

Partial characterization of the *Nicotiana tabacum* actin gene family: Evidence for pollen-specific expression of one of the gene family members

Madan Thangavelu⁵, Dmitry Belostotsky⁴, Michael W. Bevan¹, Richard B. Flavell², Hilary J. Rogers³, David M. Lonsdale¹

¹ Department of Molecular Genetics, Cambridge Laboratory and

² John Innes Institute, JI Centre for Plant Science Research, Colney Lane, Norwich NR4 7UJ, UK

³ Laboratory of the Government Chemist, Queens Road, Teddington, Middlesex, UK

⁴ Department of Genetics, University of Georgia, Athens, GA 30602, USA

⁵ Department of Botany, Wye College, Nr Ashford, Kent, UK

Received: 21 August 1992/Accepted: 23 February 1993

Abstract. The actin gene family of *Nicotiana tabacum* has been partially characterised by Southern hybridisation and by isolating λ EMBL4 recombinants from a genomic library having homology to the soybean actin gene, *Sac3*. The number of actin genes with homology to *Sac3* is estimated at between 20 to 30, based on Southern hybridisation and library screening, though the total gene family may be larger. Twenty-four recombinant λ clones were isolated, 18 had unique restriction profiles and from these, 2 clones, *Tac9* and *Tac25*, were selected for further study. The region of *Tac25* hybridizing to *Sac3* was sequenced and shown to contain an open reading frame (ORF) with homology to actin. Partial sequencing of *Tac9* revealed a sequence with homology to the third exon of *Tac25* and *Sac3*. The two tobacco actin sequences were compared to other reported actin gene sequences; *Tac25* was closely related to the allelic potato actins, *Pac58* and *Pac85*, while *Tac9* was more related to *Pac79* than to other plant actins. Northern hybridisation analysis showed that while *Tac9* detected actin transcripts in RNA from root, leaf, stigma and pollen, *Tac25* transcripts were only detected in pollen RNA.

Key words: Tobacco – *Nicotiana tabacum* – Actin – Pollen – Gene family

Introduction

Actin is found in all eukaryotic cells and is the principal component of the microfilaments of the cytoskeleton. Cytological and biochemical studies in a range of organisms also suggest that actin is involved in a number of indispensable cellular processes such as cell division, cytoplasmic streaming and organelle movement (Stai-

ger and Schliwa 1987). From the limited number of plant actin genes that have been either fully or partially sequenced, it is evident that there is considerably more intra- and interspecific variation at both the DNA and amino acid sequence levels than in the non-plant actins (Meagher 1991). This variation is believed to be due either to the fact that the vascular plant actins arose from an ancient class of genes, which may not have shared a common ancestor, or to be the result of the co-evolution of specific genes with specialised cell types, which suggests that certain actins may have evolved specialised functions (Meagher 1991).

In the metazoa there are two functionally distinguishable isoforms: the muscle-specific and general cytoplasmic actins. Both tissue-specific and developmental regulation of actin genes have been reported for a number of these isoforms (Melloul et al. 1984; Fyrberg et al. 1983). There is no evidence for such a clear division of function in the plant kingdom. Expression studies, in soybean and rice, have shown that some of the members of these actin gene families are differentially expressed (Hightower and Meagher 1985; McElroy et al. 1990; McLean et al. 1990) though there is no evidence to show that individual members of these gene families are subject to strict cell-specific expression.

We report here the preliminary analysis of the tobacco actin gene family. The entire coding region of one of the gene family members, *Tac25*, has been completely sequenced while another, *Tac9*, has been partially sequenced. *Tac25* appears to be an ortholog of the allelic potato actin genes *Pac58* and *Pac85* (Drouin and Dover 1990). Northern analysis suggests that the expression of *Tac25* may be limited to mature pollen while *Tac9* hybridises to RNA isolated from a number of tissues.

