Plant Molecular Biology 12: 367-383, 1989
© 1989 Kluwer Academic Publishers. Printed in Belgium

Phenylalanine ammonia-lyase gene organization and structure

Carole L. Cramer[1,5], Keith Edwards[2], Michel Dron[1,7], Xaiowu Liang[1,3], Sandra L. Dildine[1], G. Paul Bolwell[4], Richard A. Dixon[4,6], Christopher J. Lamb[1]* and Wolfgang Schuch[2]
[1]*Plant Biology Laboratory, Salk Institute for Biological Studies, P.O. Box 85800, San Diego, CA 92138, USA (* Author for correspondence);* [2]*Plant Biotechnology Group, Imperial Chemical Industries plc, The Heath, Runcorn, Cheshire WA7 4QE, UK;* [3]*Graduate Program in Biology, University of California San Diego, La Jolla, CA 92037, USA;* [4]*Department of Biochemistry, Royal Holloway and Bedford New College, University of London, Egham, Surrey TW20 OEX, UK;* [5]*Present address: Department of Plant Pathology, Physiology and Weed Science, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA;* [6]*Present address: Plant Biology Division, Samuel Roberts Noble Foundation, P.O. Box 2180, Ardmore, OK 73402, USA;* [7]*Present address: Laboratoire Biologie Moleculaire Végétale, Université Paris Sud, 91405-Orsay Cédex, France*

Received 4 November 1988; accepted in revised form 14 December 1988

Key words: elicitor, gene family, gene sequence, *Phaseolus vulgaris*, phenylalanine ammonia-lyase, phenylpropanoid biosynthesis

Abstract

Phenylalanine ammonia-lyase (PAL; EC 4.3.1.5) genomic sequences were isolated from bean (*Phaseolus vulgaris* L.) genomic libraries using elicitor-induced bean PAL cDNA sequences as a probe. Southern blot hybridization of genomic DNA fragments revealed three divergent classes of PAL genes in the bean genome. Polymorphic forms were observed within each class. The nucleotide sequences of two PAL genes, gPAL2 (class II) and gPAL3 (class III), were determined. gPAL2 contains an open reading frame encoding a polypeptide of 712 amino acids, interrupted by a 1720 bp intron in the codon for amino acid 130. gPAL3 encodes a polypeptide of 710 amino acids showing 72% similarity with that encoded by gPAL2, and contains a 447 bp intron at the same location. At the nucleotide level, gPAL2 and gPAL3 show 59% sequence similarity in exon I, 74% similarity in exon II, and extensive sequence divergence in the intron, 5′ and 3′ flanking regions. S1 nuclease protection identified transcription start sites of gPAL2 and gPAL3 respectively 99 bp and 35 bp upstream from the initiation codon ATG, and showed that gPAL2 but not gPAL3 was activated by elicitor, whereas both were activated by wounding of hypocotyls. The 5′ flanking region of both genes contain TATA and CAAT boxes, and sequences resembling the SV40 enhancer core. gPAL2 contains a 40 bp palindromic sequence and a 22 bp motif that are also found at similar positions relative to the TATA box in 5′ flanking regions of other elicitor-induced bean genes.

Introduction

Phenylalanine ammonia-lyase (PAL; EC 4.3.1.5) is a key regulatory enzyme in plant metabolism catalyzing the deamination of L-phenylalanine to yield *trans*-cinnamic acid and NH_4^+. This is the first reaction in the biosynthesis of a wide variety of natural products based on the phenylpropane skeleton, including the cell wall structural polymer lignin, flavonoid pigments and UV protectants, furanocoumarin and isoflavonoid phytoalexins, and wound protectant hydroxycinnamic acid esters [15, 23]. In addition, the phenylpropanoid-derived wound metabolite acetosyringone and certain flavones have recently been identified as signals for activation of virulence genes in the tumor-inducing bacterium *Agrobacterium tumefaciens* and nodulation genes in the symbiotic nitrogen-fixing bacteria of the Rhizobiaceae respectively [17, 36].

