Mol Gen Genet (1991) 229:238–244
002689259100300U

MGG

A 2.6 kb intron separates the signal peptide coding sequence of an anther-specific protein from the rest of the gene in sunflower

Claire Domon, Jean-Luc Evrard, Datta T.N. Pillay [1], and André Steinmetz

Institut de Biologie Moléculaire des Plantes, 12 rue du Général Zimmer, F-67084 Strasbourg, France
[1] University of Windsor, Windsor, Ontario, Canada N9B 3P4

Received May 6, 1991

Summary. We have isolated and sequenced an anther-specific gene from sunflower which encodes an 800-nucleotide transcript detectable in the peripheral anther cells. It contains an intron of 2615 bp, which separates the first exon (77 bp) coding for a putative signal peptide of 21 amino acids, from the second exon (563 bp) coding for a 100 amino acid polypeptide. The 5′ and 3′ untranslated regions comprise respectively 13 and 264 bp. The *SF2* gene is present in the sunflower genome in several copies, all or most of which contain a closely related intron.

Key words: *Helianthus annuus* – DNA sequence – Multicopy gene

Introduction

Anthers are floral structures specialized for pollen production in higher plants. Their development, including that of the pollen grains, brings into play a considerable number of different genes, estimated to number about 25000 in tobacco, of which approximately 10000 are anther-specific (Kamalay and Goldberg 1980, 1984). In *Tradescantia* and maize, at most 5000 of the 20000 to 24000 different mRNAs found in pollen are believed to be exclusively expressed in pollen (Mascarenhas 1990). As stamens develop from their primordia following floral induction, specific subsets of cells enter different pathways and differentiate into a limited number of different cell types: microspore mother cells (which will produce the pollen grains), tapetum, connective tissue, vascular tissue, stomium, endothecium and epidermis. The structural and functional features of these various cell types are genetically determined; their development is regulated by several independent gene expression programs (Koltunow et al. 1990).

Offprint requests to: A. Steinmetz

In a first approach to the elucidation of these various regulatory processes, several laboratories have isolated and sequenced anther-specific cDNAs and/or genes from tomato (McCormick et al. 1989; Ursin et al. 1989; Twell et al. 1989), tobacco (Smith et al. 1990; Seurinck et al. 1990), maize (Hanson et al. 1989), *Oenothera* (Brown and Crouch 1990) and *Tradescantia* (Stinson et al. 1987). Although upstream DNA fragments which are able to promote anther-specific expression of a reporter gene have been isolated from tomato and petunia genes (McCormick et al. 1989; van Tunen et al. 1990), there is no clear evidence as to what sequence elements are really involved. A 122 bp region, which is both necessary and sufficient to program tapetal-specific expression, has recently been identified in tobacco (Koltunow et al. 1990). A putative "anther-box" has been described in chalcone flavanone isomerase (CHI) genes from petunia (van Tunen et al. 1989), but this assertion requires for verification a detailed experimental analysis of the promoter region.

We have recently isolated several flower-specific cDNA clones from sunflower (Herdenberger et al. 1990). The nucleotide sequences of two of these cDNAs (SF2 and SF18) have been determined and were found to code for structurally related anther-specific proteins, with no homology to any protein sequence published to date (Domon et al. 1990). In situ hybridization experiments have shown that the SF2 and SF18 mRNAs are found exclusively in anther wall cells (Evrard et al. 1991). In the present paper we describe the structure and nucleotide sequence of the anther-specific gene *SF2* which was isolated from a genomic sunflower DNA library using the SF2 cDNA clone as a probe.

Materials and methods

Plant material. Sunflower (*Helianthus annuus* L.) seeds from the inbred male-fertile line HA401B were obtained from Cargill (France). For isolation of nuclear DNA,

seeds were surface-sterilized in a 10% sodium hypochlorite solution for 15 min. They were then thoroughly rinsed with tap water, soaked for 2 h in water and subsequently spread on several layers of filter paper. They were grown for 3 to 4 days at 24° C in the dark.

