fungizone (GibcoBRL). A 3 cm^2 piece of sterile Whatmans No. 1 paper and a 2 cm^2 sterile nylon-membrane (Genescreen Plus™) were placed in 9 cm diameter petri dishes containing MS agar medium, 3% w/v sucrose and 200 µg/ml carbenicillin. To produce a target area of 1.0 to 1.5 cm^2, 0.5 ml (10 mg) pollen was pipetted onto the sterile nylon-membrane support. Following bombardment of the pollen samples, the plates were incubated from 16 h at 25° C.

GUS and LUC assays

Pollen was harvested into an agate mortar and ground in 500 µl of 100 mM potassium phosphate buffer pH 7.0, 1 mM DTT, 1 mg/ml BSA. The extract was clarified by micro-centrifugation at 13000 rpm for 5 min at 4° C, immediately after which the luciferase activity of each extract was assayed according to Ow et al. (1986). Following injection of the luciferin activity was measured in a Berthold luminometer (LUMAT LB9501). Light emission was measured for 10 s.

Glucuronidase activity was measured fluorometrically as detailed in Jefferson (1987).

All assays were performed in triplicate. Means were used to determine the ratio of GUS to LUC activity for each sample.

Quantification

Aliquots of the GUS/LUC DNA mixtures were digested with suitable restriction enzymes to generate fragments of both the GUS gene fusion and the LUC gene fusion. The restriction fragments were separated by agarose gel electrophoresis and the relative ratios of the two plasmids determined by densitometry of proximate restriction fragments using a UVP Gel Analysis System SW2000. Variations from the E^{260}-determined 7:3 ratio were used to correct the calculated activities of the constructs. This correction overcomes the variation which occurs as a result of low molecular weight DNA contamination and/or residual ethidium bromide contamination of the plasmid preparations.

The GUS reference plasmid pPL-A173::GUS was used in all experiments. The mean GUS activity of this plasmid relative to pTAC25-Nco::LUC was set to 100.

A minimum of five replicate bombardments for each construct were performed. The standard deviation was calculated according to Snedecor (1967), using the appropriate low number correction factor and omitting the upper and lower values. Probability values were determined by applying Student's *t*-distribution to the data.

Results and discussion

Sequence analysis and homologies

The sequence of the promoter including the 5′-UTL is shown in Fig. 1. Comparison of this sequence to the EMBL database (v37) revealed no significant homologies. A putative TATAA box, starting at −34 bp, can be identified based on homology to the reported consensus (Joshi 1987). Analysis of the Npg1 promoter sequence also revealed the presence of an imperfect palindrome, −619 bp to −588 bp relative to the transcription start site (motif 1, Fig. 1).

More detailed comparisons to reported anther and pollen-specific promoter sequences revealed a number of short related sequence motifs. One of the longer homologies identified was with the maize polygalacturonase promoter (Allen and Lonsdale 1993). It is located −149 bp to −127 bp in the tobacco promoter (motif 6, Fig. 1) and between −117 bp and −95 bp in the maize promoter. Comparison of the sequences reveals a 9 bp conserved core and 3 bp flanks, separated by a 4 bp variable region: $AAC^{T}/_{C}{}^{C}/_{T}T^{T}/_{C}AATTAGTAA^{A}/_{T}AC^{A}/_{C}AAG$ (mismatches are indicated: tobacco, superscript; maize, subscript).

