

Characterization of a family of genes encoding a fruit-specific wound-stimulated protein of bell pepper (*Capsicum annuum*): identification of a new family of transposable elements

Javier Pozueta-Romero, Mathieu Klein¹, Guy Houlné, Marie-Luce Schantz, Beatrice Meyer and Rodolphe Schantz*

*Institut de Biologie Moléculaire des Plantes, Université Louis Pasteur, 12 Rue du Général Zimmer, 67084 Strasbourg Cedex, France (*author for correspondence); ¹Current address: Institut für Genbiologische Forschung, Ihnestrasse 63, 1000 Berlin 33, Germany*

Received 9 February 1995; accepted in revised form 30 May 1995

Key words: *Alien* transposable element, *Capsicum annuum*, fruit-specific expression, genomic organization, terminal inverted repeat, wound induction

Abstract

Using a fruit-specific cDNA as a probe we isolated and sequenced the two corresponding homologous genes (*Sn-1* and *Sn-2*) of the bell pepper (*Capsicum annuum*) genome. Both genes have a single intron and numerous unusual long inverted repeat sequences. The introns share 87% homology and *Sn-2* contains one 450 bp additional sequence with structural features of a transposable element, which is highly repetitive in the bell pepper genome. Surprisingly, analysis in data banks showed that genes encoding the potato starch phosphorylase (EC 2.4.1.1) and patatin contain a similar element, named *Alien*, in their 5'-upstream region. *Alien* elements are characterized by a conserved 28 bp terminal inverted repeat (TIR), small size, high AT content, potential to form stable DNA secondary structures and they have probably been inserted in TA target sites. Interestingly, the TIR of the *Alien* elements shares high homology with sequences existing in the TIR of extrachromosomal linear pSKL DNA plasmid of *Saccharomyces kluyveri*. Northern blot analyses detected *Sn-1* transcripts principally in the red fruit whereas no *Sn-2* transcripts were detected in neither of the samples monitored. Western blot analyses detected a 16.8 kDa Sn protein principally in the ripe red fruit and wounded areas of green unripe fruit. A comparison of the deduced amino acid sequence of Sn-1 with protein sequences in data banks revealed a significant homology with proteins likely involved in the plant's disease resistance response. Analyses at the subcellular level showed that Sn-1 is localized in the membrane of vacuoles.

Introduction

Fruit ripening is a genetically programmed process which involves increasing rates of respiration

and ethylene production and shifts in metabolism of carbohydrates, organic acids, proteins, etc. [15]. In many fruits, chloroplasts differentiate into non-photosynthetic chromoplasts with a subse-

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X79230 (*Sn-1*) and 79231 (*Sn-2*).

quent degradation of chlorophylls and starch and extensive structural reorganization [8].

Climacteric and non-climacteric fruits differ in their ability to produce ethylene autocatalytically in response to threshold levels of ethylene. Commonly referred to as the ripening hormone, ethylene has a cascade effect in climacteric fruits leading to 'one rotten apple spoils the whole bushel' syndrome [40]. The role of ethylene in the ripening of climacteric fruits such as tomato has been demonstrated unequivocally by inhibiting endogenous ethylene production through the expression of antisense RNAs that block the accumulation of 1-aminocyclopropane-1-carboxylate synthase [36]. In sharp contrast, bell pepper, which is also a member of the Solanaceae, has been classified as a non-climacteric fruit on the basis of the small changes of carbon dioxide and ethylene production upon wound stress and during fruit ripening [30]. This particular behavior of bell pepper has prompted us to investigate the mechanisms involved in its fruit development and ripening at the molecular level as a possible clue to explain the differences between climacteric and non-climacteric fruits. Towards this end we have characterized the two genes of a family encoding a fruit-specific protein and we discuss the possible correlation between their structural features and their respective expression pattern. We report for the first time the existence in the bell pepper genome of small highly repetitive sequences with structural characteristics of transposable elements.

Materials and methods

Plant material

Bell pepper seeds (cvs. Lamuyo and Yolo Wonder) were germinated, and plants were grown under greenhouse conditions. For wounding of fruits, cuts of 3 cm length were carried out in the pericarp tissue using a sterilized razor. Samples of 1 cm² were collected every 15 h from the wounded areas.

Isolation of a ripe fruit-specific cDNA

A λ gt11 cDNA library of ripening fruit (c.v. Lamuyo) was differentially screened with labelled first-strand cDNA synthesized from poly(A)⁺ RNA of ripe fruits and leaves [14]. A fruit-specific cDNA of 400 bp was then selected. Preliminary northern blot analysis using the labelled [13] cDNA probe showed the existence of a transcript of ca. 700 bp. Therefore we performed a second screening and a 643 bp cDNA was isolated and cloned into the *Eco* RI site of pBluescript KS + (Stratagene).

Protein expression

After determination of the coding strand by single-strand hybridization one open reading frame of 440 bp was identified. The 643 bp cDNA was then subcloned into the *Bam* HI/*Sal* I site of the pGEX-KG expression vector [19] in frame in order to produce the protein as a glutathione *S*-transferase fusion protein. Induction of the fusion protein in *Escherichia coli* was obtained by addition of isopropyl-thio- β -D-galactoside (IPTG) 0.25 mM. Bacteria were lysed by sonication and the proteins were partially purified by SDS gel electrophoresis and subsequent electroelution. Antibodies were then raised by injection of the proteins into rabbits.

Construction of the genomic library

Lambda Gem-11 (Promega) *Bam* HI arms were prepared as described by Kaiser and Murray [23]. Genomic DNA of bell pepper cv. Yolo Wonder was isolated and purified according to the method described by Draper and Scott [11]. After partial digestion with *Sau* 3AI, restriction fragments of 15–20 kbp were purified on NaCl gradient, ligated into the lambda arms and packaged using Giga-pack II Gold packaging extracts (Stratagene) following the manufacturer's instructions. The resulting library contained 1.8×10^6 recombinants, representing around ten bell pepper genomes which were used to transfect ER1647 bacteria.

Screening of the genomic library

Library plates bacteria grown overnight at 37 °C were transferred in duplicate to 0.45 µm nitrocellulose membranes (Schleicher & Schuell). The filters were subsequently hybridized in 5 × SSC (1 × SSC is 0.15 M NaCl, 0.015 M sodium-citrate), 5 × Denhardt's solution (1 mg/ml Ficoll, 1 mg/ml polyvinyl-pyrrolidone, 1 mg/ml BSA), 1% SDS, 20 µg/ml salmon sperm DNA and 50% formamide for 12 h at 42 °C with the ³²P-labelled 643 bp cDNA. Filters were first washed at 65 °C in 1 × SSC, 0.1% SDS and then in 0.1 × SSC, 0.1% SDS prior to the autoradiography. Positive genomic clones were selected after three rounds of screening.