Isolation of nuclear DNA. Nuclear DNA was isolated from 3- to 4-day-old hypocotyls using a modification of a protocol described by Dellaporta et al. (1983). Thirty grams of hypocotyls were ground in liquid nitrogen (using mortar and pestle) to a fine powder which was subsequently suspended in 200 ml homogenization buffer (100 mM TRIS-HCl pH 8.0, 50 mM EDTA, 500 mM NaCl and 10 mM β-mercaptoethanol) and lysed upon addition of 20% SDS (1% final concentration). After 30 min incubation at 65° C, 0.3 volume of 3 M potassium acetate, pH 4.8, was added and, after gentle shaking, the mixture was put on ice for 30 min. The dense, white precipitate was pelleted by centrifugation and the nucleic acids in the supernatant were precipited with 2 volumes of ethanol. Nucleic acids were recovered by a low speed centrifugation ($2000 \times g$), washed with 70% ethanol, and subsequently dissolved in 15 ml TE buffer (10 mM TRIS-HCl, 1 mM EDTA pH 8). Fifteen grams of CsCl and 400 µl ethidium bromide (EtBr; 10 mg/ml) were added and the DNA was gradient-purified by centrifugation for 6 h at 65000 rpm in a VTi65 rotor. After collection of the DNA band and removal of EtBr with isoamyl alcohol, the DNA solution was diluted with 1 volume of distilled water and ethanol-precipitated, DNA fibers were collected with a pasteur pipette, washed in 70% ethanol, air dried for 5 min and dissolved in 1 ml TE.

Isolation of phage and plasmid DNA. For large-scale growth of phage Charon 40 (Dunn and Blattner 1987) or its recombinant derivatives, 1 ml of overnight culture of *Escherichia coli* ED8767 were infected with 2.5×10^7 phage (15 min at 37° C) in the presence of infection salts containing 10 mM $MgCl_2$ and 10 mM $CaCl_2$ and subsequently transferred into 1 l of NZCM growth medium in a 2 l Erlenmeyer flask. After vigorous shaking at 37° C for 17 h, 60 g NaCl, 5 ml of chloroform and 125 µl DNase I (10 mg/ml) were added and shaking was continued until all salts were dissolved. Unlysed cells and cell debris were removed by centrifugation and phage was precipitated from the supernatant by addition of 70 g polyethylene glycol (PEG 8000). After 4 h at 4° C, phage was pelleted by 15 min centrifugation at $5000 \times g$, resuspended in 2 ml of lambda dilution buffer (10 mM TRIS-HCl pH 7.4, 10 mM $MgSO_4$) and chloroform-extracted. Following centrifugation the aqueous phase was recovered and phage was purified by two runs on CsCl block gradients as described by Davis et al. (1980). DNA was isolated from purified phage using the formamide lysis method (Davis et al. 1980). Plasmid DNA was isolated from 100 ml bacterial cultures (NM522) grown overnight in the presence of the appropriate antibiotic (ampicillin) according to the alkaline lysis method (Birnboim and Doly 1979).

Construction of the genomic library. Partial digestion of genomic DNA with *Sau*3A was carried out in three 1.5 ml Eppendorf tubes, each containing 100 µg DNA, 1 × low-salt restriction buffer and 4 units of *Sau*3A restriction enzyme (New England Biolabs) in a reaction volume of 1 ml. Digestion was allowed for 35 min at 37° C, following which the enzyme was inactivated by addition of EDTA to a final concentration of 50 mM and by heating at 75° C for 10 min. DNA was ethanol-precipitated, pelleted by centrifugation and dissolved in 300 µl of TE buffer. Fragments were then size-fractionated on NaCl gradients as described by Kaiser and Murray (1985), except that discontinuous gradients (25%, 20%, 15%, 10% and 5%) were used. Fractions of 150 µl were collected and aliquots (10 µl) were checked on a 0.3% agarose gel in the presence of size markers (λgt10 cut with *Bgl*II which produces fragments of 24, 11 and 9 kb). Fractions containing fragments in the size range 11–24 kb were diluted with 1 volume of water, ethanol-precipitated, dissolved in 10 µl TE buffer and stored at −20° C. Ten micrograms of Charon 40 arms were then ligated to 5 µl of size-fractionated DNA in a 50 µl reaction volume overnight at 12° C. Recombinant phage DNA was subsequently packaged (2.5 µl of the ligation per set of packaging extracts), and the phage particles were used to transfect ED8767 bacteria.

