Mol Gen Genet (1989) 217:240–245

MGG

Isolation and expression of an anther-specific gene from tomato

David Twell, Rod Wing, Judy Yamaguchi, and Sheila McCormick
Plant Gene Expression Center, USDA-ARS, 800 Buchanan St., Albany, CA 94710, USA

Summary. We have isolated and sequenced an anther-specific cDNA clone and a corresponding genomic clone from tomato. The gene (LAT52) encodes an 800-nucleotide-long transcript that is detectable in pollen, anthers and at 20- to 50-fold lower levels in petals. LAT52 mRNA is not detectable in pistils, sepals or non-reproductive tissues. Steady-state levels of LAT52 mRNA are detectable in immature anthers containing pollen at the tetrad stage and increase progressively throughout microsporogenesis until anthesis (pollen shed). The LAT52 gene contains 5′ and 3′ untranslated regions of 110 and approximately 150 nucleotides, respectively, and a single intron with a highly repetitive sequence. A TATA box motif is located 28 nucleotides upstream of the transcription start site. The gene encodes a putative protein of 18 kDa that is cysteine rich and has an N-terminal hydrophobic region with characteristics similar to eucaryotic secretory signal sequences. LAT52 is a single or low copy gene in tomato and shares homology with sequences in tobacco.

Key words: *Lycopersicon esculentum* – Anther – Microsporogenesis – Pollen

Introduction

The development and function of the male gametophyte (pollen) is of central importance in the reproductive success of most plant species. Along with cytological differentiation, pollen formation is accompanied by biochemical changes such as the synthesis of complex polysaccharides, pigments and storage products (Mascarenhas 1975). Underlying these processes is a complex and unique pattern of gene expression involving genes expressed from both the sporophytic and gametophytic generations.

Several independent studies have demonstrated extensive overlap (60%–90%) between genes expressed in pollen and those expressed in vegetative tissues. Estimates obtained from reassociation kinetics of labelled cDNAs from pollen mRNA hybridized to poly(A)$^+$ RNA from different organs indicate that about 20000 different mRNAs are expressed in pollen, whereas more than 30000 are expressed in vegetative shoots (Willing and Mascarenhas 1984; Willing et al. 1988). Furthermore, there is biochemical and genetic evidence demonstrating the presence of pollen-specific isozymes in many plant species (Tanksley et al. 1981; Sari Gorla et al. 1986; Pedersen et al. 1987). Examples of genes that are expressed exclusively in floral tissues but not necessarily pollen-specific include an a1-tubulin gene in *Arabidopsis* that is preferentially expressed in flowers at anthesis (Ludwig et al. 1988), and a chalcone isomerase gene in petunia that is only expressed in anthers (van Tunen et al. 1988).

We are interested in the molecular processes underlying the development and function of the male gametophyte. These processes inevitably involve the specific expression of genes both in the anther and in pollen. The isolation and characterization of such genes will not only provide molecular markers for the analysis of pollen development, but also allow the basis of their tissue-specific regulation to be investigated. Our analyses have focused on tomato for several reasons: flower development and microsporogenesis are physiologically and genetically well characterized; numerous single gene mutants exist that affect flower development and function, including many male sterile mutants (Rick and Butler 1956); and routine methods for *Agrobacterium*-mediated gene transfer exist (McCormick et al. 1986). Furthermore, since accumulating evidence suggests that microsporogenesis may be a generic developmental pathway, information from tomato should shed light on this process in other plants.

Pollen-specific cDNA clones have been isolated from maize and *Tradescantia* (Stinson et al. 1987), but to date there have been no reports on the structure of such genes. Here we describe the isolation of a floral-specific cDNA from tomato that encodes an abundant mRNA expressed in pollen, and the nucleotide sequence of the corresponding genomic clone.

Materials and methods

Plant material. Lycopersicon esculentum cv VF36 plants were grown under normal greenhouse or field conditions. Harvested plant tissues were frozen in liquid nitrogen and stored at −80° C prior to RNA and DNA isolation. Flowers at three different stages of development were collected, using bud length from the top of the pedicel to the tip of the sepals as an approximate measure of developmental stage. In general, 6–10 mm buds were in meiotic through tetrad stages of microsporogenesis (immature anthers), >14 mm stage buds were at late microspore development with green petals visible through the sepals (green petal

Offprint requests to: S. McCormick

anthers), and mature flowers at anthesis contained mature pollen, with yellow anthers, reflexed petals and sepals (mature anthers). Petals were from mature flowers, and immature seeds (1–3 mm in length with some attached pericarp) were dissected from immature (2–4 cm) green fruit. Leaves were fully expanded and harvested from 3-month-old plants. Three-week-old seedlings were used as the source of RNA for the differential screens because they possess roots, a stem, true leaves and cotyledons, but no floral meristems.