DNA sequencing

Subfragments of the genomic clones were cloned into the pBluescript KS[−] vector (Stratagene). *Sure* (rec A[−]) (Stratagene) bacteria were used as the recipient strain. Subsequently, plasmid DNA was isolated according to the rapid DNA preparation of Saunders and Burke [37] and used for double strand sequencing [47] as well as for isolation of DNA inserts. Pairwise and multiple alignments of sequences were performed using the BESTFIT and PILEUP (gap penalty of 5.00 and gap length of 0.30) programs, respectively. Minimum energy folding of sequences was performed using the program FOLD.

Northern hybridization

Total RNAs from roots, shoots, leaves and cotyledons, flowers and fruits were extracted according to Bogorad *et al.* [1]. Subsequently, 10 µg of RNA were separated on 1.1% agarose-6% formaldehyde gels, transferred to Hybond-N⁺ membranes (Amersham) and hybridized as described above.

S1 mapping

For determination of the transcription start and polyadenylation sites, S1 mapping was carried out as described [18].

Southern hybridizations

Plasmid and genomic DNAs were digested with various restriction enzymes. Fragments were separated on 0.7% agarose gels, denatured, and transferred to Hybond-N⁺ membranes. Hybridization was carried out as indicated above for genomic library screening.

Immunochemical techniques

Twenty µg of proteins from different plant tissues, extracted by the phenol method [31], but using 0.1 M LiCl, 1% SDS, 10 mM EDTA and 0.1 M Tris-HCl pH 8 as extracting buffer were separated on SDS-PAGE and transferred to nitrocellulose filters for western blot analyses [42].

Immunogold labelling and electron microscopy

Protoplasts from fresh pericarp tissue of fruits at the orange stage were obtained according to the procedure of Veidt *et al.* [46] but 0.1 mg/ml caryolase M3 (Cayla Laboratories, Toulouse, France) was added in the enzyme solution and digestion took place overnight at 28 °C without shaking. Fixation, permeabilization and subsequent immunolabelling of cell preparations were performed as described [38], with an additional step of silver-enhancement using the IntenSE BL Kit (Amersham) before embedding in Epon. Sections were made using a Ultracut E microtome (Reichert) and observed on a Hitachi H600 transmission electron microscope.

Results

Structure and organization of *Sn* genes

Labelled 643 bp *Sn* cDNA (see Materials and methods) was hybridized with restricted pepper genomic DNA (Fig. 1a). The pattern of the hybridizing genomic restriction fragments suggested the existence in the pepper genome of more than a single gene homologous to the *Sn* cDNA. The restriction maps obtained from the different ge-

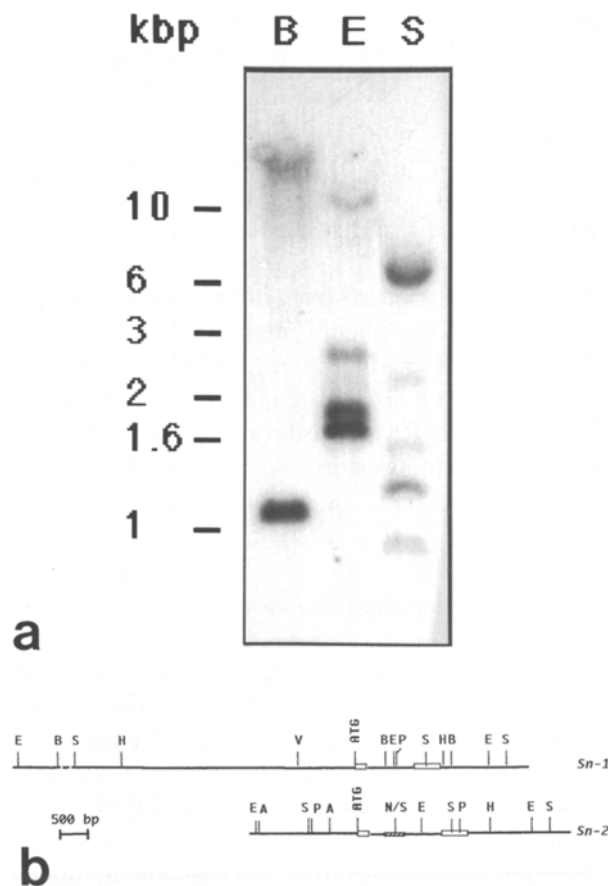


Fig. 1. Structure of *Sn* genes. a. Southern hybridizations of 10 μ g genomic DNA restricted by *Eco* RI (E) *Sst* I (S) and *Bam* HI (B). The blot was probed with ^{32} P-labelled 643 bp *Sn* cDNA. b. Restriction endonuclease site maps of *Sn-1* and *Sn-2* from genomic clones. Translated sequences and the 450 bp intronic additional sequence are shown by open and dashed boxes, respectively. Restriction sites are indicated as following: H, *Hind* III; V, *Eco* RV; P, *Spe* I; N, *Nco* I; A, *Sna* BI.

nomics clones isolated from the library of pepper nuclear DNA confirms the existence of two *Sn* genes, *Sn-1* and *Sn-2* (Fig. 1b). To investigate the structure of these genes, we determined their respective nucleotide sequences and compared them with that of the cDNA (Fig. 2). Importantly, the 120 bp sequence upstream of the translation start codons from *Sn-1* and *Sn-2* genes share 91% homology. The coding sequences of either *Sn-1* and *Sn-2* have a relatively high AT content (ca. 63%) and although they share high homology with the cDNA (100% and 94.5%, respectively) important differences are found between positions 160 and 220 bp downstream the TAA termination codon. A similar case of variability of the 3'-untranslated region between a family of genes has been reported elsewhere [4]. The variation between the genomic and cDNA clones is ascribable to the utilization of different cultivars for generating their respective libraries.

Both *Sn-1* and *Sn-2* contain one intron inserted at 188 bp downstream the translation start codon. The sizes of the introns are 739 bp and 1130 bp for *Sn-1* and *Sn-2* respectively. Their sequences share 87% homology and the *Sn-2* intron has an additional sequence of 450 bp (position 584 to 1034) with a 17 bp tandemly repeated sequence (CCATTAGGGGTTGTTT) at its origin (position 584 to 618). Introns from both genes have splice junctions obeying the GT....AG rule of eukaryotic mRNA introns [3]. The high AT content (69% and 72% for *Sn-1* and *Sn-2*, respectively) is consistent with that expected for a plant intron [16].

An important feature of the structure of both *Sn* genes is the existence of numerous long inverted repeat sequences which are found throughout the entire gene structure. These sequences are predicted to potentially participate in secondary and tertiary interactions forming stable hairpins in the pre-mRNA. Particularly, sequences located in the additional sequence of the *Sn-2* intron (nucleotide positions 606 to 1034) and in the 3' end of *Sn-2* (nucleotide position 2093 to 2466) are predicted to form stable secondary structures with calculated free energies of -412 kJ/mol and -312 kJ/mol, respectively.

Fig. 2.