Screening of genomic library. Infection doses of 2×10^6 pfu in ED8767 bacteria were plated on 8 Nunc plates (24.5 × 24.5 mm), grown overnight at 37° C and transferred to nitrocellulose membranes as described by Benton and Davis (1977). The filters were subsequently hybridized (in 5 × SSC, 5 × Denhardt's, 0.5% SDS) for 12 h at 65° C with ^{32}P-labelled SF2 cDNA at a concentration of 2 ng/ml of hybridization solution. Filters were washed at 65° C in 5 × SSC, (1 × SSC is 0.15 M NaCl, 15 mM sodium citrate), in 1 × SSC, 0.1% SDS and finally in 0.1 × SSC, 0.1% SDS. They were then dried and autoradiographed at −80° C using intensifying screens.

Radiolabelling of DNA probes. Labelling was by random priming (Feinberg and Vogelstein 1984) of the 550 bp cDNA fragment isolated from clone SF2.

Southern hybridizations. Plasmid and genomic DNAs were digested with various restriction enzymes; fragments were separated on 0.7% agarose gels, denatured, and transferred to nitrocellulose or HyBond membranes. Hybridization was carried out as indicated above for genomic library screening.

DNA sequencing. Subfragments of the genomic clone HA200 were cloned into either pUC19 or M13mp19 vectors. Double-, and single-stranded DNAs respectively were sequenced according to Sanger et al. (1977), using synthetic oligonucleotides as primers. For sequencing of double-stranded DNA, the denaturation and annealing conditions described by Zhang et al. (1988) were used.

Primer extension. Primer extension was carried out as described (Gualberto et al. 1988). The oligonucleotide

used was a 24-mer (GATATTGCAAGAACGAAGACGAGC) complementary to the 3′ region of the first exon sequence.

Northern hybridizations. Total floral RNA (Herdenberger et al. 1990) was separated on (1%) agarose–(6%) formaldehyde gels, transferred to Hybond membranes and hybridized as described above for genomic library screening.

Results

Isolation and mapping of the anther-specific gene SF2

Out of 2×10^6 plaques of the genomic library screened by plaque hybridization using the radiolabelled SF2 cDNA as a probe, approximately 80 positive clones were detected. The clone selected for the studies described here contained a 12.5 kb genomic DNA fragment and was designated HA200. For physical mapping of the insert, single and multiple digestions were performed on the recombinant phage HA200 using the restriction enzymes *Eco*RI, *Hin*dIII, *Kpn*I, *Xba*I and *Xho*I. By Southern hybridizations using the radioactively labelled SF2 cDNA as probe the coding region was located in the middle of the 12.5 kb insert (Fig. 1).

Structure of the SF2 *gene*

Figure 2 shows the nucleotide sequence of a 3739 bp segment containing the coding region delimited by a *Pst*I and the second *Xba*I site (see Fig. 1). We identified the 3′ end of the gene in the 0.5 kb *Xba*I fragment by comparison with the cDNA sequence. This small *Xba*I fragment, however, contained only part of the gene, more precisely the last 178 nucleotides of the untranslated 3′ end. On the adjacent large *Xba*I fragment the cDNA sequence matched the genomic sequence over an additional 386 bp, but the 5′ end of the coding sequence (77 bp) could only be identified 2615 bp further upstream. These data show that the *SF2* gene consists of two exonic sequences separated by a 2.6 kb intron. Curiously the first exon stops precisely after the sequence coding for a putative signal peptide. The second exon is 563 bp long and continues the reading frame for an additional 100 amino acids.

Determination of the transcription start site by primer extension (Fig. 3) shows a major transcript initiated at an A residue one nucleotide before the 5′ end of the SF2 cDNA. Three weaker bands, also visible on the autoradiograph, could correspond to minor transcription start sites. The cDNA sequence of the SF2 clone can therefore be considered close to full-length (only the first nucleotide is missing). The length from the transcription start site to the polyadenylation site is 639 nucleotides (the SF2 cDNA has a 1 bp deletion in the 3′ untranslated region with respect to the genomic sequence). The difference in length observed between the cDNA and the mRNA (estimated at 800 nucleotides) can be accounted for by a polyA tail of 150–200 nucleotides in the mRNA. The 5′ untranslated leader sequence is surprisingly short (13 bp for the major transcript).