RNA isolation and analysis. Total and poly(A)$^+$ RNA was isolated as in Rochester et al. (1986). Poly(A)$^+$ RNA was electrophoresed on 1.3% formaldehyde gels and blotted to Nytran membranes using standard techniques (Maniatis et al. 1982). Hybridization was performed with random-primer labelled probes (Feinberg and Vogelstein 1984) in 5× SSPE (0.9 M NaCl, 50 mM sodium phosphate, pH 8.3, 5 mM EDTA), 2× Denhardts (0.04% Ficoll, 0.04% polyvinylpyrrolidone, 0.04% bovine serum albumin), 0.1% SDS, 100 μg/ml denatured salmon sperm DNA and 100 μg/ml poly(A) at 68° C. Filters were washed in 0.3× SSPE, 0.1% SDS at 68° C.

cDNA library construction and screening. A cDNA library from mature anther poly(A)$^+$ RNA was constructed in lambda gt10 essentially according to Huynh et al. (1985). Clones were screened differentially, on duplicate plaque filters and subsequently on Southern filters, with ^{32}P-labelled single-stranded probe prepared from either mature anther or seedling poly(A)$^+$ RNA.

Genomic clone isolation. A tomato (*L. esculentum* Mill cv VF36) genomic library constructed in bacteriophage vector EMBL3 (Frischauf et al. 1983; gift of C. Gasser) was screened according to Benton and Davis (1977) with the random-primer ^{32}P-labelled insert of pLAT52. Hybridizing restriction fragments of the corresponding genomic clone, gLAT52, were subcloned into plasmid vectors pGEM7Zf(+) (Promega) and pBluescript (Stratagene) according to standard methods (Maniatis et al. 1982).

DNA sequence analysis. DNA sequencing was performed by the dideoxy chain termination method (Sanger et al. 1977) using modified T7 DNA polymerase supplied by U.S. Biochemicals. Both strands of the cDNA and genomic clone corresponding to LAT52 were sequenced using single-stranded M13, and double-stranded plasmid DNA templates, with synthetic oligonucleotides as primers. Computer analysis of the DNA sequence and predicted protein sequence was performed using the PCGene (Intelligenetics) and University of Wisconsin sequence analysis software (Devereaux et al. 1984).

Plant DNA isolation and analysis. DNA was isolated from tomato leaves as in Bernatzky and Tanksley (1986a). DNA was digested with restriction endonucleases, electrophoresed through 0.8% agarose gels and blotted to Nytran membranes. Hybridization and washing conditions were as described for Northern blot analysis.

Primer extension and RNase protection analysis. The 5′ terminus of the LAT52 gene transcript was determined by the primer extension method according to Gidoni et al. (1988). In each reaction, 0.2 pmol ^{32}P-5′ end-labelled primer was annealed with 1 μg poly(A)$^+$ RNA for 5 h at 45° C. Following extension with AMV reverse transcriptase for 45 min at 37° C, RNA was base hydrolysed, the products analysed on an 8% sequencing gel and visualized by autoradiography. The 3′ terminus of the LAT52 gene transcript was determined by RNase protection mapping as described by Melton et al. (1984). In each reaction, in vitro ^{32}P-labelled antisense transcript was annealed with 1 μg poly(A)$^+$ RNA for 16 h at 42° C. Following RNase A (30 μg/ml) and RNase T1 (840 units/ml) digestion for 60 min at 35° C, the final protection products were analysed on a 4% sequencing gel and visualized by autoradiography.

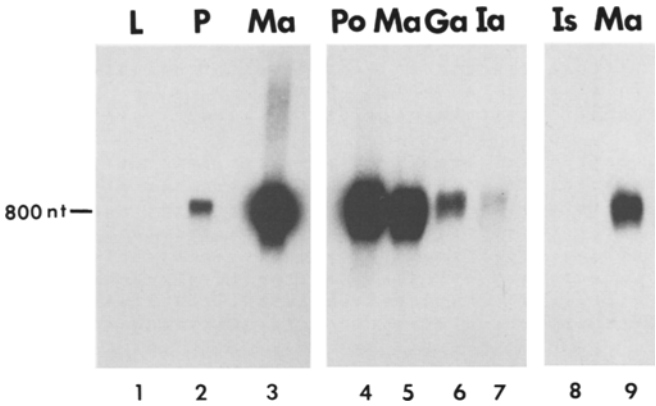


Fig. 1. Northern blot analysis of LAT52 mRNA in tomato. Filters with 2 μg poly(A)$^+$ RNA isolated from: leaves (L), petals (P), immature seeds (Is), pollen (Po), and anthers at the immature (Ia), green petal (Ga) and mature (Ma) stages of development. The RNA was hybridized with ^{32}P-labelled cDNA clone pLAT52. The size of the LAT52 mRNA is shown in nucleotides (nt)