PAL activity is highly regulated during development associated with cell-type specific synthesis of lignin or flavonoid pigments, and by an array of environmental stimuli including wounding, infection, and light associated with synthesis of specific phenylpropanoid products involved in adaptation or protection [15, 23]. Fluctuations in PAL levels are a key element in the regulation of phenylpropanoid synthesis and labeling of the enzyme *in vivo* and mRNA translation *in vitro* have shown environmental control over PAL *de novo* synthesis [15, 16, 19, 23]. cDNA clones have been identified that are complementary to PAL mRNA from irradiated parsley cells [24], and bean cells treated with a fungal cell wall elicitor of phytoalexin production [20]. RNA blot hybridization and nuclear run-off experiments show that elicitor, wounding, infection and irradiation rapidly stimulate transcription of PAL genes leading to marked accumulation of PAL mRNA and hence increased enzyme synthesis and activity [9, 12, 13, 20, 24, 25].

In the present study we have used bean PAL cDNA sequences to identify PAL genomic sequences and to initiate the characterization of the structure and organization of PAL genes in relation to their function and regulation. We demonstrate that in bean, PAL is encoded by a small gene family and describe the complete nucleotide sequence, deduced amino acid sequence, intron-exon structure and transcription start sites of two differentially regulated members of the gene family. This study represents the first molecular characterization of this important family of plant genes.

Materials and methods

Plant material

Cell suspension cultures of bean (*Phaseolus vulgaris* L.) cv. Canadian Wonder and cv. Tendergreen were generated and maintained as previously described [4, 12]. Germination and growth of bean seedlings were as described [34].

cDNA clones

Generation, identification and characterization of the PAL cDNA clones pPAL1 through pPAL5, complementary to an elicitor-induced PAL transcript, have been described [20].

Isolation of genomic clones

Genomic DNA from leaves of bean cv. Canadian Wonder was digested with *Eco* RI to completion and cloned into λgtWES *Eco* RI arms [26]. This genomic library, and a previously described library of a partial *Mbo* I digest of bean cv. Tendergreen genomic DNA cloned in λ1059 [38], were plated and clones containing PAL sequences identified by *in situ* hybridization with nick-translated pPAL5 cDNA sequences. Purified clones containing PAL sequences were analyzed by restriction digestion and blot hybridization, and representative restriction fragments subcloned into M13 [35] or pSP6 vectors [32] for further analysis (Table 1).

DNA preparation

Phage particles were purified by CsCl banding of eluates from plates of *Escherichia coli* grown to confluent lysis. Plasmid and phage DNA was prepared by standard methods [30]. Bean genomic DNA was isolated from germinated seedlings or suspension-cultured cells of cv. Tendergreen or cv. Canadian Wonder using the method of Murray and Thompson [33].

Blot hybridization analysis

Bean genomic DNA samples were digested with various restriction enzymes, fractionated by electrophoresis on agarose gels and blotted as described [10, 30]. Hybridizations to genomic DNA from cv. Tendergreen were performed at 42 °C in a solution of 50% formamide, 4 × SSC (1 × SSC is 0.15 M NaCl and 0.015 M sodium citrate), 5 × Denhardt's solution (1 × Denhardt's is 0.02% Ficoll, 0.02% polyvinyl pyrrolidone (PVP), 0.02% bovine serum albumin (BSA), 5 mM EDTA, 0.1% SDS and 100 μg/ml sheared, denatured calf thymus DNA. Filters were incubated for 16 to 24 h in buffer without probe prior to hybridization for 24 to 48 h with probe [^{32}P]-labeled (> 10^8 cpm/μg DNA) by nick translation [30]. Prior to autoradiography, blots were washed twice in 2 × SSC, 1 mM EDTA and 0.1% SDS for 15 min at room temperature and twice in 0.2 × SSC, 1 mM EDTA and 0.1% SDS for 1 h at room temperature or 37 °C. Blots of genomic DNA from cv. Canadian Wonder were pre-hybridized in 5 × SSPE (20 × SSPE is 3.6 M NaCl, 0.2 M NaH_2PO_4, 20 mM EDTA, pH 7.4), 1% SDS and 100 μg/ml sheared herring sperm DNA at 65 °C for 3 h. Hybridization with probe was performed under the same conditions for 18 h. For low stringency hybridization, filters were washed in 2 × SSC, 1% SDS at 65 °C for 4 h. For high-stringency hybridization, filters were washed in 0.1% SSC, 1% SDS at 65 °C for 5 h.