Features of the intron

The intron is 2615 bp long, with splice junctions obeying the GT...AG rule of eukaryotic mRNA introns (Breathnach and Chambon 1981). With this large size it is considerably longer than most plant introns so far described, the mean length of which is estimated to be about 250 bp (Hawkins 1988). It is AT-rich (67%) as expected for a plant intron (Goodall and Filipowicz 1989), and contains occasional GC-rich segments (the most important one being a 44 bp stretch with 77% G+C content; see Fig. 2) as well as repeated palindromic sequences (the undecanucleotide TACGCA/TGCGTA is repeated seven times within 620 bp; underlined in Fig. 2).

A computer study (Zuker and Stiegler 1981) suggests that this intron is very highly structured. Such a structure could be required to guide the binding of splicing factors to the correct splice sites by occluding all other potential 3′ splice sites (more than 60 can be detected in the intron) in a highly organized secondary and tertiary structure. The presence of potential 3′ splice sites in the intron raises the possiblity that additional reading frames in the intron could serve as second exons to produce, by alternative splicing, other polypeptides carrying the same signal peptide and showing the same expression pattern as SF2. Although the intron contains three significant open reading frames (close to 100 amino acids in length) we were unable to detect their sequences in floral mRNAs by Northern hybridization using reading frame-specific DNA probes (data not shown).

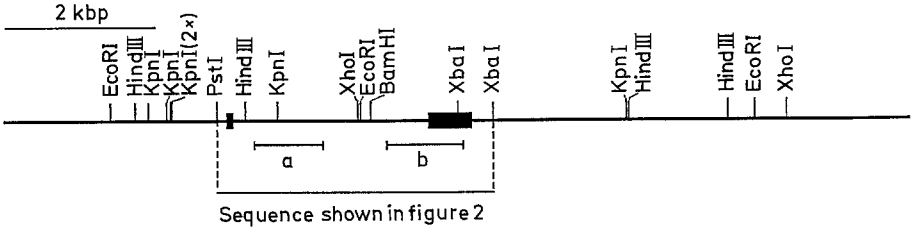

Fig. 1. Restriction map and location of the *SF2* gene on the 12.5 kb insert. Exons are shown as *heavy bars*. Transcription is from left to right. The nucleotide sequence spanning the region between the *Pst*I site and the second *Xba*I site (indicated by the longer *bar*) is shown in Fig. 2. Bars marked *a* and *b* correspond to DNA probes used in the Southern hybridizations; see Fig. 4)