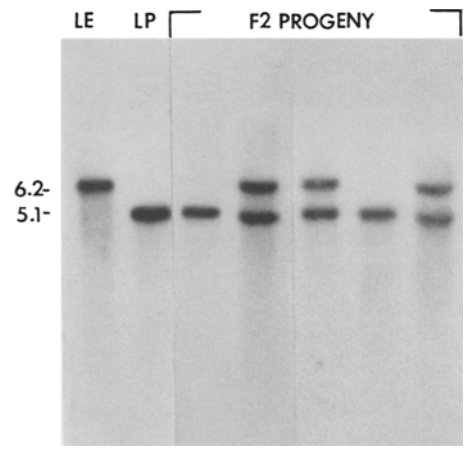
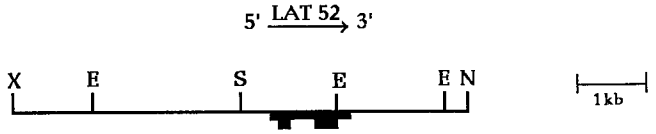


Fig. 2. Southern blot analysis of genomic sequences corresponding to cDNA clone pLAT52. Filters with 10 μg genomic DNA from *Lycopersicon esculentum* (LE), *L. pennellii* (LP) and five F_2 progeny plants digested with *Eco*RV were hybridized with ^{32}P-labelled pLAT52. The sizes of hybridizing fragments are shown in kb. Among 48 F_2 plants only 2 were homozygous for the *L. esculentum* restriction fragment length polymorphism (none are shown above). A similar over-representation of *L. pennellii* alleles in F_2 progeny was found for other cDNA markers (Bernatzky and Tanksley 1986b).



A

```
*           *           *           *           *
CCTATACCCCTTGGATAAGGGTAGCTCTATCTATATAGTCAATTATTGTCTTCTGTCTGT  -497
TGGTGTCGACATACTCGACTCAGAAGGTATTGAGGAATGATCGATTCTGGGTCATTTGIG  -437
IGGIIAATCACCCTCCAAATCAACTAAGTCATCCTGAAGGACAATATCCTATTTTTCTC  -377
TCGTAGGTTTATCATTTAAATTACTATCGCGTGATAATTTTGTAACGTAGAAAAATAATA  -317
CCATTAATCCAAACGTTATATTCATTAAAATAATTATGATACATTTAAAAATATTTCGTG  -257
ACCTCTCAATTATTGCAAATTCTAAGCCATCCCAAGTTTTGAGGCTAATTTTTTTTACTA  -197
TACTATTTTTACAACCACAAAAACATAAAAAATAAAAAATAAAAAAATAAACCGAGTCA  -137
ATTGCTACAATCACTTCATTATTAATTTTAATTAATATTAIGIGGIIATATATGAAACTG   -77
TTAGAGAAATAATAGCTCCACCATATTTTTTTCTCAATTTATTTTCACTATAAAAAGGCT   -17
ATTTCATTATAATCAAaacaagacacacacaaagagaaggagcaataaaataaaagtaaa    44
caacaatttgtgtGtttaaaaaaaaaaaaaaagtacacacaccaaaaaaaaaaattccaa   104
tttaaaATGGCAAAGGCTATTGTGCTCCTCTCTGCTCTTTGCATTTTGGCCCTTGCCAAT   164
        M  A  K  A  I  V  L  L  S  A  L  C  I  L  A  L  A  N
TTTGCACATTGCCGCCCTGAAGTTTTTGATGTTGAAGGAAAGGTCTACTGTGATACTTGC   224
F  A  H  C  R  P  E  V  F  D  V  E  G  K  V  Y  C  D  T  C
CGCGTTCAATTCGAAACAAAACTCAGCGAGAACCTTGAAGgtatatactatattttcatt   284
R  V  Q  F  E  T  K  L  S  E  N  L  E  G
cacttcgttctatcagactcatttcatcttatatatatttacattcactttgttctattc   344
aggctcatttcatcttatataaatttacattcacttcgttctattcaggctcatttcatc   404
ttatgtatatttacattcacttagttctattcaggctcatttcatcttatatatatttac   464
attcacttcgttctattcagactcatttcatcttatatatatttacattcactttgtttt   524
attcagactcatttcatcttatataaatttacattcacttcgttctattcaggctcattt   584
catcttatgtatatttacattcacttcgttctattcagactcatttcatcttatatatat   644
atacattcactttgttctattcagactcatttcatctaataactccgtattgattatgtt   704
ttgttttgttgcaaatagGTGCAACTGTGAAGTTGCAATGCAGGAACATCAGCACAGAGG   764
                    A  T  V  K  L  Q  C  R  N  I  S  T  E  A
CTGAGACGTTTTCAGTGGAAGGTGTGACTGATAAAGATGGCAAGTACAAATTAACCGTTA   824
  E  T  F  S  V  E  G  V  T  D  K  D  G  K  Y  K  L  T  V  N
ATGGAGACCACGAGAACGATATTTGCGAGGTGACAGTTGTGAAAAGCCCAAGGGAGGACT   884
  G  D  H  E  N  D  I  C  E  V  T  V  V  K  S  P  R  E  D  C
GCAAAGAGAGTGTGTCTGGATATGAAAAGGCAAGAATTGAGTGTAGTGACAATGTTGGTA   944
  K  E  S  V  S  G  Y  E  K  A  R  I  E  C  S  D  N  V  G  I
TCCACAATGCTGTGAGATTTGCCAATCCACTTTTCTTCATGAAGGCTGAGTCTGTTCAAG  1004
  H  N  A  V  R  F  A  N  P  L  F  F  M  K  A  E  S  V  Q  G
GATGCAAAGAAGCTCTTGATGAGTTGGGTTTATTCCCCCTTGAATTCTAAaaaagtctttt 1064
  C  K  E  A  L  D  E  L  G  L  F  P  L  E  F  STOP
ttctttttttttaacgtgacagaaaataaattagtccttttcatgttgtaatgcgaaata  1124
ataaaaatatgaggaaaagatggatttctcatttatatgtattcatatgggatgataagt  1184
tccatgtatttgttttttaatgagaagatcttgatcatatttaAGAAACTATGTTCTCTTC  1244
TTGGTGTTTG
```