Nucleotide sequence analysis

DNA was subcloned in M13 vectors and sequenced by the dideoxy chain termination method [35]. Fragments for sequencing were generated by restriction endonuclease digestion or by rapid-deletion cloning [14]. Regions in gPAL2 that gave compressions in dideoxy sequencing gels were also analyzed by the method of Maxam and Gilbert [31].

S1 nuclease protection

Isolation and purification of total cellular RNA from elicitor-treated cell cultures [12] and excision-wounded hypocotyls [34] was as previously described. [^{32}P]-labeled single-stranded DNA probes [3], prepared as described in the relevant figure legends, were hybridized to 25–100 μg of total cellular RNA in 20 mM Pipes, pH 6.8, containing 50% formamide, 0.5 M NaCl, 5 mM EDTA and 10 μg calf liver tRNA at 50 °C for 18 h. Protection against digestion by S1 nuclease was performed as described [30], and products analyzed on 7% sequencing gels.

Results

Gene structure

pPAL5 cDNA sequences were used to identify clones containing bean PAL genomic sequences. Screening of the 2 independent genomic libraries identified 14 clones which were plaque purified and grouped into 3 classes by restriction fragment length analysis and Southern blot hybridization (Table 1 and Fig. 1). Class I and Class II were found in the libraries of cv. Tendergren and cv. Canadian Wonder genomic DNA respectively, whereas class III was represented in both libraries (Table 1).

Nucleotide sequence analysis showed that class I comprised a set of genomic clones that contained truncated versions of the gene, designated gPAL1, encoding the elicitor-induced

Table 1. Isolated bean PAL genomic clones.

Clone	Description[a]	Source[c]
1	gPAL2[a]	C
2	gPAL2[b]	C
3	gPAL3, 3′-truncated[b]	C
4	gPAL2[a]	C
5	gPAL2[a]	C
6	gPAL3[a]	C
7	gPAL3[a]	T
8	gPAL3[a]	T
9	gPAL1, 5′-truncated[b]	T
10	gPAL1[a]	T
11	gPAL1[a]	T
12	gPAL3, 5′-truncated[b]	T
13	gPAL3[a]	T
14	gPAL3[b]	T

[a] Analyzed by restriction mapping and hybridization with gene-specific probes.
[b] Confirmed by sequencing, for further information see text.
[c] C, Canadian Wonder genomic library in λgtWES; T, Tendergreen genomic library in λ1059.

transcript represented in cDNA clone pPAL5 (Fig. 1). The 5.3 kb *Eco* RI insert of gPAL2 (class II; clone 2) and a 3.6 kb *Bam* HI fragment of gPAL3 (class III; clone 14) were sequenced and found to contain complete PAL genes including flanking regions (Fig. 2A). Comparison of the nucleotide sequences of gPAL2 and gPAL3 by dot matrix analysis illustrates the presence of two conserved, colinear regions of homology, which correspond to the protein coding sequences, separated by an intron that is divergent with respect to both size and nucleotide sequence (Fig. 3).