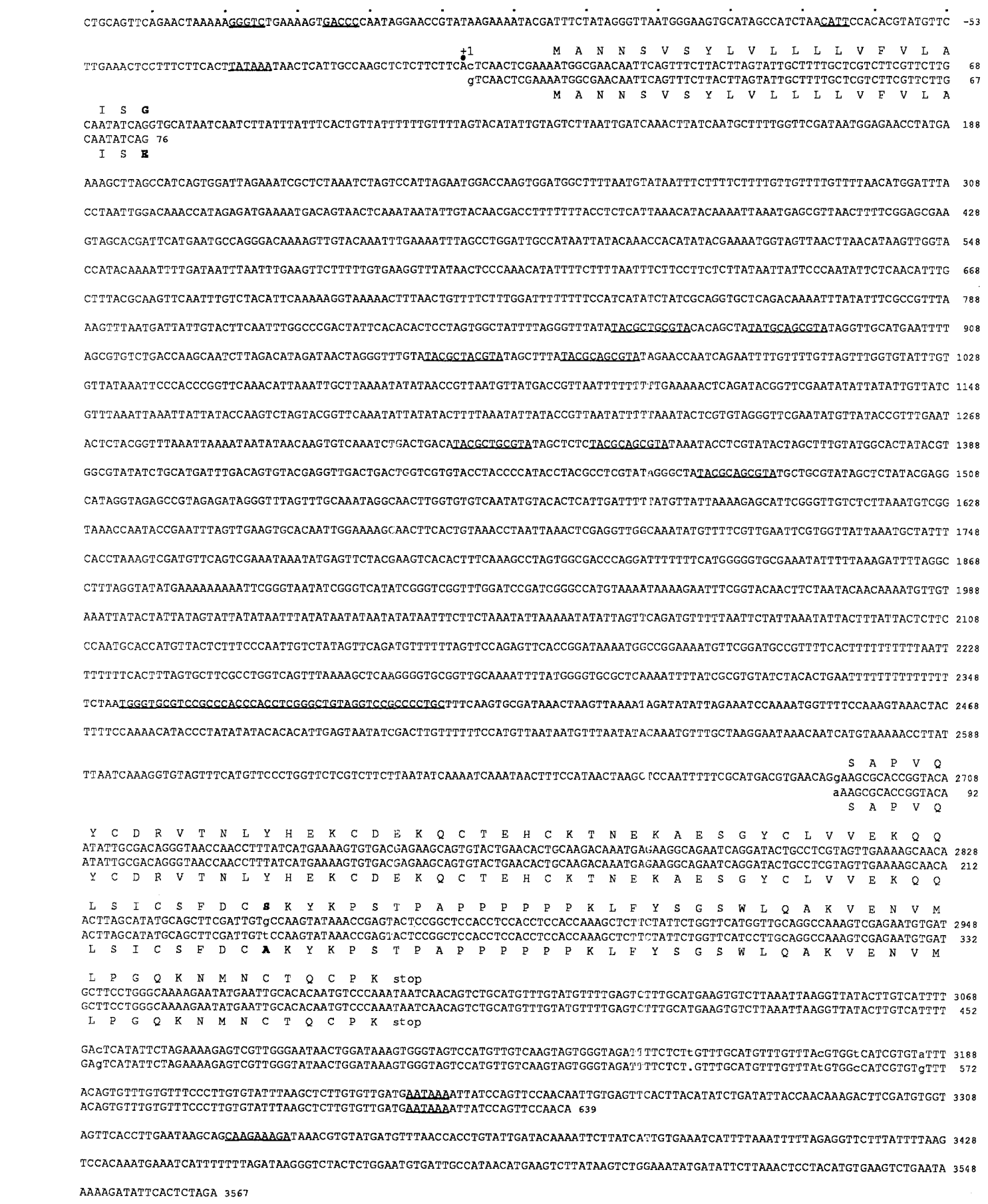

Fig. 2. Nucleotide sequence of the *SF2* gene. The cDNA sequence of *SF2* is aligned with the genomic sequence. Also shown are the peptide sequences derived from the genomic sequence (*upper*) and the cDNA sequence (*lower*). Amino acids differing in the genomic and cDNA sequences are indicated in *boldtype* letters. See the text for comments on underlined sequences

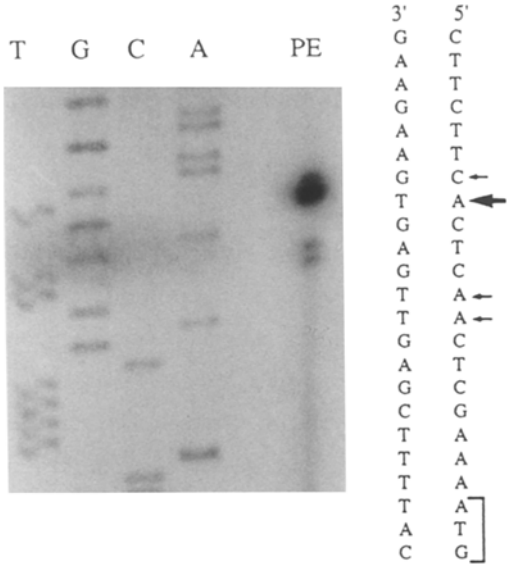

Fig. 3. Determination of the transcription start site by primer extension. The nucleotide sequence around the transcription start site is shown. The initiation codon ATG is also indicated. PE, products of primer extension. The *arrows* indicate the positions of the major and minor start sites