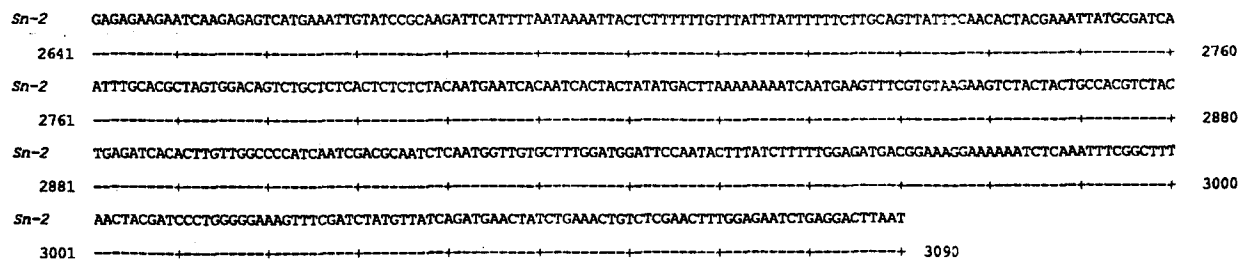


Fig. 2. Nucleotide sequences and predicted amino acid sequence of genomic *Sn* clones and *Sn* cDNA. Only different nucleotides of *Sn-2* and the cDNA towards *Sn-1* are presented. The cDNA sequence begins and ends at the nucleotide positions 131 and 1951, respectively. Deletions in the nucleotide sequence of either *Sn-1* or *Sn-2* are represented by dots. TIRs of the two transposable elements of *Sn-2* are indicated by open arrows and their respective flanking direct repeats are indicated by bold arrows. The ATG translation start and the termination TAA codons are located at positions 139 and 1743, respectively. The sequence upstream the translation initiation codon is consistent with Kozak's consensus [24]. Putative AATAAA or AATAAT consensus polyadenylation signals [22] are found at positions 1816, 1862 (for either *Sn* cDNA, *Sn-1* and *Sn-2*) and 1931 (only for the *Sn* cDNA). Repeated sequences in the 3'-untranslated region, nucleotide positions 1793 to 1811, 1837 to 1850 and 1887 to 1903, are underlined.

Identification of a new family of transposable elements

Labelled DNA fragments from the additional intronic 450 bp and 3' region flanking the termination codon which have long inverted repeats were hybridized with digested genomic DNA under high-stringency conditions (Fig. 3a, b). Results showing the cross-hybridization of the probes with many fragments of genomic DNA demonstrate the existence in the bell pepper genome of multiple and highly conserved copies of sequences existing in the additional 450 intronic bp and also in the 3' end of *Sn-2*. Interestingly, these highly repetitive sequences have the structural features of unusual inverted repeat transposable elements. In the case of the 450 bp additional intronic sequence of *Sn-2*, a 428 bp DNA fragment with potential to form stable hairpin structures has a 28 bp imperfect terminal inverted repeat at its termini (TIR, position 608 to 636 and 1001 to 1032) (see Fig. 2) flanked by TA direct repeats; the 3' end of *Sn-2* contains an AT-rich (60%) 374 bp DNA fragment (position 2093 to 2466) with a 16 bp TIR flanked by AT direct repeats. To determine if these sequences are members of larger families, computer-assisted sequence similarity searches of the GenBank and EMBL nucleic acid databases were performed using the two inverted

repeat sequences as queries. To our surprise, this examination revealed that the 5'-flanking regions of the genes coding for the potato starch phosphorylase (position -1022 to -303) [9] and patatin [43] contain a small previously unreported element with structural and sequence similarities to the *Sn-2* 428 bp transposable-like element (Fig. 4a). These elements have been designated as *Alien* family members. The *Alien* element inserted in the 5'-flanking region of the patatin gene is truncated and it lacks the 5'-TIR. *Aliens* have a potential to form DNA secondary structures ($\Delta G = -412$ KJ/mol and -832 KJ/mol for the *Sn-2* and starch phosphorylase *Alien* elements, respectively), they are AT-rich (74% and 76% for the *Sn-2* and starch phosphorylase *Alien* elements, respectively), they have small size (428 bp and 719 bp for *Sn-2* and starch phosphorylase *Alien* elements, respectively), they have TIRs (Fig. 4b), and they have probably been inserted in TA target sites. Most importantly, the sequence searches also revealed that the 28 bp TIR consensus sequence of the *Alien* elements is highly homologous to a 27 bp sequence repeated eight times within the 483 bp TIR of the extrachromosomal linear DNA pSKL plasmid of *Saccharomyces kluyveri* (nucleotide position 253 to 280 and 14003 to 14030 of the pSKL plasmid) [21] (Fig. 4a). On the other hand, analysis in data banks did not

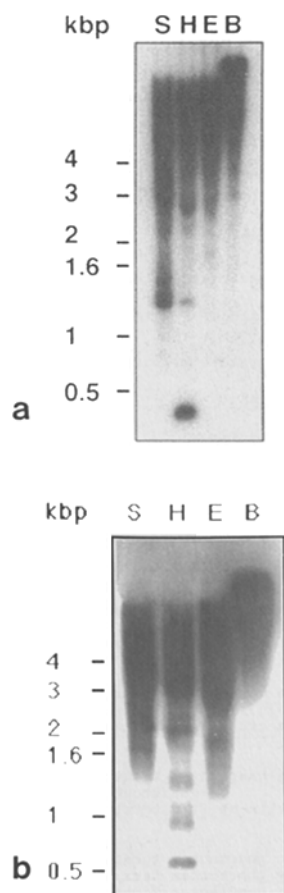


Fig. 3. Autoradiography of a bell pepper genomic DNA blot hybridized under standard conditions with a) a probe specific of *Sn-2* intronic sequence (nucleotide position 521 to 777) and b) the *SpeI/HindIII* fragment containing the 3' termination codon of *Sn-2* (nucleotide position 1838–2301). Ten μ g of DNA was restricted with *StyI* (S), *HindIII* (H), *EcoRI* (E) or *BamHI* (B) and separated by agarose gel electrophoresis.

reveal a significant homology between the transposable element of the 3' end of *Sn-2* and any nucleic acid sequence.

Expression of the *Sn* genes

Northern blot analysis using the 643 bp cDNA as probe shows that among the different parts of the plant monitored, a ca. *Sn* 700 b transcript is detectable exclusively in flowers and fruits (Fig. 5a). The relative amount of the transcript observed in flowers remains low and constant during the early

fruit development and increases dramatically during ripening from mature green to red fruit. Subsequent blots using the *StyI/BamHI* fragment specific from *Sn-1* show an identical pattern of expression to that observed using the cDNA probe (Fig. 5b). The transcription start and polyadenylation sites of *Sn-1* determined by S1-nuclease protection experiments are located at positions 103 and 1972, respectively and the exact size of the *Sn-1* major transcript is 692 b (Fig. 6a, b). Results showing the presence of one alternative transcription polyadenylation site at position 2138 (see Fig. 6b) indicates the existence of a minor transcript of nearly 858 b. On the other hand, S1-nuclease protection experiments using *Sn-2*-specific probes did not detect any transcript in neither of the five samples monitored (data not shown).