Materials and methods

Construction of the genomic library. Green leaves of *Nicotiana tabacum* var. Havana were ground to a powder in liquid nitrogen before being dispersed in 0.1 M TRIS-

Communicated by D.M. Lonsdale

EMBL Accession numbers: *Tac9*, X69885; *Tac25*, X63603

Correspondence to: D.M. Lonsdale

HCl pH 8.0, 0.1 M NaCl, 50 mM ethylenediaminetetraacetic acid (EDTA), 2% sodium dodecyl sulphate (SDS), 200 µg/ml proteinase K. The lysate was heated at 55°C for 1 h before being extracted twice with phenol/chloroform (1:1 (v/v)). The DNA was isolated by spooling from an isopropanol precipitation and purified by banding twice in CsCl/ethidium bromide gradients. The purified DNA was partially digested with *Eco*RI. Fragments with a size of 15–22 kb were purified by sucrose gradient centrifugation and ligated into similarly purified arms of λ EMBL4 before packaging. A library of 5×10^7 recombinants was obtained.

DNA sequencing and sequence analysis. pEMBL single-stranded DNA templates were sequenced by the dideoxy chain-termination method (Sanger et al. 1977). The sequence was compiled and analysed using the Staden (1986) and Wisconsin programs (Devereux et al. 1984).

Results and discussion

Estimation of the tobacco actin gene family size

*Hind*III- and *Eco*RI-digested DNAs from wheat, tobacco, petunia, potato and soybean were probed with a 640 bp *Bg*II-*Eco*RI coding fragment of the soybean actin gene clone, Sac3 (Shah et al. 1982). In the dicot species, the Sac3 probe hybridised to a relatively large number of bands with variable intensity (Fig. 1); in wheat only a few discrete bands were detected. There are several possible reasons for this finding: firstly, detection sensitivity is reduced as the wheat genome size, at approximately 34 pg, is 7 to 10 times the size of the genomes of the other species (Bennett and Smith 1976) and secondly, the actin gene sequences of wheat are sufficiently divergent to effectively reduce hybridisation to the soybean actin gene probe.

In petunia, previous studies (Baird and Meagher 1987; McLean et al. 1990) have shown that actin is encoded by a superfamily of genes, with as many as 200 members. The observed hybridisation pattern is indeed complex, consistent with the gene family being large (Fig. 1). The potato hybridisation pattern is similar to petunia in its complexity, suggesting a family of equivalent size. In soybean digests there are fewer fragments but these are significantly more intense, suggestive of more copies per band and again indicating a large number of genes. By comparison to petunia, potato or soybean, tobacco has a relatively simple hybridisation pattern, indicative of a relatively small gene family. Approximately 15 to 20 restriction bands can be distinguished with *Hind*III- or *Eco*RI-digested DNA. Based on the copy number standards (Fig. 1), we have estimated that there are at least 25–30 genes, assuming one gene per band and 100% homology to the Sac3 actin probe.

To characterize the gene family further, a *N. tabacum* (var. Havana) genomic library was constructed. Genomic DNA was partially digested with *Eco*RI and fragments with sizes between 15 and 22 kb cloned into λ EMBL4. Approximately 4.5×10^5 plaques were screened using the