B

Fig. 3A and B. Physical map and DNA sequence of genomic clone gLAT52. **A** Restriction map of a region of gLAT52 showing LAT52 gene structure. *Solid bars* represent transcribed regions, with exons as *thick bars*, and untranslated DNA (including the single intron) as *thin bars*. Restriction sites are indicated as single letter abbreviations: E, *Eco*RI; N, *Nsi*I; S, *Sal*I; X, *Xba*I. **B** Nucleotide sequence of the LAT52 gene including 5′ and 3′ flanking DNA. Numbering is from the major transcription start site (+1). Transcribed, but untranslated regions are in *lower case*. The most 5′ nucleotide of cDNA clone pLAT52 is shown in *bold type* (+58). Single letter amino acid codes are shown below the predicted coding DNA and a putative N-linked glycosylation site (NXS/T) is *underlined*. Features mentioned in the text are indicated as follows: the major (+1) and minor (+26) transcriptional start sites and the putative poly(A) addition sites are *arrowed*; core TATA box and polyadenylation signal homologies are *boxed*; repeated sequences in the 5′ flanking DNA are *underlined*; *solid lines* indicate repeats that share TCAATT, *broken lines* indicate repeats containing SV40 core enhancer sequence

Results

Isolation of cDNA clones

The general strategy to isolate anther-specific genes has been described (McCormick et al. 1987). Five anther-expressed clones were selected for further analysis on the basis of their lack of cross-hybridization to each other. Here we report the results with pLAT52; results with the other clones will be reported elsewhere.

Analysis of LAT52 mRNA expression

Northern blot analysis of poly(A)$^+$ RNA isolated from vegetative (leaf, stem and root) and reproductive (anther, pistil, fruit and immature seed) tomato organs showed that LAT52 mRNA is only found in flowers (Fig. 1 and McCormick et al. 1987). The pLAT52 cDNA insert hybridized to an 800-nucleotide-long RNA species found exclusively in mature anthers and petals. By comparison of hybridization signals using densitometry (data not shown) it was estimated that steady-state levels of LAT52 mRNA in mature anthers were 20- to 50-fold higher than in petals and at least 200-fold higher than in vegetative or other floral organs, where no signal was detected (Fig. 1).

The developmental profile of LAT52 mRNA showed that it is expressed throughout anther development with maximum accumulation in mature anthers (Fig. 1, lanes 5–7). A significant fraction of pLAT52 mRNA in mature anthers can be attributed to expression in the pollen (Fig. 1, lane 4). This is also suggested by analysis of several male sterile mutants of tomato which lack mature pollen and show markedly reduced levels of LAT52 mRNA in anthers (Ursin et al., in preparation). These data demonstrate that the LAT52 gene is expressed both in haploid (pollen) and diploid (petal) flower tissues.