The localization of the transcription start and polyadenylation sites of *Sn-1* permitted to us to characterize their respective flanking regions; TATA and consensus sequences [20] are present 37 and 61 bp, respectively, upstream to the transcription start site and the high AT content of the 3'-untranslated region (73%) is consistent with previous reports arguing for the important role of the high AT content in this region as a mRNA-destabilizing factor [41].

Predicted amino acid sequence and comparison with other proteins

The predicted polypeptide from the nucleotide sequence of the *Sn* cDNA (see Fig. 2) has 147 amino acid residues with a size of 16.8 kDa, a global charge of -8 and an isoelectric point of 5.8. It has a putative N-glycosylation site at the level of aminoacids 50 to 52. The amino acid sequences deduced from the coding sequence of *Sn-1* and *Sn-2* share 100% and 90% identity with respect to that of the cDNA. A hydropathy plot constructed for the predicted polypeptide of the cDNA shows that the N-terminal region of the peptide does not have any significant structural homology with transit peptide thus totally excluding a plastidial localization (Fig. 7a).

a

1 taggggttgt ttggtagagt gtattaagaa atataatgca tatattaggt ...gtgtgta ttagtagtac 70
 Phospho ..ggggata. ttggtagagt gtattgaaaa \$
 Patatin taggggttgt ttggtagagt gtattgaaaa at.taatgca tgtattagtt taagtgtgta ttactaatac
 Sn-2

71 cttgtttggc acactttttc atgccatgta taactaatgc atgtgtatta ctaataccaa ggaattctag 140
 Phospho cttatttggc atacttttgc accctattta taactaatgc
 Patatin
 Sn-2

141 gtattagtaa taaatagcat tttaacactt gcattagatc aaataattac aaaactaccc ttaagcatt 210
 Phospho
 Patatin
 Sn-2

211 ttcattttct ttgtgtcat aagtttttat ttttattttt atttgctttt cggtatcttt taatttggtg 280
 Phospho
 Patatin
 Sn-2

281 gtgtcttaat agactttatg gccttttaag tatcttttta aaaaaaatct aatttcaata taattttaa 350
 Phospho
 Patatin
 Sn-2

351 ttttttttac tattgtgaca ataatattga taaaaaaat tatttgccaa ctttcacaaa aatattttga 420
 Phospho
 Patatin
 Sn-2

421 cgcattagta taactattta .atactattt ttttattttt tatttataaa aaagatgaag agttaatgat 490
 Phospho
 Patatin
 Sn-2

491 gttttaacaa agattttttt ttgtatgitt tagcaaaaaa cttcttgca aaggaagtgt acaatataat 560
 Phospho
 Patatin
 Sn-2

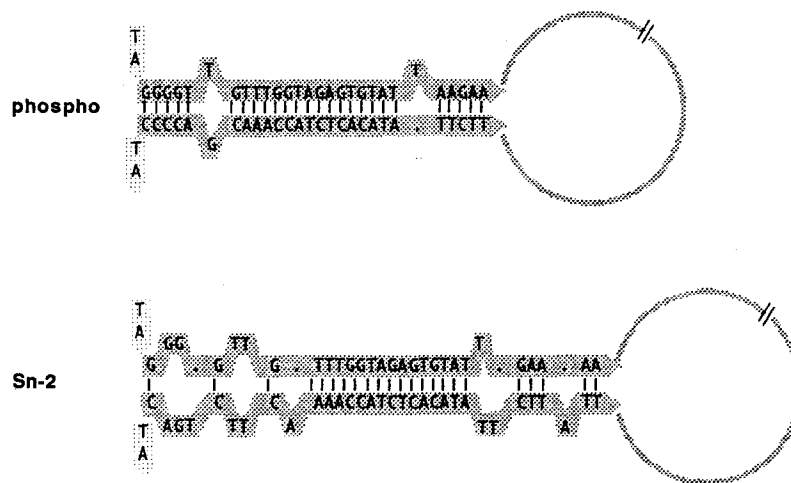
561 aaagtgtgaa gggatatttt gtaaacatat attatttaat agtaattatg caagatttat tatttttaat 630
 Phospho
 Patatin
 Sn-2

631 acatcaaac aaacaatgta taagaataaa tacttgcata actaatgcac gcactactaa tgcaagcatt 700
 Phospho
 Patatin
 Sn-2

701 actaatgcac catattttgt attgttctt atacactcta ccaaacgacc oct.ta 756
 Phospho
 Patatin
 pSKL
 Sn-2

attaatcac catattttgc a.ttttctt atacactcta ccaaaccttc tgacta

b



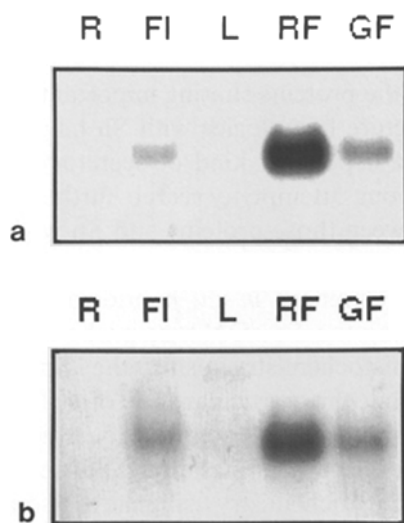


Fig. 5. Northern analysis of the *Sn* transcript in various organs and during fruit development. Ten μ g of total RNA isolated (see Material and methods) from roots (R), leaves (L), flowers (FI), young green unripe fruit (GF) and red ripe fruit (RF) were hybridized with the radiolabelled a) 643 bp *Sn* cDNA and b) *StyI/Bam* HI fragment from *Sn-1* (nucleotide position 1711–1993).

A computer search of protein sequence data bases (Blast NCBI) revealed an identity of 54% of the predicted amino acid sequence with that of the 691 bp FB7-4 cDNA [33] and an identity of 32% with two major latex proteins expressed in opium poppy (MLP 15 and MLP 22 [34, 35], respectively) (Fig. 7b). We must emphasize that in contrast to *Sn-1* whose transcription takes place during fruit development FB7-4 is expressed in the basal internodes and roots of tobacco during flowering [33].