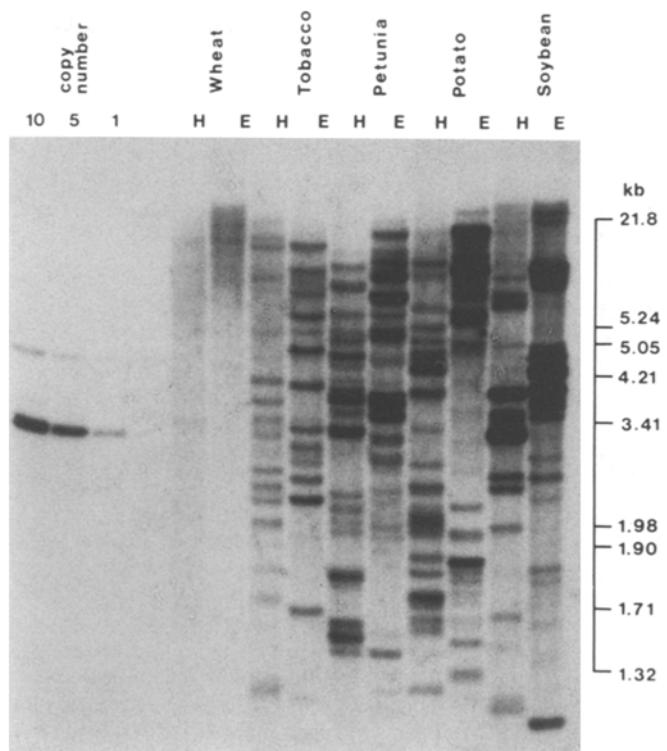


Fig. 1. Southern blot analysis of plant DNAs using the soybean actin Sac3 probe. Approximately 10 µg DNA was digested with *Hind*III (H) or *Eco*RI (E) and electrophoresed on a 0.8% agarose gel. A 10, 5 and 1 copy (\approx 3.6 pg) reconstruction using a *Hind*III digest of pSac3 are in lanes 1, 2 and 3 respectively. The restriction fragments were transferred to nitrocellulose and probed with the 640 bp *Bg*II-*Eco*RI coding sequence fragment of Sac3 (Shah et al. 1982). Hybridisation was carried out in $6 \times$ SSC, $5 \times$ Denhardt's, 0.2% SDS and 100 µg/ml herring sperm DNA at 50°C. The filter was washed at 50°C in $3 \times$ SSC prior to autoradiography

640 bp *Bg*II-*Eco*RI Sac3 actin probe. Using the same hybridisation and washing conditions as for the Southern analysis (Fig. 1), approximately 100 positive hybridising plaques were obtained. Of these, 24 phage plaques were purified and the phage DNA subjected to restriction analysis; 18/24 recombinant phages had unique restriction profiles and seven of these were mapped in detail (Fig. 2). Compared to the Southern analysis, the number of hybridising plaques (assuming 75% are unique) equates to about 30 genes per haploid genome ($1C = 3.6 \times 10^9$ bp; Bennett and Smith 1976); this figure agrees well with the estimate based on Southern analysis. However, as with the wheat genomic DNA, to which Sac3 did not hybridise particularly well, there may be tobacco actin genes to which Sac3 does not hybridise, and therefore the size of the *N. tabacum* actin gene family estimated at 25 to 30 genes should be considered a minimum.

Sequences related to Tac9 and Tac25 in the tobacco genome

The 5.6 kb *Eco*RI-*Sst*I fragment of λ Tac9 and the 2.6 kb *Bam*HI fragment of λ Tac25 (see Fig. 2) were used to

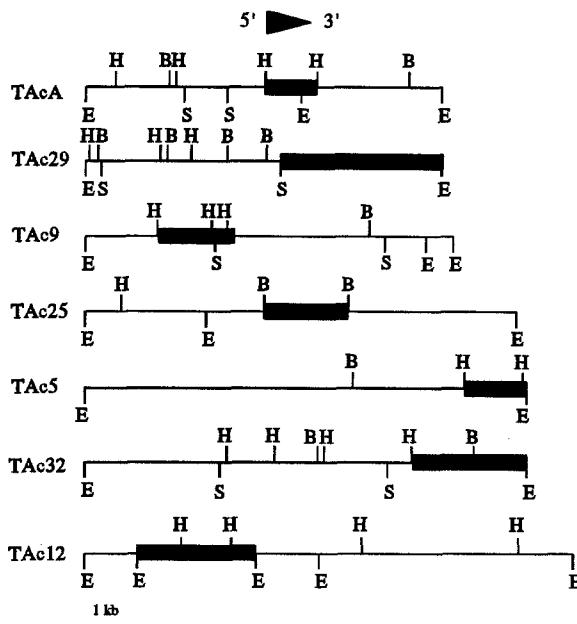


Fig. 2. Restriction enzyme maps of different recombinant phage carrying *Nicotiana tabacum* actin genes. The phage DNAs were mapped with *Bam*HI (B), *Eco*RI (E), *Hind*III (H) and *Sst*I (S). The maps were orientated relative to the orientation of the actin gene which was determined using 5' and 3' specific sub-clones of pSac3 (Shah et al. 1982). The solid box indicates the location of the region hybridizing to the entire Sac3 probe