Southern blot analysis

Genomic DNA was isolated from *L. esculentum* and *L. pennellii* (parental species) and several F_2 progeny plants de-

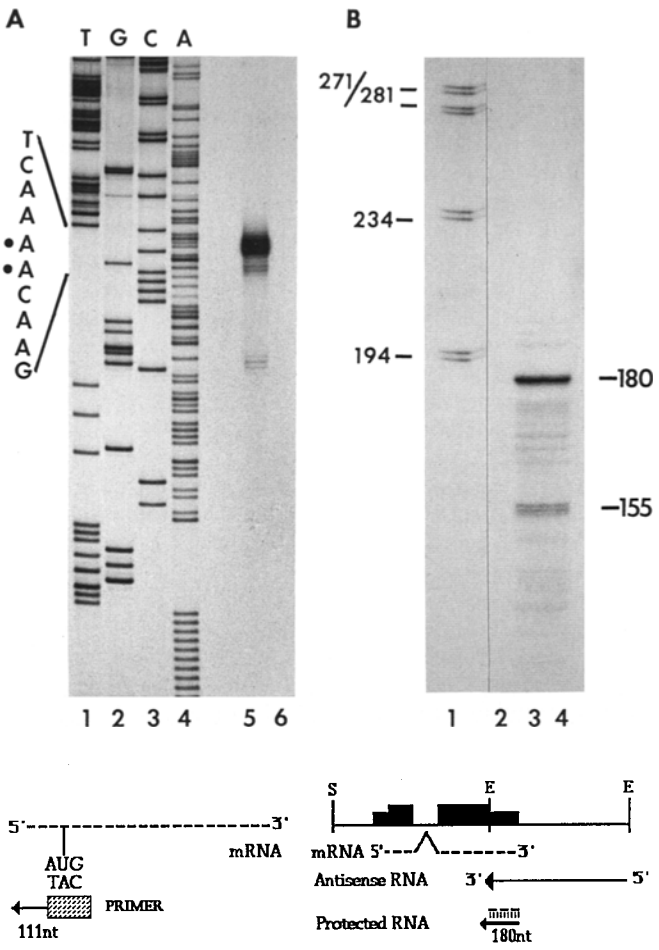


Fig. 4A and B. Primer extension and RNase protection analysis of LAT52 mRNA. **A** Primer extension was carried out using 1 μg poly(A)$^+$ RNA from mature anthers (lane 6) or tRNA (lane 5) as templates, and a 20mer oligonucleotide primer, as shown below the resulting autoradiograph. The sequence shown to the *left* (lanes 1–4) was obtained from gLAT52 DNA using the same primer. Bases corresponding to the major extension products are indicated as *dots*. **B** RNase protection was carried out using 1 μg poly(A)$^+$ RNA from mature anthers (lanes 3 and 4) or tRNA (lane 2), hybridized with in vitro ^{32}P-labelled antisense RNA made from a 1.7 kb *Eco*RI fragment of gLAT52, as shown below the resulting autoradiograph. Hybridization was at 37° C (lanes 2 and 3) or 45° C (lane 4). The approximate sizes (in bp) of the major protected RNAs are shown to the *right*. Size markers (in bp) were 5′ end-labelled *Hae*III fragments of phage φX174 (lane 1)

rived from a single F_1 hybrid between the parental species. DNA was digested to completion with restriction endonuclease *Eco*RV and subjected to Southern blot analysis using pLAT52 as a probe. In *L. esculentum* and *L. pennellii* single differently sized hybridizing bands were detected (Fig. 2, lanes 1 and 2). This restriction fragment length polymorphism (RFLP) showed segregation among the F_2 individuals (Fig. 2, lanes 3–6). Taken together these data strongly suggest that LAT52 corresponds to a single copy gene.

Isolation of genomic clone gLAT52

An amplified tomato cv VF36 genomic library constructed in bacteriophage EMBL3 was screened by plaque hybridization using pLAT52 as a probe. Two hybridizing clones were obtained from 2×10^6 plaques and one of these, gLAT52, was characterized further by restriction mapping and hybridization to cDNA sequences. All sequences hybridizing to cDNA pLAT52 were within a 1.5 kb *Sal*I-*Eco*RI restriction fragment (Fig. 3a). Hybridization to the adjacent 1.7 kb *Eco*RI fragment was not detected because pLAT52 did not contain 3′ untranslated sequences.