The computer search analysis has also revealed the existence of a motif between amino acids 88 to 93 (Ile-Glu-Gly-Asp-Val-Leu) homologous to sequences described for the major allergen proteins Car b I of hornbeam and Cor a I of hazel

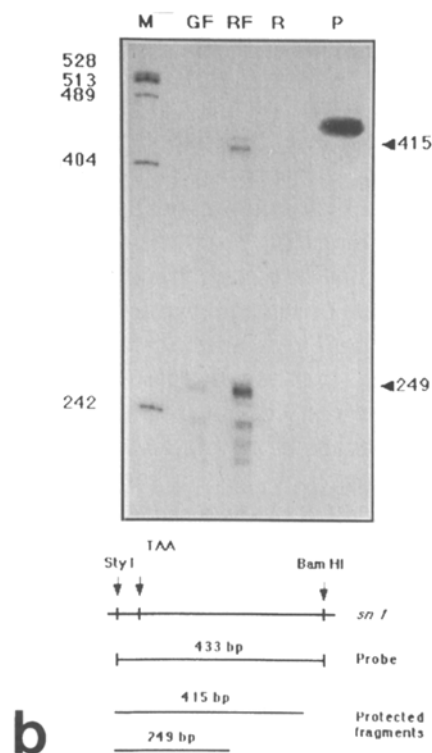
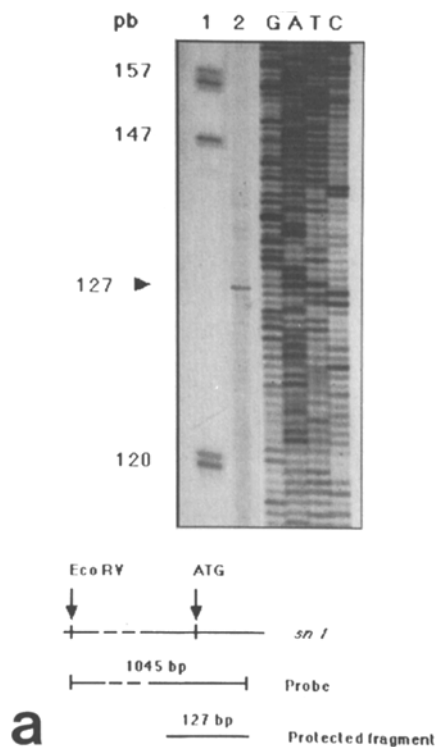
pollen ([4, 26], respectively) and PR proteins from potato (pSTH-21 and pSTH-2) and parsley (PR2) ([27, 29], respectively) (Fig. 7b). It is noteworthy that this motif is also homologous to an amino acid sequence of the immunologically active CD domain of the fish allergen [12].

All of these proteins lack signal peptides, they are ca. 17 kDa acidic proteins whose pI values are between 4.0 to 6.0 and circumstantial evidence suggests an important role of some of them in the plant's early disease resistance response [5].

Sn expression is stimulated by wounding

Western blot analysis using the specific antisera raised against the overexpressed protein encoded by the *Sn* cDNA detected a unic band of 17 kDa in samples corresponding to green and red fruits (Fig. 8a). As we have noted for the *Sn-1* transcripts (see Fig. 5), the relative amount of Sn-1 protein dramatically increased during the transition from mature-green to red fruit. The high homology of Sn with proteins likely involved in the protection of the plant prompted us to analyse the effect of wound stress in the expression of *Sn* genes and western blot analyses using samples from mature green fruits upon wounding were performed. To our surprise, we observed that the relative amount of Sn dramatically increased after 15 h of wounding (Fig. 8b) which further strengthens the view that Sn is likely involved in the plant's resistance response. Independent S1 mapping analyses using *Sn-1*- and *Sn-2*-specific probes demonstrated that the increase of the relative amount of Sn is ascribable to the stimulation of *Sn-1* rather than to the induction of *Sn-2* gene (data not shown).

Fig. 4. a. Sequence alignments of *Alien* elements from the promoter regions of the starch phosphorylase gene (phospho), patatin gene and intron of *Sn-2*. The positions of the direct repeats are indicated by solid arrows and the TIRs by open arrows. Conserved nucleotides are indicated by bold letters. For the case of the pSKL plasmid of *Saccharomyces kluyveri* only the TIRs delimited by '§' are represented, omitting the 14 kb between them. b. Potential hairpin structures formed by the *Alien* elements of starch phosphorylase and of *Sn-2*.



Localization of Sn protein

Some of the proteins sharing important sequence and structure homologies with Sn have been localized in a peculiar kind of secretory vacuoles [17]. In our attempt to search further possible links between those proteins and Sn we have investigated the tissular and subcellular localization of Sn. Preliminary *in situ* hybridization experiments using the *Sn* cDNA as probe and immunogold histochemistry using the Sn antisera showed that this particular type of cell is principally localized in the external exocarpic cell layers (manuscript in preparation). Subsequent immunogold cytochemistry using the antisera raised against the overexpressed protein encoded by the *Sn* cDNA showed that Sn-1 is exclusively associated to the membrane of a small vacuole which in many cases is localized close to the plasmalemma (Fig. 9). Controls using the gold particle-labeled IgG alone did not label the vacuoles (data not shown). The most obvious morphological feature of the cells containing these small vacuoles is that the cytoplasm is relatively abundant and they have numerous vesicles, lipid granules and organelles.

Discussion

It is well established that fruit ripening is a highly regulated developmental process requiring ex-

Fig. 6. Determination of *Sn-1* transcription start (a) and stop (b) sites by S1 nuclease mapping. In both cases the protected fragment is indicated by arrows and the radiolabelled *Hpa* II digests of the pBluescript KS⁻ were used as markers. a. The labelled 1045 bp single-strand DNA fragment existing between the *Eco* RV and nucleotide position 232 was hybridized with 20 μ g of total RNA extracted from ripe red fruit. The probe was synthesized in presence of α^{32} P-dCTP by *in vitro* extension of an oligonucleotide priming at positions 213 to 232. The G,A,T and C sequence ladder were obtained from the *Sn-1* genomic clone using the same primer. b. Twenty μ g of total RNA extracted from young green fruit (GF), red fruit (RF) and roots (R) were probed with the 433 bp *Sty* I/*Bam* HI DNA fragment (P). This fragment was labelled on the *Sty* I site in the presence of α^{32} P-dCTP and DNA polymerase I large (Klenow) fragment.