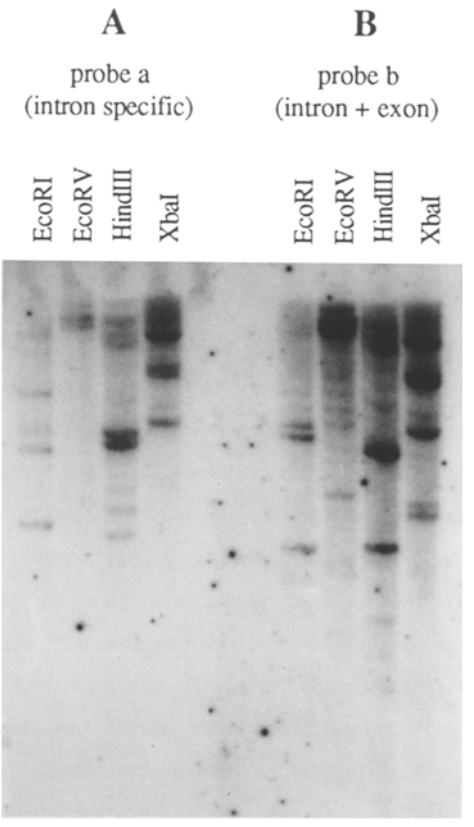

Fig. 4A and B. Southern hybridizations on genomic blots carrying *Eco*RI, *Eco*RV, *Hin*dIII and *Xba*I restricted DNA fragments. **A** DNA probe *a* (see Fig. 1), containing an intron-specific sequence (nucleotides 415 to 1334 in Fig. 2). **B** DNA probe *b* (see Fig. 1), containing intron as well as exon 2 sequences (nucleotides 2130 to 3132 in Fig. 2). The two probes were obtained by PCR amplification

Features of the flanking regions

Figure 2 shows the DNA sequence of the gene including 167 bp of the upstream region (up to the *Pst*I site) and 311 bp downstream of the polyadenylation site. A TATA box (TATAAA) is found around position −30 and a "CAT-like" box (CCAT) around position −75. This region also contains a short inverted repeat (GGGTC..8 bp..GACCC). No experiment has yet been performed to see whether these or other motifs act as binding sites for transcription factors.

As mRNAs are processed immediately following their transcription in the nucleus, by removal of sequences downstream of the polyadenylation site, their transcription termination sites are difficult to determine; hence transcription termination signals are difficult to identify. However, signals involved in 3′ processing have been identified in various organisms; they include sequences in the untranslated trailer (such as the polyadenylation signal AATAAA), as well as sequences downstream of the polyadenylation signal which are removed during the processing step (Birnstiel et al. 1985). In sea urchin histone genes (whose mRNAs are not polyadenylated) a conserved downstream sequence (CAAGAAAGA) is an essential element for the formation of 3′ termini of the H3 histone mRNA by RNA processing (Georgiev and Birnstiel 1985). The same nonanucleotide sequence is found 74 nucleotides downstream of the polyadenylation site of the *SF2* gene and could be involved in processing of the SF2 mRNA. Another striking feature is the presence of many short GT-rich sequences in the untranslated 3′ region; some of these sequences could serve as polyadenylation signals (Proudfoot 1991), or they could be involved in the stability of the SF2 mRNA.

Gene copy number

The comparison of the genomic DNA sequence with that of the SF2 cDNA suggests that several copies of the *SF2* gene must be present in the sunflower nuclear genome. In fact the two sequences differ by 9 nucleotides, of which two lead to an amino acid change in the polypeptide (indicated by boldtype in Fig. 2). The nucleotide sequence of another anther-specific cDNA clone (SF1) differs by only a few nucleotides from that of the SF2 cDNA as well as that of the genomic DNA (data not shown; see comparison of peptide sequences in Evrard et al. 1991); most noticeable in the SF1 protein is the presence of two additional prolines in the cluster of six proline residues. This brings to at least 3 the number of genomic copies of this anther-specific gene.

The multicopy nature of the *SF2* gene has been confirmed by Southern hybridization on genomic blots carrying *Eco*RI, *Eco*RV, *Hin*dIII and *Xba*I digested DNA (Fig. 4). Several bands were detected with a DNA probe containing the second exon and part of the intron (probe *b*; see Fig. 2), as well as with an intron-specific probe (probe *a*: see Fig. 2), suggesting that most, if not all, of the copies of this gene are split and contain a highly related intron.

Discussion

The *SF2* gene is characterized by three unusual features: (i) it is an anther-specific gene whose expression is restricted to peripheral anther cells (Evrard et al. 1991); (ii) the gene is interrupted by an intron of an unusual length; and (iii) the two exons code for very specific protein domains, one being a signal peptide, the other one corresponding to the mature protein. There is, to date, little information relating to the mechanism by which the expression of anther-specific genes is regulated. Although several of these genes have been sequenced, no common nucleotide pattern has been found in their upstream region. This is probably due in part to the different nature of the cell types in which the anther-specific mRNAs have been detected.