Structure of the LAT52 gene

The complete cDNA insert (536 bp) and 1810 bp of the genomic clone gLAT52 (including 666 bp of 5′ and 200 bp of 3′ flanking DNA) were sequenced (Fig. 3b). Comparison of the cDNA and genomic sequences showed the presence of a single intron of 458 nucleotides. The intron is characteristically AT-rich (75%) compared with the coding DNA (55% AT), and the exon-intron junction obeys the GT-AG rule (Breathnach and Chambon 1981). The cDNA and genomic sequences contain an identical open reading frame of 483 nucleotides. The only difference between the two sequences is the presence of two additional A residues (position +62) within the 5′ untranslated region of the cDNA, which we suggest may be a cloning artefact. In addition, cDNA clone pLAT52 does not include sequences immediately 3′ of the TAA translational stop codon, and so does not contain 3′ flanking DNA or a poly(A) tail.

Primer extension of mature anther RNA was used to determine the transcriptional start site of LAT52 (Fig. 4A). The major extension products map 110 nucleotides upstream of the putative translational initiation codon. The sequences around this AUG codon agree well with the consensus derived for other plant genes (Lutcke et al. 1987). A TATA box sequence (Messing et al. 1983) is present 29 nucleotides upstream of the major transcriptional start site. There is also a minor cluster of extension products which map to position +26, 35 nucleotides downstream of a second TATA box sequence. The significance of this feature is unknown although multiple start sites have been observed for several other genes expressed in plants (Langridge and Feix 1983; Tischer et al. 1986; Bruce et al. 1987). The 5′ untranslated DNA contains two poly(dA) regions, which contribute to a leader sequence with an unusually high (61%) adenylate content (Fig. 3b).

The length of the 3′ untranslated region of the LAT52 mRNA was determined by RNase protection. Two major protected RNA species were obtained (Fig. 4B), indicating that two alternative poly(A) addition sites are used, approximately 147 and 172 nucleotides 3′ of the TAA stop codon. The predicted length(s) of the 3′ untranslated region agree well with those determined for other plant genes (Messing et al. 1983). The putative poly(A) addition sites of the LAT52 mRNA are not preceded by the consensus polyadenylation signal AATAAA in the expected location. However two examples of this sequence occur 35 and 70 nucleotides 3′ of the translational stop codon.

An unusual intron structure

Analysis of the single intron within the LAT52 gene showed that it is composed almost entirely of a directly repeated sequence of 46 nucleotides (Fig. 5). The repeat occurs nine times as a tandem array and is flanked by 8 and 37 nucleotides of unique DNA at the 5′ and 3′ ends of the intron,

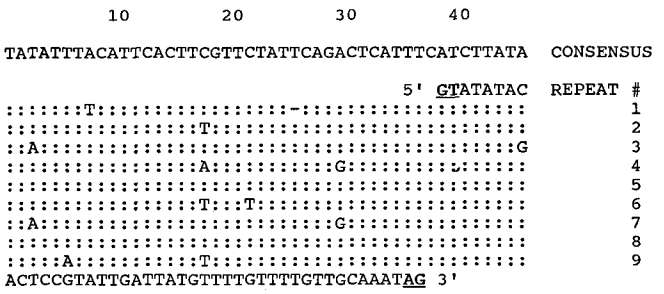


Fig. 5. LAT52 intron structure. The complete intron sequence is represented, with exon-intron borders *underlined*. Only nucleotides differing from those of the derived consensus repeat sequence are shown. A single gap (–) was inserted into repeat 1 to maximize alignment

respectively. Throughout the intron the repeats are highly conserved with the positions of sequence divergence appearing somewhat clustered at three positions. Repeated sequences within other plant gene introns have been noted (Dennis et al. 1984; Mignery et al. 1988), but in these cases the repeats extend into exon DNA and/or do not account for most of the intron length.

Coding region of pLAT52

The cDNA clone pLAT52 contains an open reading frame that codes for a putative protein of 17.8 kDa. This is in accord with the estimated size (18 kDa) of the in vitro translation product synthesized from RNA hybrid-selected from mature anther poly(A)$^+$ RNA with pLAT52 (data not shown). The predicted protein sequence shows a strongly hydrophobic domain at the N-terminus with two potential signal cleavage sites at 17 and 22 amino acid residues downstream of the initiator methionine, both of which obey the "(−3, −1) rule" (von Heijne 1983). The predicted protein is hydrophilic and acidic (pI = 4.6), with a potential N-linked glycosylation site at amino acid position 61, and a relatively high cysteine (5.6%) content. No significant similarities to the predicted protein sequence of LAT52 were detected in computer searches of the NBRF-PIR, Genbank and EMBL databases.