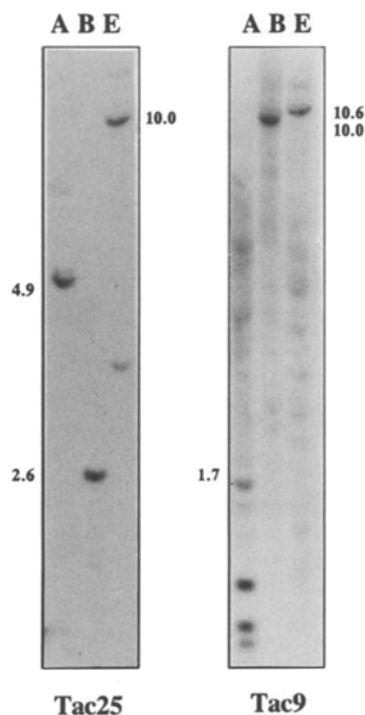


Fig. 3. Southern blot analysis of *N. tabacum* (var. Havana) probed with the 2.6 kb *Bam*HI fragment of Tac25 and 5.6 kb *Eco*RI-*Sst*I fragment of Tac9 (Fig. 2). Ten micrograms of DNA was digested with *Afl*III (A), *Bam*HI (B) and *Eco*RI (E). Hybridization was in 6 × SSC, 10 × Denhardt's, 1.0% SDS and 100 µg/ml herring sperm DNA at 65°C. The filter was finally washed in 0.05% SSC at 65°C prior to autoradiography. The sizes of hybridising bands are given alongside the gel in kb

probe tobacco genomic DNA digested with *Afl*III, *Bam*HI and *Eco*RI (Fig. 3). Both Tac25 and Tac9 hybridised to *Bam*HI and *Eco*RI fragments of the same size as those identified in the λ clones (Fig. 2). The 2.6 kb *Bam*HI fragment of Tac25 detected only the homologous genomic fragment, whereas in the *Eco*RI and *Afl*III digests one or two additional fainter bands were evident, suggesting one or two closely related sequences. This was in contrast to the digests hybridised with the Tac9 probe where a background of additional bands was detected. This clearly demonstrates that Tac9 is much more closely related to other members of the tobacco actin gene family than is Tac25.

Sequence analysis of Tac9 and Tac25

The 2.6 kb *Bam*HI fragment of Tac25 was sequenced entirely (Fig. 4) while 820 bp of Tac9 was sequenced (data not shown). The sequences were compared to the reported sequences of the soybean and maize actins (Shah et al. 1982, 1983). The sequence of Tac9 corresponded to the entire third exon of the gene. The sequence of the *Bam*HI fragment of Tac25 contained an open reading frame (ORF) encoding a polypeptide of 377 amino acids with a molecular mass of 41 739 daltons. The ORF was disrupted by three intervening sequences of lengths 570 bp (intron I), 155 bp (intron II) and 296 bp (intron III). The first and third intron lengths are relatively large compared to the lengths of other plant actin introns, as a result of imperfect sequence duplications (Fig. 4). There is a direct repeat, of approximately 158 bp, in the first intron and an inverted repeat, of approximately 59 bp, in the third intron. The positions of the introns are perfectly conserved in relation to the groups of codons in which they are inserted. The first intron is between codons 20 (Lys) and 21 (Ala); the second intron disrupts the glycine codon 152 between the first and second nucleotide, and the third intron lies between codon 356 (Gln) and 357 (Met).