The SF2 mRNA is an abundant mRNA, as estimated from Northern blot experiments. This observation is reflected by the abundance of SF2-related sequences in the cDNA library: the SF2 cDNA cross-hybridizes with three other cDNA clones (SF1, SF6 and SF7) isolated from a total of 30 flower-specific clones. Sequence comparisons have shown that the two cDNAs (SF2 and SF1) and the genomic clone are not identical, but that the sequence divergence is very low (less than 5%), in the translated as well as untranslated regions. This indicates that the various copies of this new gene family arose by recent duplications of a single copy gene which, according to the data obtained by Southern hybridization, was already split.

In a survey comparing intron and exon lengths in higher plant, fungal, insect and vertebrate genes (Hawkins 1988), noticeable differences were observed between these groups of organisms. Fungi were found to have the smallest introns, with a mean length of 86 bp (for over 120 introns considered) whereas the mean length of plant and insect introns was 249 and 622 bp respectively (for about 200 introns considered in each of the two groups). Vertebrates seem to have the largest introns with a mean length of 1127 nucleotides (for 1941 introns considered). The figures given here for insects and vertebrates apply to internal introns; 5′ external introns are usually much larger, especially in the case of insects where a mean size of more than 5 kb was found. In more recent years, with additional genes having been studied, this intron length pattern has been confirmed: in plants, most of the introns described are smaller than 300 bp. Occasionally, introns of up to 600 bp are found, as for example in the soybean glycinin genes (Nielsen et al. 1989). Plant introns of several kb are rare exceptions; as in insects, they appear to be confined to the 5′ external introns. This is the case for the rice phytochrome gene where a 2585 bp intron was observed approximately 100 bp upstream of the initiation codon (Kay et al. 1989). This gene also contains a 747 bp 3′ external intron, suggesting that long introns are rather restricted to untranslated regions of the genes.

Hence, the 2615 bp intron found in the *SF2* gene is unique in its length and position: as one of the longest plant introns described so far it separates the sequences coding for two very distinct protein domains. The first domain, a putative signal peptide, characterized by a short sequence of hydrophobic amino acids (see Fig. 2B in Domon et al. 1990), would direct the growing polypeptide chain to the rough endoplasmic reticulum and target the protein to the extracellular space. Many signal peptide sequences have been described to date in various organisms but, apart from their hydrophobic nature, no consensus sequence could be established (von Heijne 1983). Interestingly, we found evidence for the presence of a very similar signal peptide sequence at the amino-terminus of another anther-specific protein (SF18) during the study of its cDNA (Domon et al. 1990). Three different mechanisms could account for the presence of the same signal peptide-coding sequence in the SF2 and SF18 mRNAs: (i) *trans*-splicing of two independent transcripts; (ii) alternative splicing from a common precursor mRNA; (iii) exon transfer at the gene level (exon shuffling). In the light of what is known about the structure of the *SF2* gene, the latter situation seems the most likely; to confirm this hypothesis an SF18 genomic clone is presently being studied.

Acknowledgements. This work was supported in part by grant A-1984 (NSERC) to D.T.N.P. and by a Competitive Research Grant from Pioneer Hi-Bred International. C.D. was supported by a "Bourse Boston" from the city of Strasbourg. We are grateful to A. Hoeft for the synthesis of oligonucleotides.

References

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

Birnboim HC, Doly J (1979) A rapid extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Birnstiel M, Busslinger M, Strub K (1985) Transcription termination and 3′ processing: the end is in site! Cell 41:349–359

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

Brown SM, Crouch ML (1990) Characterization of a gene family abundantly expressed in *Oenothera organensis* pollen that shows sequence similarity to polygalacturonidase. Plant Cell 2:263–274

Davis RW, Botstein D, Roth JR (1980) Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: Version 2. Plant Mol Biol Reporter 1:19–21

Domon C, Evrard JL, Herdenberger F, Pillay DTN, Steinmetz A (1990) Nucleotide sequence of two anther-specific cDNAs from sunflower (*Helianthus annuus* L). Plant Mol Biol 15:643–646