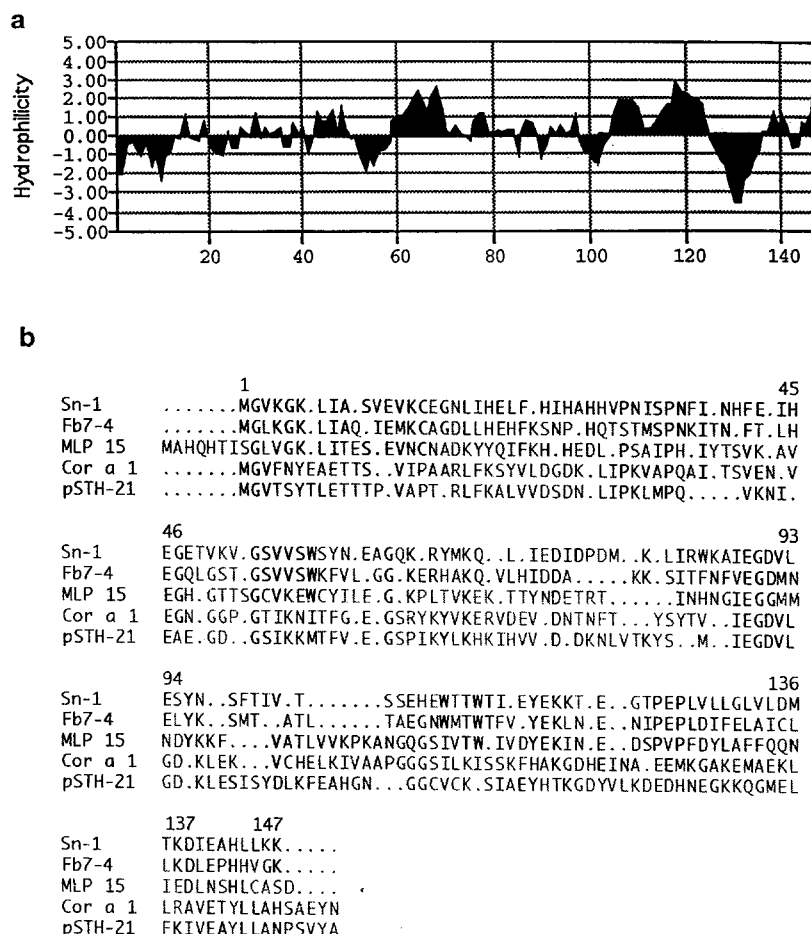


Fig. 7. Sn amino acid sequence analysis. a. Hydropathic profile displayed using the method of Kyte and Doolittle [25]. b. Amino acid sequence alignment of the deduced Sn-1 protein with similar polypeptides from tobacco (FB7-4), opium poppy (major latex protein MLP15), hazel (major allergen protein Cor AI) and potato (pathogenesis-related protein pSTH-21). Bold letters indicate residues common to the deduced Sn-1 protein.

pression of a large number of genes. Most of the work describing some ripening-specific genes [39] and their interactions with DNA-binding proteins [10, 32] as well as the functional maps of their respective promoters [44, 45] has been performed using the climacteric tomato fruit as a model system. In view of the scarcity of information concerning the expression of fruit- or ripening-specific genes of non-climacteric fruits we have analyzed for the first time the complete structure of a bell pepper gene family whose levels of the corresponding fruit-specific protein dramatically increase in the ripe stage.

Taken together, the results shown in Fig. 7b suggest that the Sn protein, FB7-4, the major

latex proteins, some pathogenesis related proteins and some major allergen proteins likely belong to the same group of proteins having common allergenically/antigenically active domains and which most likely participate in the plant's early disease resistance response [5]. A spectacular wound response of some plants to prevent the invasion of pathogens into the plant is the stimulation of secretion of some of these proteins [28]. Results in Fig. 8b show that Sn is a wound-stimulatable protein and its sequence homology with the latex proteins (see Fig. 7b) as well as its localization in a peculiar kind of small vacuole of cells principally localized in the external parts of the pericarp tissue (see Fig. 9) suggests that Sn is

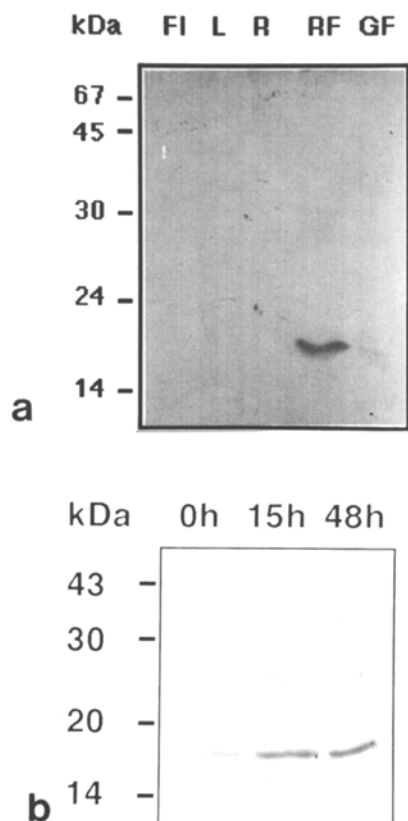


Fig. 8. Immunological analysis of samples from (a) flowers, leaves, roots and fruits at various developmental stages and (b) green fruit after 0 h, 15 h and 48 h of wounding. In each case 10 μ g of proteins were subjected to SDS-PAGE, followed by immunoblotting using antisera against the overexpressed protein encoded by the *Sn* cDNA.

likely a secretory protein having a role in the protection of the bell pepper fruit against pathogens. The analysis of the phenotypes obtained by reducing or increasing the expression of *Sn-1* and *Sn-2* will aid us in assigning biological functions to these genes.

In climacteric plants, ethylene enhances ripening of mature-green fruits. Ethylene also appears

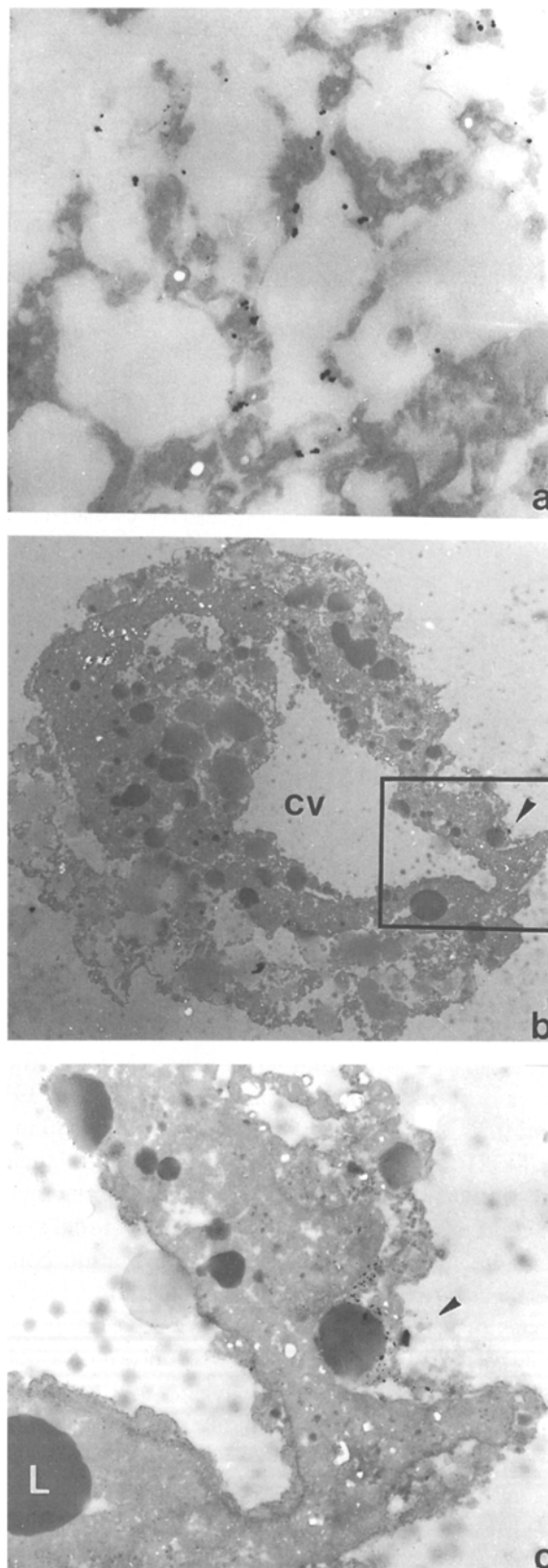


Fig. 9. Immunocytochemical localization of Sn-1 in protoplasts isolated from the pericarp tissue of ripe pepper fruit (see Materials and methods). a. Exclusive localization of Sn-1 in the membrane of vacuoles (Bar = 1 μ m). b. Transversal section of a typical kind of protoplast producing Sn-1 (cv = central vacuole, Bar = 10 μ m). c. Detailed picture of a vacuole containing Sn-1 in its membrane (indicated by arrow in b) (L = lipid granule, Bar = 1 μ m).