Discussion

Northern blot hybridization experiments demonstrate that the expression of LAT52 is strictly floral-specific. LAT52 mRNA was detectable only in anthers and, at 20- to 50-fold lower levels in petals. This demonstrates that control of LAT52 gene expression is exerted not only in the floral lineage, but also during the differentiation of individual organs within the flower. We have not determined whether transcriptional and/or post-transcriptional control mechanisms are responsible for the observed organ specificity.

The increase in abundance of LAT52 mRNA during anther development resembles that observed for pollen-specific mRNAs cloned from maize and *Tradescantia* (Stinson et al. 1987). These mRNAs accumulate progressively after microspore mitosis, while other mRNAs expressed in pollen (e.g. actin mRNA) decrease in abundance following late pollen interphase. The abundant expression of LAT52 mRNA in mature pollen may largely account for the observed accumulation during anther development, although in situ hybridization studies show that LAT52 mRNA is also expressed in the anther wall (Ursin et al., in preparation).

Southern blot analysis showed that pLAT52 hybridized to a single restriction fragment in *L. esculentum* and to a single different sized fragment in *L. pennellii* (Fig. 2). This RFLP segregated in F_2 plants derived from *L. esculentum* × *L. pennellii* F_1 hybrids (Fig. 2). These data strongly suggest that LAT52 represents a single copy gene, with no other closely related sequences in the tomato genome. LAT52 and several of the other anther-expressed genes also share sequence homology with genomic DNA of tobacco (data not shown). This evolutionary conservation may indicate that similar programs of gene expression will be conserved in microsporogenesis, and should allow us to analyse this process in other species.

Since pollen contains a store of presynthesized mRNAs (Mascarenhas et al. 1984), structural features of these mRNAs may contribute to their stability. In this regard the LAT52 transcript has a relatively long 5′ untranslated leader (Joshi 1987) with an unusually high (61%) content of adenylate residues. The significance of the latter feature is unknown, although it would prevent the occurrence of any extensive secondary structure in the leader.

The abundance of LAT52 mRNA in mature pollen suggests that the translation product may play a role during germination or early tube growth. A striking feature of the LAT52 protein is the presence of nine cysteine residues, which provide the potential for intra- and intermolecular crosslinking through disulfide bridges. The putative signal sequence of the LAT52 protein may indicate that it is secreted. This is in accord with the occurrence of many proteins in the pollen wall that are thought to be secreted (Knox and Heslop-Harrison 1970). Although we cannot predict the role of LAT52 mRNA found in petals and the anther wall, antibodies to the LAT52 protein will allow us to determine whether the mRNA is translated in these organs and in pollen prior to germination. This will also allow the cellular and subcellular localization of the protein during microsporogenesis and pollen tube growth.

Finally, analysis of the 5′ flanking DNA of the LAT52 gene identified several directly repeated sequences that may be involved in transcriptional regulation (Fig. 3b). Most notable are three pairs of 8–10 bp repeats all of which share the hexanucleotide TCAATT and an 8 bp repeat that includes the SV40 core enhancer sequence (Khoury and Gruss 1983). We are currently conducting a functional analysis of these sequences and characterizing promoters of other anther expressed genes in order to identify *cis*-acting sequence elements important in anther-specific expression.

Acknowledgements. S.M. thanks Robb Fraley and Monsanto for providing support to initiate the project. We thank Chuck Gasser for the genomic library, Susan Larabell for Southern blot analysis, Elizabeth Wang and Ryan Alfonso for dissection of flowers and Steven Goff for synthesis of oligonucleotides. This work was supported by USDA CRIS number 5335-22230-002-00D

References

Benton WD, Davis RW (1977) Screening lambda gt recombinant clones by hybridization to single plaques *in situ*. Science 196:180–182

Bernatzky R, Tanksley SD (1986a) Genetics of actin related sequences in tomato. Theor Appl Genet 72:314–321

Bernatzky R, Tanksley SD (1986b) Majority of random cDNA correspond to single loci in the tomato genome. Mol Gen Genet 203:8–14

Breathnach R, Chambon P (1981) Organization and expression of eukaryotic split genes coding for proteins. Annu Rev Biochem 50:349–383

Bruce WB, Gurley WB (1987) Functional domains of a T-DNA promoter active in crown gall tumors. Mol Cell Biol 7:59–67

Dennis ES, Gerlach WL, Pryor AJ, Bennetzen JL, Inglis A, Llewellyn D, Sachs MM, Ferl RJ, Peacock WJ (1984) Molecular analysis of the alcohol dehydrogenase (Adh1) gene of maize. Nucleic Acids Res 12:3983–4000