Dunn IS, Blattner F (1987) Charons 36 to 40: multi enzyme, high capacity, recombination deficient vectors with polylinkers and polystuffers. Nucleic Acids Res 15:2677–2698

Evrard JL, Jako C, Saint-Guily A, Weil JH, Kuntz M (1991) Anther-specific, developmentally regulated expression of genes encoding a new class of proline-rich proteins in sunflower. Plant Mol Biol 16:271–281

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

Georgiev O, Birnstiel ML (1985) The conserved CAAGAAAGA spacer sequence is an essential element for the formation of

3′ termini of the sea urchin H3 histone mRNA by RNA processing. EMBO J 4:481–489
Goodall GJ, Filipowicz W (1989) The AU-rich squences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell 58:473–483
Gualberto JM, Wintz H, Weil JH, Grienenberger JM (1988) The genes coding for subunit 3 of NADH dehydrogenase and for ribosomal protein S12 are present in the wheat and maize mitochondrial genomes and are cotranscribed. Mol Gen Genet 215:118–127
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
Hawkins JD (1988) A survey on intron and exon lengths. Nucleic Acids Res 16:9893–9908
Herdenberger F, Evrard JL, Kuntz M, Tessier LH, Klein A, Steinmetz A, Pillay DTN (1990) Isolation of flower-specific cDNA clones from sunflower (*Helianthus annuus* L.) Plant Sci 69:111–122
Kaiser K, Murray N (1985) In: Glover DM (ed) DNA cloning, vol 1. IRL Press, Oxford, pp 1–47
Kamalay JC, Goldberg RB (1980) Regulation of structural gene expression in tobacco. Cell 19:935–946
Kamalay JC, Goldberg RB (1984) Organ-specific nuclear RNAs in tobacco. Proc Natl Acad Sci USA 81:2801–2805
Kay SA, Keith B, Shinozaki K, Chye ML, Chua NH (1989) The rice phytochrome gene: structure, autoregulated expression, and binding of GT-1 to a conserved site in the 5′ upstream region. Plant Cell 1:351–360
Koltunow AM, Truettner J, Cox C, Wallroth M, Goldberg RB (1990) Different temporal and spatial gene expression patterns occur during anther development. Plant Cell 2:1201–1224
Mascarenhas JP (1990) Gene activity during pollen development. Annu Rev Plant Physiol 41:317–338
McCormick S, Twell D, Wing R, Ursin V, Yamagushi J, Larabell S (1989) Anther-specific genes: molecular characterization and promoter analysis in transgenic plants. In: Lord E, Bernier G (eds) Plant reproduction: From floral induction to pollination. American Society of Plant Physiologists Symposium Series, vol 1, pp 128–135
Nielsen NC, Dickinson CD, Cho T-J, Thanh VH, Scallon BJ, Fischer RL, Sims TL, Drews GN, Goldberg RB (1989) Characterization of the glycinin gene family in soybean. Plant Cell 1:313–328
Proudfoot N (1991) Poly(A) signals. Cell 64:671–674
Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467
Seurinck J, Truettner J, Goldberg RB (1990) The nucleotide sequence of an anther-specific gene. Nucleic Acids Res 18:3403
Smith AG, Gasser CS, Budelier KA, Fraley RT (1990) Identification and characterization of stamen and tapetum-specific genes from tobacco. Mol Gen Genet 222:9–16
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
Twell D, Wing R, Yamagushi J, McCormick S (1989) Isolation and expression of an anther-specific gene from tomato. Mol Gen Genet 247:240–245
Ursin VM, Yamagushi J, McCormick S (1989) Gametophytic and sporophytic expression of anther-specific genes in developing tomato anthers. Plant Cell 1:727–736
van Tunen AJ, Mur LA, Brouns GS, Rienstra JD, Koes RE, Mol JNM (1990) Pollen and anther-specific promoters from *Petunia*: tandem promoter regulation of the *chi*A gene. Plant Cell 2:393–401
von Heijne G (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 133:17–21
Zhang H, Scholl R, Browse R, Somerville C (1988) Double stranded DNA sequencing as a choice for DNA sequencing. Nucleic Acids Res 16:1220
Zuker M, Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleid Acids Res 9:133–148

Communicated by J. Schell