to function as a wound hormone, being an early, active response of plants to the perception of wound stress [2]. Thus, it is not surprising that wound stress enhances the ripening of climacteric fruits at the mature-green stage [40]. On the other hand, on the basis of its response to externally added ethylene or wound stress, bell pepper has been classified as a non-climacteric fruit [30]. Surprisingly, results in Fig. 8b show that production of Sn, a ripening-stimulated protein, is also stimulated by wound stress. Moreover, our presently ongoing experiments have shown that, as we have noted with Sn, wounding of the green pepper fruit also induces the expression of other ripening-specific genes, some of them participating in the mechanism of disease resistance response, i.e. a gene coding for a thionin (manuscript in preparation). These results could suggest that bell pepper has to some extent a climacteric-like behavior. Obviously further analyses will be necessary to elucidate the role of ethylene in the development and ripening of the bell pepper fruit.

Alien forms a unique family of plant small inverted repeat transposable elements. Standard hybridization protocols have been useful in identifying this new family (Fig. 3) indicating that, in contrast to the case of the previously reported *Tourist* and *Stowaway* small transposable elements [6, 7] respectively) *Alien* elements of bell pepper share a very high identity over their length. Although the *Alien* elements share no significant sequence similarity with *Tourist* and *Stowaway* transposable elements ([6, 7] respectively) these families of elements have similar structural features, thus suggesting that they were probably transposed by a similar 'cut-and-paste' mechanism. It is noteworthy that *Sn-1* possess a 41 bp fragment (nucleotide position 992 to 1034) at the same position than the *Sn-2 Alien* element. This sequence corresponds to a transposable element footprint arising from the excision of the *Sn-2 Alien* element. A similar case of footprints arising from the excision of small transposable element has been reported for the *Tourist* element in the maize *Adh1-C^m* allele [6].

Due to the limited coding capacity, the *Tourist*, *Stowaway* and *Alien* elements almost certainly do

not encode a transposase specifically recognizing their TIRs and they were probably mobilized by a transposase encoded at another chromosomal locus. TIRs from the *Alien* elements are strikingly homologous to sequences existing in the TIR of the linear DNA pSKL plasmid of *Saccharomyces kluyveri* (Fig. 4a). Interestingly, these sequences have associated a protein which is believed to be required for autoreplication of the linear DNA molecule by mechanisms similar to the DNA molecules of adenoviruses and bacteriophage $\phi 29$ of *Bacillus subtilis* [21]. The existence of these sequences in the genome of higher plants raises the question about their functional and evolutionary roles.

The sequences upstream the translation start codons from *Sn-1* and *Sn-2* share high homology (see Fig. 2) and thus both genes are expected to have a similar expression pattern. Nevertheless we have not detected any *Sn-2* transcript in neither of the samples monitored. It is therefore tempting to speculate that as we have noted with *Sn-1*, *Sn-2* is transcribed in the fruit but the predicted $\Delta G = -412$ kJ/mol hairpin of its *Alien* element acts as a mRNA destabilizing agent imposing a spatial or temporal pattern of expression different to that of *Sn-1*. It is particularly interesting that nearly 50% of the small *Stowaway* transposable elements are located in the introns and in the regions flanking the 5' end of stress-inducible or stress-stimulatable plant genes [7].

Studies of the possible interaction between the *Sn-2 Alien* element and flanking regions likely involved in the control of *Sn-2* expression are presently in progress. Structure of the promoter and of DNA binding factors of *Sn-1* and other genes that we have identified with enhanced expression at various stages of fruit development (manuscript in preparation) will be useful tools in future studies of gene expression and development in non-climacteric fruits.

Acknowledgements

We express our most sincere thanks to Dr Nicholas S. Foulkes (LGME du CNRS, Faculté

de Médecine, Strasbourg) for his critical review of the paper and improvement of the text. We also thank Dr Rohfritsch (IBMP du CNRS, Université Louis Pasteur, Strasbourg) for assistance in the electron microscopic analysis. J.P.-R. acknowledges a postdoctoral student fellowship from the CEE.