Devereaux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Feinberg AP, Vogelstein B (1984) A technique for radiolabelling DNA fragments to high specific activity. Anal Biochem 137:266–267

Frischauf AM, Lehrach H, Poustka A, Murray N (1983) Lambda replacement vectors carrying poly linker sequences. J Mol Biol 170:827–842

Gidoni D, Bond-Nutter D, Brosio P, Jones J, Bedbrook J, Dunsmuir P (1988) Coordinated expression between two photosynthetic petunia genes in transgenic plants. Mol Gen Genet 211:507–514

Huynh TV, Young RA, Davis RW (1985) Constructing and screening cDNA libraries in lambda gt10 and gt11. In: Glover D (ed) DNA cloning techniques, a practical approach. IRL Press, Oxford, pp 49–78

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

Khoury G, Gruss P (1983) Enhancer elements. Cell 33:313–314

Knox RB, Heslop-Harrison J (1970) Pollen-wall proteins: localization and enzyme activity. J Cell Sci 6:1–27

Langridge P, Feix G (1983) A zein gene of maize is transcribed from two widely separated promoter regions. Cell 34:1015–1022

Ludwig SR, Oppenheimer DG, Silflow CD, Snustad DP (1988) The a1-tubulin gene of *Arabidopsis thaliana*: primary structure and preferential expression in flowers. Plant Mol Biol 10:311–321

Lutcke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA (1987) Selection of AUG initiation codons differs in plants and animals. EMBO J 6:43–48

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

Mascarenhas JP (1975) The biochemistry of angiosperm pollen development. Bot Rev 41:259–314

Mascarenhas NT, Bashe D, Eisenberg A, Willing RP, Xiao CM, Mascarenhas JP (1984) Messenger RNAs in corn pollen and protein synthesis during germination and pollen tube growth. Theor Appl Genet 68:323–326

McCormick S, Niedermeyer J, Fry J, Barnason A, Horsch R, Fraley R (1986) Leaf disc transformation of cultivated tomato (*L. esculentum*) using *Agrobacterium tumefaciens*. Plant Cell Rep 5:81–84

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

Melton DA, Krieg PA, Rebagliatti MR, Maniatis T, Zinn K, Green MR (1984) Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res 12:7035–7056

Messing J, Geraghty D, Heidecker G, Hu N-T, Kridl J, Rubenstein I (1983) Plant gene structure. In: Kosuge T, Meredith CP, Hollaender A (eds) Genetic engineering of plants. Plenum, New York, pp 211–227

Mignery GA, Pikaard CS, Park WD (1988) Molecular characterization of the patatin multigene family of potato. Gene 62:27–44

Pedersen S, Simonsen V, Loeschcke V (1987) Overlap of gametophytic and sporophytic gene expression in barley. Theor Appl Genet 75:200–206

Rick CM, Butler L (1956) Cytogenetics of the tomato. Adv Genet 8:267–382

Rochester DE, Winter JA, Shah DM (1986) The structure and expression of maize genes encoding the major heat shock protein, hsp70. EMBO J 5:451–458

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

Sari Gorla M, Frova C, Binelli G, Ottaviano E (1986) The extent of gametophytic-sporophytic gene expression in maize. Theor Appl Genet 72:42–47

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

Tanksley SD, Zamir D, Rick CM (1981) Evidence for extensive overlap of sporophytic and gametophytic gene expression in *Lycopersicon esculentum*. Science 213:453–455

Tischer E, DasSarma S, Goodman H (1986) Nucleotide sequence of an alfalfa glutamine synthetase gene. Mol Gen Genet 203:221–229

van Tunen AJ, Koes RE, Spelt CE, van der Krol R, Stuitje AR, Mol JNM (1988) Cloning of two chalcone flavanone isomerase genes from *Petunia hybrida*: coordinate, light-regulated and differential expression of flavonoid genes. EMBO J 7:1257–1263

von Heijne G (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 133:17–21

Willing RP, Mascarenhas JP (1984) Analysis of the complexity and diversity of mRNAs from pollen and shoots of *Tradescantia*. Plant Physiol 75:865–868

Willing RP, Bashe D, Mascarenhas JP (1988) An analysis of the quantity and diversity of messenger RNAs from pollen and shoots of *Zea mays*. Theor Appl Genet 75:751–753

Communicated by E. Meyerowitz

Received November 20, 1988

Note added in proof

The predicted protein sequence of LAT52 shows 32% amino acid identity to the predicted protein sequence of a pollen specific gene (Zm13) from corn (Hanson, DD et al. (1989) The Plant Cell 1:173–179).