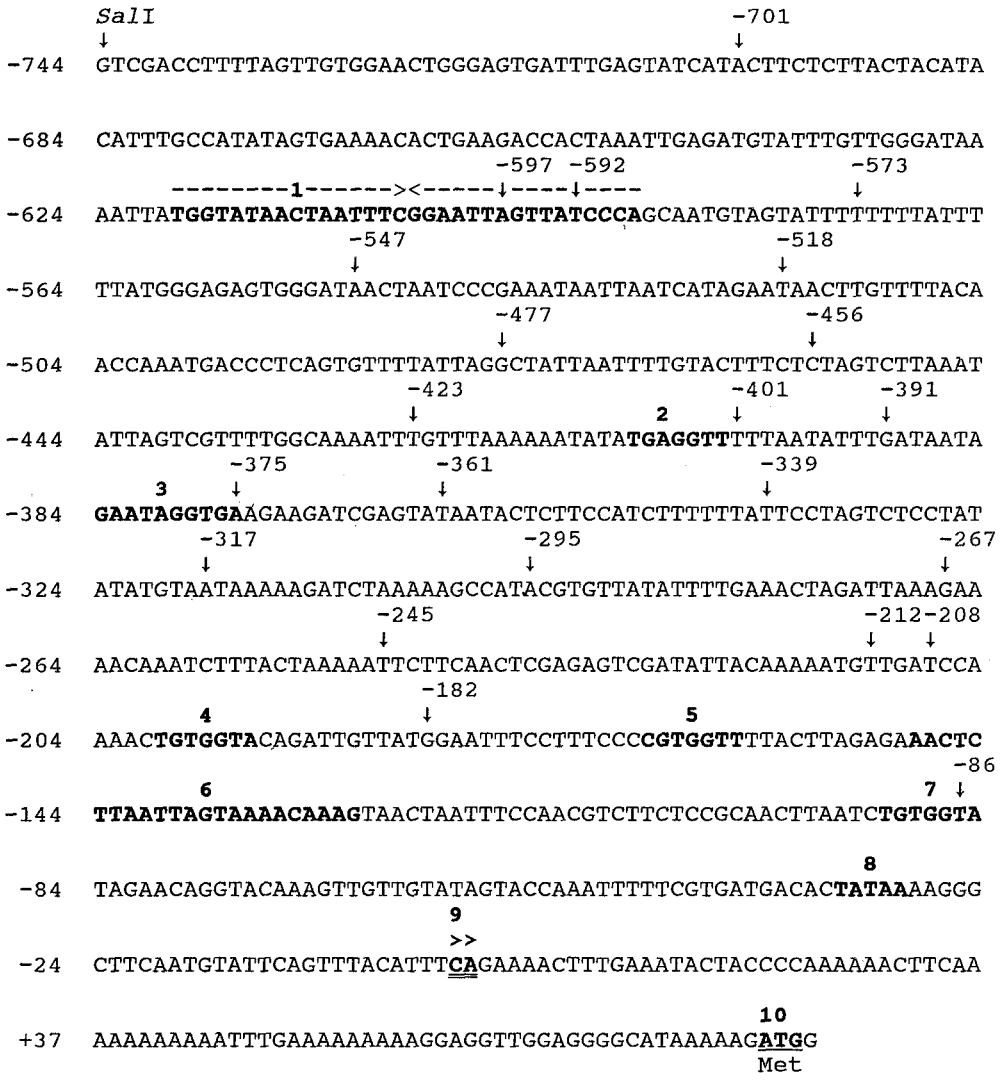

Fig. 1 Sequence of 5′ region of Npg1, showing the 744 bp promoter and the 81 bp UTL. The 5′-nucleotide of each promoter deletion and its length are shown. *1* The imperfect palindromic element; *2, 4, 5, 7* sequences which resemble the LAT52/56 box core motif; *3* a putative pollen box similar to the LAT56/59 box; *6* the PG box; *8* the putative TATAA; *9* transcriptional start site; *10* translational start site

We called this the 'PG box', since it is not discernable in other anther or pollen-specific promoters.

Sequence motifs homologous to the cis-acting sequences in the LAT52, 56 and 59 promoters identified by Twell et al. (1991) can also be identified in the Npg1 promoter. These have been shown by mutational scanning to be required for expression of the LAT genes in tobacco (Twell et al. 1991). The LAT56/59 element has been identified in many other pollen-specific promoters including those of the maize Zm13 (Hamilton et al. 1989) and the petunia CHI-A P_{A2} genes (van Tunen et al. 1989). In tobacco Npg1, GAATAGGTGA (motif 3, Fig. 1) is the only sequence displaying similarity to the LAT56/59 box (GAAATTGTGA), from which it differ at three positions (underlined). Alterations of the LAT56/59 box at these positions have only a marginal effect, if any, on expression (Twell et al. 1991).