In addition to the introns that disrupt the coding sequence, all reported actin gene sequences have a large intron in the untranslated leader with the 3' splice junction located approximately 10 bp from the ATG (Shah et al. 1983; Drouin and Dover 1990); McElroy et al. 1990; Meagher 1991). Though we have not established the presence of a leader intron directly, a PCR product of the *Bam*HI fragment of Tac25 (positions 3–226, see Fig. 4), which is the sequence located immediately 5' to the ATG, fails to hybridise to RNA on Northern blots (data not shown) while the entire *Bam*HI fragment hybridises to a 1.5 kb transcript (Fig. 5). The inability of the DNA immediately 5' to the ATG to detect the 1.5 kb transcript is suggestive of the presence of a leader intron. A comparison of the third exon nucleotide sequences of Tac9 and Tac25 showed that they were 81.5% homologous while amino acid homology was 92.7%. A comparison of the Tac25 5' and 3' non-coding regions, and the intron sequences to the corresponding sequences of other plant actins and to the EMBL database (release 33) showed no significant homologies.

Fig. 4. Sequence of the *Bam*HI fragment of Tac25. The non-coding sequences are shown in *lower case*. The nucleotides comprising the polymerase chain reaction (PCR) primers used to amplify the 5'

non-transcribed region are *overlined*. The direct repeat in intron 1 is *underlined and double-underlined*. The inverted repeats in intron 3 are *italicised and underlined*.

Expression of Tac9 and Tac25

To investigate whether there was any specificity associated with the expression of these genes, mRNA was isolated from *N. tabacum* pollen, stigma, leaf and root and *Brassica napus* pollen and leaf and probed with restriction fragments from within exon III of Tac9 (Fig. 5A) and Tac25 (Fig. 5B). The results show quite clearly that whereas the Tac9 probe hybridises to RNAs from all tissues, with increased hybridisation to pollen and root, Tac25 hybridises only to pollen mRNA; no hybridisation to the mRNA of the other tissues could be detected even on long over-exposures, and neither frag-

ment hybridises significantly to pollen and leaf RNA isolated from *B. napus*.

Comparison of Tac9 and Tac25 to other plant actins

Plant actin sequences recovered from the EMBL database (release 33) were compared to nucleotide and amino acid sequences of Tac9 and Tac25 using the program PILEUP (Devereux et al. 1984). Exon 3 of Tac9 was compared to all the actins and was most closely related to the potato actin, Pac79, showing 94.2% and 93.3% homology between the amino acid and nucleotide

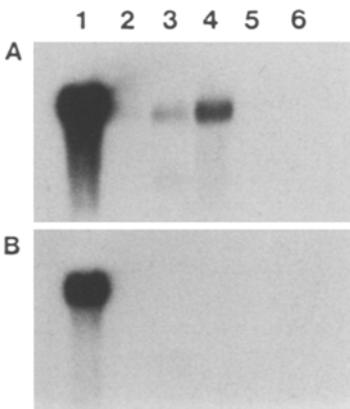


Fig. 5. Northern hybridization using **A** Tac9 and **B** Tac25. Five micrograms poly(A)⁺ RNA from various tissues of *N. tabacum*: lane 1, pollen; lane 2, stigma; lane 3, leaf; lane 4, root and *Brassica napus*: lane 5, pollen; and lane 6, leaf were electrophoresed through a 0.8% formaldehyde gel. Following transfer to Genescreen Plus™ hybridisation was carried out in a solution containing 50% formamide, 10 × Denhardt's, 1% SDS, 100 µg/ml herring sperm DNA, 1 M NaCl-50 mM TRIS, pH 7.5 and 0.1% Sodium pyrophosphate. **A** The filter was probed with the Tac9 5' *Hind*III-*Sst*I fragment (Fig. 2); subsequently it was stripped and **B**, reprobed with the Tac25 *Pst*I-*Scal* fragment (Fig. 4). Filters were washed at 65°C in 0.3 × SSC and autoradiographed