References

1. Bogorad L, Gubbins EJ, Krebbers E, Larrinua IM, Mulligan BJ, Muskavitch KMT, Orr EA, Rodermel SR, Schantz R, Steinmetz AA, DeVos G, Ye YK: Cloning and physical mapping of maize plastid genes. *Meth Enzymol* 97: 524–554 (1983).
2. Boller T: Ethylene in pathogenesis and disease resistance. In: Mattoo AK, Suttle JC (eds) *The Plant Hormone Ethylene*, pp. 293–314. CRC Press, Boca Raton, FL (1991).
3. Breathnach R, Chambon P: Organization and expression of eukaryotic split genes coding for proteins. *Annu Rev Biochem* 50: 349–383 (1981).
4. Breiteneder H, Ferreira F, Hoffmann-Sommergruber K, Ebner C, Breitenbach M, Rumpold H, Kraft D, Scheiner O: Four recombinant isoforms of *Cor a I*, the major allergen of hazel pollen, show different IgE-binding properties. *Eur J Biochem* 212: 355–362 (1993).
5. Breiteneder H, Pettenburger K, Bito A, Valenta R, Kraft D, Rumpold H, Scheiner O, Breitenach M: The gene coding for the major birch pollen allergen *Betvl*, is highly homologous to a pea disease resistance response gene. *EMBO J* 8: 1935–1938 (1989).
6. Bureau TE, Wessler SR: *Tourist*: a large family of small inverted repeat elements frequently associated with maize genes. *Plant Cell* 4: 1283–1294 (1992).
7. Bureau TE, Wessler SR: *Stowaway*: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. *Plant Cell* 6: 907–916 (1994).
8. Camara B, Bousquet J, Cheniclet C, Carde JP, Kuntz M, Evrard JL, Weil JH: Enzymology of isoprenoid biosynthesis and expression of plastid and nuclear genes during chromoplast differentiation in pepper fruits (*Capsicum annuum*). In: Boyer CD, Shannon JC, Hardison RC (eds) *Biochemistry and Genetics of Nongreen Plastids*, pp. 141–156. American Society for Plant Physiology (1989).
9. Camirand A, St-Pierre B, Marneau C, Brisson N: Occurrence of a copia-like transposable element in one of the introns of the potato starch phosphorylase gene. *Mol Gen Genet* 224: 33–39 (1990).
10. Cordes S, Deikman J, Margossian LJ, Fischer RL: Interaction of a developmentally regulated DNA-binding factor with sites flanking two different fruit-ripening genes from tomato. *Plant Cell* 1: 1025–1034 (1989).
11. Draper J, Scott R: The isolation of plant nucleic acids. In: Armitage P, Walden R (eds) *Plant Genetic Transformation and Gene Expression: A Laboratory Manual*, pp. 201–236. Blackwell Scientific Publications, Oxford (1988).
12. Elsayed S, Apold J, Holen H, Vik H, Florvaag E, Dybendal T: The structural requirements of epitopes with IgE binding capacity demonstrated by three major allergens from fish, egg and tree pollen. *Scand J Clin Lab Invest Suppl* 204: 17–31 (1991).
13. Feinberg AP, Vogelstein B: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 137: 266–267 (1983).
14. Gasser CS, Budelier KA, Smith AG, Shah DM, Fraley RT: Isolation of tissue-specific cDNAs from tomato pistils. *Plant Cell* 1: 15–24 (1989).
15. Gillaspay G, Ben-David H, Gruissem W: Fruits: a developmental perspective. *Plant Cell* 5: 1439–1451 (1993).
16. Goodall GJ, Filipowicz W: The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. *Cell* 58: 473–483 (1989).
17. Griffing LR, Nessler CL: Immunolocalization of the major latex proteins in developing laticifers of opium poppy (*Papaver somniferum*). *J Plant Physiol* 134: 357–363 (1989).
18. Gualberto JM, Wintz H, Weil JH, Grienemberger JM: 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 (1988).
19. Guan KL, Dixon JE: Eukaryotic proteins expressed in *Escherichia coli*: an improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. *Anal Biochem* 192: 262–267 (1991).
20. Heidecker G, Messing J: Structural analysis of plant genes. *Annu Rev Plant Physiol* 37: 439–466 (1986).
21. Hishinuma F, Hirai K: Genome organization of the linear plasmid, pSKL, isolated from *Saccharomyces kluyveri*. *Mol Gen Genet* 226: 97–106 (1991).
22. Joshi CP: Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. *Nucl Acids Res* 15: 9627–9640 (1987).
23. Kaiser K, Murray NE: The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. In: Glover DM (eds) *DNA Cloning: A Practical Approach*, pp. 1–47. IRL Press, Oxford/Washington DC (1985).
24. Kozak M: Compilation and analysis of sequences upstream from the translation start site in eukaryotic mRNAs. *Nucl Acids Res* 12: 857–872 (1984).
25. Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157: 105–132 (1982).
26. Larsen JN, Stroman P, Ipsen H: PCR based cloning and sequencing of isogenes encoding the tree pollen major allergen *Car b I* from *Carpinus betulus*, hornbeam. *Mol Immunol* 29: 703–711 (1992).

27. Locht U, Meier I, Halbrock K, Sommsich IE: A 125 bp promoter fragment is sufficient for strong elicitor-mediated gene activation in parsley. *EMBO J* 9: 2945–2950 (1990).
28. Lürssen K: Ethylene and Agriculture. In: Mattoo AK, Suttle JC (eds) *The Plant Hormone Ethylene*, pp. 315–326. CRC Press, Boca Raton, FL (1991).
29. Matton DP, Brisson N: Cloning, expression and sequence conservation of pathogenesis-related gene transcripts of potato. *Mol Plant-Microbe Inter* 2: 325–331 (1989).
30. Mikal E, Saltveit J: Carbon dioxide, ethylene and color development in ripening mature green bell peppers. *J Am Soc Hort Sci* 102: 523–525 (1977).
31. Monroy AF, Schwartzbach SD: Photocontrol of the polypeptide composition of *Euglena*. *Planta* 138: 249–258 (1983).
32. Montgomery J, Goldman S, Deikman J, Margossian LJ, Fischer R: Identification of an ethylene-responsive region in the promoter of a fruit ripening gene. *Proc Natl Acad Sci USA* 90: 5939–5943 (1993).
33. Neale AD, Wahleithner JA, Lund M, Bonnett HT, Kelly A, Ry Meeks-Wagner D, Peacock WJ, Dennis ES: Chitinase, β -1,3-glucanase, osmotin, and extensin are expressed in tobacco explants during flower formation. *Plant Cell* 2: 673–684 (1990).
34. Nessler CL, Burnett RJ: Organization of the major latex protein gene family in opium poppy. *Plant Mol Biol* 20: 749–752 (1992).
35. Nessler CL, Kurz WGW, Pelcher LE: Isolation and analysis of the major latex protein genes of opium poppy. *Plant Mol Biol* 15: 951–953 (1990).
36. Oeller PW, Min-Wong L, Taylor LP, Pike DA, Theologis A: Reversible inhibition of tomato fruit senescence by antisense RNA. *Science* 254: 437–439 (1991).
37. Saunders SE, Burke JF: Rapid isolation of miniprep DNA for double strand sequencing. *Nucl Acids Res* 18: 4948 (1990).
38. Schmit A-C, Lambert A-M: Characterization and dynamics of cytoplasmic F-actin in higher plant endosperm cells during interphase, mitosis, and cytokinesis. *J Cell Biol* 105: 2157–2166 (1987).
39. Slater A, Maunders MJ, Edwards K, Schuch W, Grierson D: Isolation and characterization of cDNA clones for tomato polygalacturonase and other ripening-related proteins. *Plant Mol Biol* 5: 137–147 (1985).
40. Theologis A: One rotten apple spoils the whole bushel: the role of ethylene in fruit ripening. *Cell* 70: 181–184 (1992).
41. Tonouchi N, Miwa K, Karasuyama H, Matsui H: Deletion of 3' untranslated region of human BSF-2 mRNA causes stabilization of the mRNA and high-level expression in mouse NIH3T3 cells. *Biochem Biophys Res Comm* 163: 1056–1062 (1989).
42. Towbin H, Staehelin T, Gordon J: Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedures and some applications. *Proc Natl Acad Sci USA* 76: 4350–4354 (1979).
43. Twell D, Ooms G: Structural diversity of the patatin gene family in potato cv. Désirée. *Mol Gen Genet* 212: 325–336 (1988).
44. Van Haaren MJJ, Houck CM: Strong negative and positive regulatory elements contribute to the high-level fruit-specific expression of the tomato 2A11 gene. *Plant Mol Biol* 17: 615–630 (1991).
45. Van Haaren MJJ, Houck CM: A functional map of the fruit-specific promoter of the tomato 2A11 gene. *Plant Mol Biol* 21: 625–640 (1993).
46. Veidt I, Bouzoubaa SE, Leiser R-M, Ziegler-Graf V, Guilley H, Richards K, Jonard G: Synthesis of full-length transcripts of beet Western Yellows virus RNA: Messenger properties and biological activities in protoplasts. *Virology* 186: 192–200 (1992).
47. Zhang H, Scholl R, Browse J, Somerville C: Double-stranded DNA sequencing as a choice for DNA sequencing. *Nucl Acids Res* 16: 1220 (1988).