Motifs 2, 4, 5 and 7 (Fig. 1) are related to the core motif of the LAT52/56 box (TGTGGTT). This motif is identical to the *RbcS* GT-1 box which binds the nuclear factor GT-1 (Green et al. 1988) and probably binds a general transcriptional enhancement factor.

Effects of deleting motifs

Deletion of the palindrome sequence (motif 1, Δ-597, Δ-592; Fig. 1, 2) has little if any effect on promoter activity. Deletion of the LAT52/56 element (motif 2, Δ-401) results in a significant activity reduction that is reversed by deletion of the next 10 bp. This sequence, TTTAATATTT, clearly suppresses the activity of the promoter. Deletion of the LAT56/59 element (motif 3, Δ-375) again results in a significant reduction of activity. Deletion of the next 14 bp (AAGAAGATCGAGTA) leads to a significant activity enhancement ($P \geq 0.01$) that restores the activity of the full length promoter. There is no sequence similarity between this and the other sequence which is immediately 3′ to the LAT52/56 box (motif 2) which also suppress promoter activity. Nor is there any similarity to the other suppressor domain (–518 to –478 and –456 to –423). The LAT56/59 element is obviously not required for promoter function as its deletion along with the sequences immediately 3′ does not affect the overall function of the promoter.

Deletions between the LAT56/59 box (motif 3) and the LAT52/56 box (motif 4) reveal two domains which