Table 1. Percentage amino acid (above) and nucleotide (below) homology of the most closely related actin genes

	<i>Tac25</i>	<i>Pac58</i>	<i>Pac97</i>	<i>Aac1</i>	<i>Rac1</i>	<i>Pac71</i>
<i>Tac25</i>	—	99.2	98.4	97.6	97.0	96.3
<i>Pac58</i>	84.3	—	98.7	97.9	97.6	96.0
<i>Pac97</i>	84.2	89.5	—	97.4	96.8	96.3
<i>Aac1</i>	80.4	82.3	83.9	—	96.5	96.6
<i>Rac1</i>	79.8	80.7	82.8	80.3	—	95.4
<i>Pac71</i>	81.1	79.1	80.5	80.9	77.7	—

sequences respectively. Tac25 displayed 99.2% amino acid homology and 84.3% nucleotide sequence homology to Pac58. Pac58 has been shown to be allelic to Pac85 (Drouin and Dover 1990). As the sequence of Pac85 is incomplete only a comparison of the coding sequence of exons 2 and 3 to Tac25 was performed. This showed 100% amino acid homology and 84.6% nucleotide sequence homology.

Both rooted and unrooted phylogenies were generated using TREEALIGN (Hein 1990) and PHYLIP (v3.41; Felsenstein 1989) respectively. For this analysis only complete or near complete sequences were compared in order to increase the validity of the comparison. A number of different phylogenetic trees were generated using the available amino acid sequences. It was evident from this analysis that the actins most closely related to Tac25, i.e. all those within the same clade of the trees, were the potato actins Pac58, 85, 71, 97, the *Arabidopsis* actin, Aat1, and the rice actin, Rac1. The amino acid and nucleotide sequence homologies of these actins are shown in Table 1.

Conclusions

Transcription of actin genes has been studied in soybean and rice (Hightower and Meagher 1985; McElroy et al. 1990). It is clear from these studies that actin gene family members differ not only in their developmental regulation, but also exhibit marked differences in the abundance of transcripts in different tissues. The latter is clearly illustrated by transcripts homologous to Tac9, which are more abundant in pollen and roots than in other tissues. In comparison, Tac25 transcripts are only detected in mature pollen. In *Tradescantia*, using the soybean actin gene Sac3 as a probe, under conditions of reduced hybridisation stringency, actin transcripts were detected in developing and mature pollen. These accumulated rapidly following the first microspore mitosis, reaching a maximum before pollen maturity (Stinson et al. 1987). This pattern of gene expression is characteristic of the gametophytic "late genes" (Mascarenhas 1990). It has been estimated that as many as 10 to 20% of the "late gene" transcripts are specific to the male gametophyte (Stinson et al. 1987) and over the last few years a number of these genes have been isolated and characterised; the majority encode pectinases (Mascarenhas 1990; McCormick 1991; Rogers et al. 1992; Allen and Lonsdale 1993). Tac25 represents the first example of a member of a plant actin gene family that is specifically expressed in gametophytic tissue. Presumably the translation products of the Tac25- and the Tac9-related transcripts contribute to the cytoskeleton of the mature pollen grain and may contribute to the cytoskeleton of the growing pollen tube.

Actin fibrils are particularly abundant in pollen tubes and have a predominantly axial orientation. They are thought to play an indispensable role in cell wall organisation, organelle movement and cytoplasmic streaming (Steer and Steer 1989; Heslop-Harrison and Heslop-Harrison 1991). Given the uniqueness of the male gametophyte it is tempting to speculate that the male gametophyte actins have evolved a unique function. The lack of hybridisation of both Tac9 and Tac25 to *Brassica* pollen mRNA suggests that if there is a group of male gametophyte-specific actins then they appear to be as variable in their DNA sequence as other actin genes. Although the DNA sequence may be variable, however, the conservation of the amino acid sequence may indicate a unique male gametophyte-specific function and this aspect is currently being investigated.