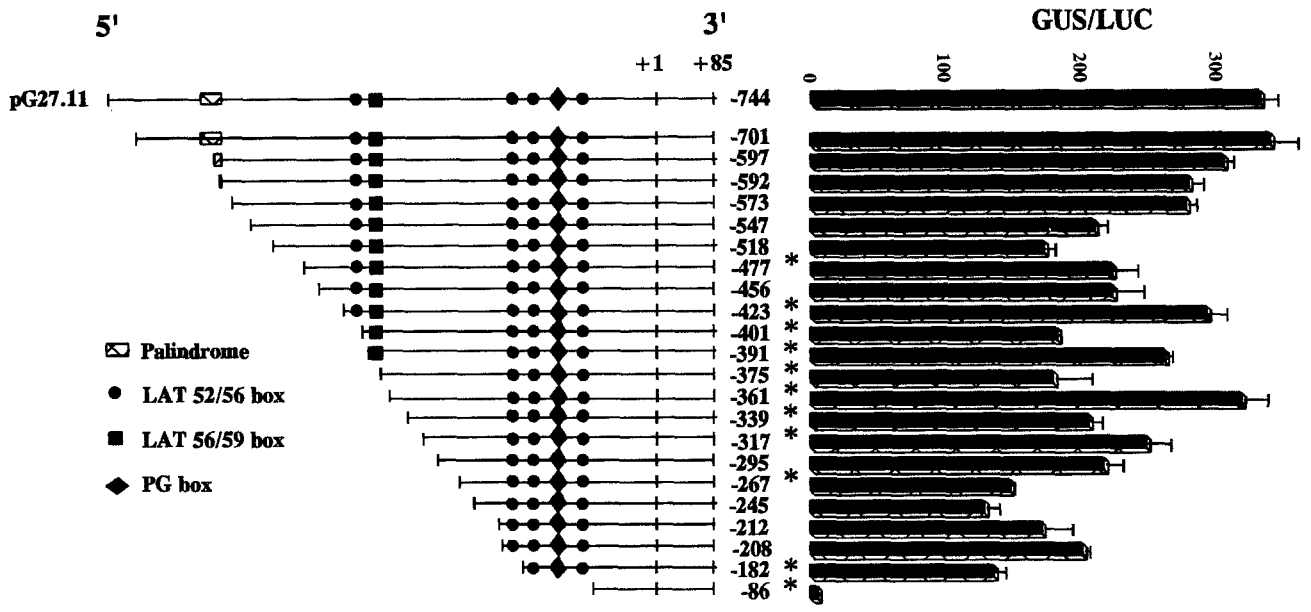

Fig. 2 Relative activities of the Npg1-promoter deletions compared to the full-length pG27.11 construct. Activity is represented by the GUS/LUC ratio, standardised against pPL-A173::GUS/pTAC25-Nco::LUC with relative activity set at 100 units. Positions of the imperfect palindrome, LAT52/56 boxes, LAT56/59 box and polygalacturonase (*PG*) box are indicated. Error bars are half the value of the standard deviation. * The difference in relative activities of adjacent deletions is statistically significant ($P \geq 0.05$)

significantly enhance expression, −361 to −340 and −295 and −268. Deletion of these sequences leads to a significant reduction in promoter activity. Deletion of the LAT52/56 element (motif 4, −208 to −182) results in a similar decrease to the deletion of motif 2, clearly implicating the LAT52/56 elements as positive regulatory elements. Finally, deletion of the two LAT52/56 boxes (motifs 2, 4) and the PG box (motifs 5, 6, 7) reduces the activity of the promoter to background levels.

Analysis of the promoter deletions

The reproducibility of the analysis and the significance of the observed differences can be analysed in several ways. Comparison of the deletion pairs Δ-744/Δ-701, Δ-592/Δ-573, Δ-477/Δ-456, Δ-317/Δ-295 and Δ-212/Δ-208 show that the activity differences between these pairs are insignificant (±σ). This is to be expected with the removal of a sequence that plays no significnt role in promoter function. Where activity differences between adjacent deletions are statistically significant (±2σ), the deletion or inactivation of a sequence element involved in promoter function is assumed. Using these terms of reference in assessing the relative activities of the deletions (Fig. 2), the promoter can be divided into a number of domains (Fig. 3).

Deletions within the *modulation domain* (−744 to −362) can significantly affect the overall activity of the promoter, although deletion of this entire region restores the activity of the full length promoter.

The *basic promoter* (−361 to −87) has the same activity as the full length promoter. Deletion of the sequence between −361 and −340 results in a 36% drop in activity. This is statistically significant ($P \geq 0.01$) and suggests that this sequence is a positive regulatory element (Eh-1). The next three deletions, Δ-339, Δ-317 and Δ-295, have little effect.

The next deletion, Δ-267, exhibits a significant reduction in acitivity ($P \geq 0.01$), suggesting another positive regulatory element (Eh-2). Deletion of the next 22 bp (Δ-245) has no effect. This level of activity is also reflected by Δ-182. Therefore, Δ-267 can be considered as a *core promoter* whose activity is up-regulated by the two upstream positive regulatory elements, Eh-1 and Eh-2 (Fig. 3). Deletion of the sequence between −245 and −212 (Δ-212) leads to a slight increase in promoter activity; it appears not to be significant because the activity of Δ-245 and Δ-182 have similar activities ±2σ.

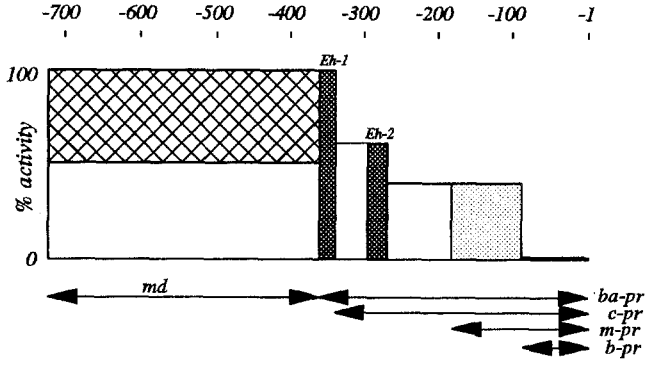

Fig. 3 Schematic of the Npg1 promoter. *md* Modulation domain (*hatched area* indicates range of observed activities), *ba-pr* basic promoter, *c-pr* core promoter, *m-pr* minimal promoter, *b-pr* basal promoter, *Eh-1, Eh-2* enhancer elements

Δ-182 represents the *minimal promoter* as the data clearly show that this region has the same activity as Δ-267. In this analysis Δ-182 is the smallest defined region which still functions as an effective promoter. It retains approximately 50% of the activity of the full length promoter and the core promoter. It can be concluded that the 96 bp domain between −183 and −85 contains sufficient information to direct pollen gene expression.

This analysis has identified several specific regions of the Npg1 promoter that appear to be involved in regulating the expression of the Npg1 gene. Experiments are now in progress to map the protein binding domains.

Acknowledgements We wish to thank Alison Harvey and Marie Bradley for technical assistance. S.J.T. was supported by an AFRC studentship.

References

Albani D, Sardana R, Robert LS, Altosaar I, Arnison PG, Fabijanski SF (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 tobacco plants. Plant J 2:331–342

Allen RL, Lonsdale DM (1993) Molecular characterisation of one of the maize polygalacturonase gene family members, which are expressed during late pollen development. Plant J 3: 261–271

Fabijanski SF, Albani D, Robert LS, Arnison PG (1992) Characterization of genes expressed during the development of *Brassica napus* pollen. In Vitro Cell Dev Biol 280:46–52

Green PG, Yong M-H, Cuozzo M, Kano-Murakami Y, Silverstein P, 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

Hamilton DA, Bashe DM, Stinson JR, Mascarenhas JP (1989) Characterization of a pollen-specific genomic clone from maize. Sex Plant Reprod 2:208–212

Hanson DD, Hamilton DA, Travis JL, Bashe DM, Mascarenhas JP (1989) Characterization of a pollen-specific cDNA clone from *Zea mays* and its expression. Plant Cell 1:173–179

Henikoff S (1987) Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol 155:156–165

Jefferson RA (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol Biol Rep 5:387–405

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

Lonsdale DM, Hodge TP, Stoehr PJ (1988) Analysis of the genome structure of plant mitochondria. In: Weissbach A, Weissbach H (eds) Methods for plant molecular biology. Academic Press, New York, pp 275–292

Lonsdale DM, Onde S, Cuming A (1990) Transient expression of exogenous DNA in intact, viable wheat embryos following particle bombardment. J Exp Bot 41:1161–1165

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant 15:473–497

Ow DW, Wood KV, DeLuca M, deWet JR, Helinski DR, Howell SH (1986) Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. Science 234:856–859

Rogers HJ, Harvey A, Lonsdale DM (1992) Isolation and characterization of a tobacco gene with homology to pectate lyase which is specifically expressed during microsporogenesis. Plant Mol Biol 20:493–502

Snedecor GW (1967) Statistical methods applied to experiments in agriculture and biology. Iowa State University Press, Ames, Iowa

Tebbutt SJ, Rogers HJ, Lonsdale DM (1994) Characterization of a tobacco gene encoding a pollen-specific polygalacturonase. Plant Mol Biol 25:283–297

Thangavelu M, Belostotsky D, Bevan MW, Flavell RB, Rogers HJ, Lonsdale DM (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

Twell D, Klein TM, Fromm ME, McCormick S (1989) Transient expression of chimeric genes delivered into pollen by microprojectile bombardment. Plant Physiol 91:1270–1274

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

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

van Tunen AJ, Hartman SA, Mur LA, Mol JNM (1989) Regulation of chalcone flavanone isomerase (CHI) gene expression in *Petunia hybrida*: the use of alternative promoters in corolla, anthers and pollen. Plant Mol Biol 12:539–551