Acknowledgements. M.T. would like to thank the Nehru Trust and the INLAKS Foundation for support. D.B. was supported by a Foreign and Commonwealth Fellowship and would like to thank the British Council and Prof. Yuri Gleba for their support. H.J.R. was supported by Agricultural Genetic Co., UK. We are grateful to Dr. George Murphy for his help and advice in running the TREEALIGN and PHYLIP programs and to Alison Harvey for technical assistance. Lastly we would like to thank Richard Meagher for providing the soybean actin clone Sac3 and for his helpful comments during the preparation of this manuscript.

References

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

Baird WV, Meagher RB (1987) A complex gene superfamily encodes actin in petunia. *EMBO J* 6:3223–3231

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. *Phil Trans R Soc B* 274:227–274

Devereux J, Haeblerli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the Vax. *Nucleic Acids Res* 12:387–395

Drouin G, Dover GA (1990) Independent gene evolution in the potato actin gene family demonstrated by phylogenetic procedures for resolving gene conversions and the phylogeny of angiosperm actin genes. *J Mol Evol* 31:132–150

Felsenstein J (1989) PHYLIP – Phylogeny interference package. *Cladistics* 5:164–166

Fyrberg EA, Mahaffey JW, Bond BJ, Davidson N (1983) Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. *Cell* 33:115–123

Hein J (1990) A unified approach to alignments and phylogenies. *Methods Enzymol* 183:626–645

Heslop-Harrison J, Heslop-Harrison Y (1991) The actin cytoskeleton in unfixed pollen tubes following microwave-accelerated DMSO-permeabilisation and TRITC-phalloidin staining. *Sex Plant Reprod* 4:6–11

Hightower RC, Meagher RB (1985) Divergence and differential expression of soybean actin genes. *EMBO J* 4:1–8

Mascarenhas JP (1990) Gene activity during pollen development. *Annu Rev Plant Physiol Plant Mol Biol* 41:317–318

McCormick S (1991) Molecular analysis of male gametogenesis in plants. *Trends Genet* 7:298–303

McElroy D, Rothenberg M, Reece KS, Wu R (1990) Characterization of the rice (*Oryza sativa*) actin gene family. *Plant Mol Biol* 15:257–268

McLean BG, Eubanks S, Meagher RB (1990) Tissue-specific expression of divergent actins in soybean roots. *Plant Cell* 2:335–344

Meagher RB (1991) Divergence and differential expression of actin gene families in higher plants. *Int Rev Cytol* 125:139–163

Melloul D, Aloni B, Calvo J, Yaffe D, Nudel U (1983) Developmentally regulated expression of chimeric genes containing muscle actin DNA sequences in transfected myogenic cells. *EMBO J* 3:983–990

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

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain termination inhibitors. *Proc Natl Acad Sci USA* 74:5463–5467

Shah DM, Hightower RC, Meagher RB (1982) Complete nucleotide sequence of a soybean actin gene. *Proc Natl Acad Sci USA* 79:1022–1026

Shah DM, Hightower RC, Meagher RB (1983) Genes encoding actin in higher plants: introns positions are highly conserved but the coding sequences are not. *J Mol Appl Genet* 2:211–216

Staden J (1986) The current status and portability of our sequence handling software. *Nucleic Acids Res* 14:217–232

Staiger CJ, Schliwa M (1987) Actin localization and function in higher plants. *Protoplasma* 141:1–12

Steer MW, Steer JM (1989) Pollen tube tip growth. *New Phytol* 111:323–358

Stinson JR, Eisenberg AJ, Willing RP, Pe ME, Hanson DD, Mascarenhas JP (1987) Genes expressed in the male gametophyte of flowering plants and their isolation. *Plant Physiol* 83:442–447

Thangavelu M (1989) The actin gene family of tobacco. PhD Thesis, University of Cambridge