Intron-exon junctions and transcription start sites were established by S1 nuclease protection experiments. Thus, a 155 bp fragment of the 536 bp *Hinc* II-*Hind* III fragment of gPAL2, which spans the intron-exon II junction (Fig. 1) is protected from S1 nuclease digestion by mRNA isolated from elicitor-treated cells (Fig. 4A). This positions the junction between two G residues at

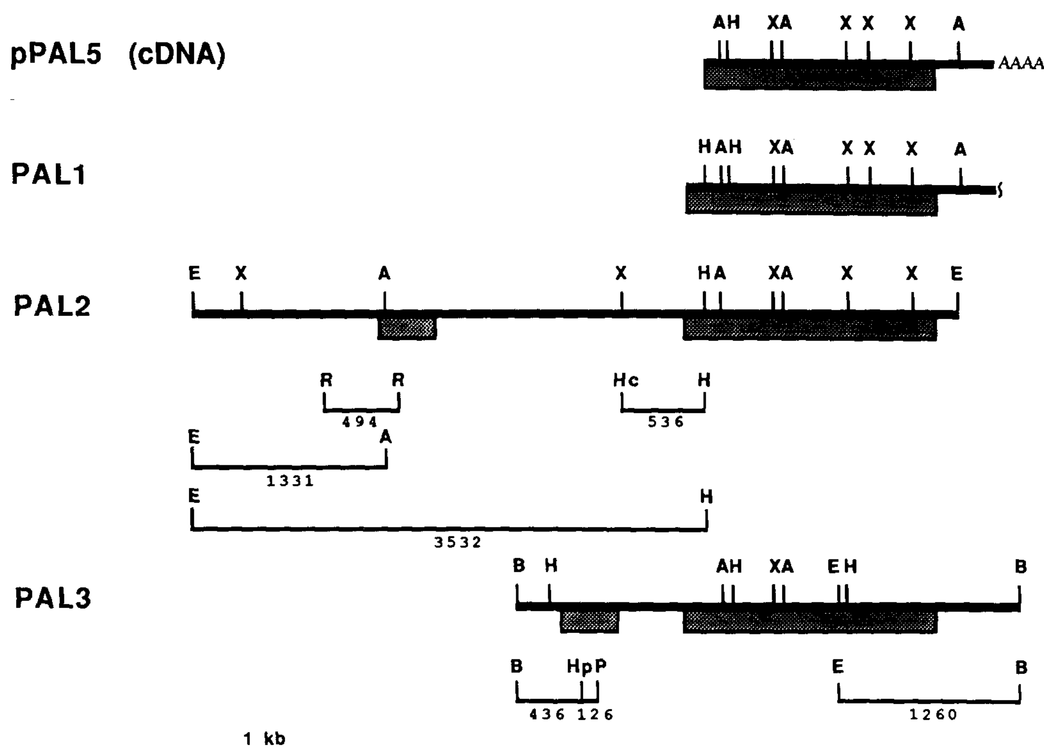

Fig. 1. Partial restriction maps of the cDNA clone pPAL5 and genomic subclones representing three classes of bean PAL genes. Protein coding regions are indicated by hatched boxes. Specific restriction fragments used in S1 nuclease protection studies or Southern blot hybridization analyses are indicated below gPAL2 and gPAL3. Restriction enzymes: A = *Ava* II, B = *Bam* HI, E = *Eco* RI, H = *Hind* III, Hc = *Hinc* II, Hp = *Hpa* II, P = *Pvu* II, R = *Rsa* I, and X = *Xho* II.

A

```
gPAL2  *GAATTCCCAGTCCAAGTCATAGTTATTGCAGGGATTAATCAAATTTTATAAAAATATGGCATTGGTGTTTAAAAGCAACATATAATAAGTTATGGATGTGGCCTGTAATAGGAGAGCAC

gPAL2  AAAGGTATTTTTGTAGCATCTTTTTATTATGAAGGATATTTACTTTTATTCAACTATTATTTTAAAAGTCTAAGCCAATATTCATTATTTTTTTATTTATGCTTAACATTTATGTTCAAG

gPAL2  CCAATAGTAACAAGAAGATGAACTGGTTTTGTATATTAATTCAATATAACCAATCCCTGTGGAGTGATTTAGTTGAAAGGATCTACAATTCTAAAGATGATGAGTTAAACCCTAGCCGAA

gPAL2  GTTGTAGTTGAATCTTAAACTTACTTTTCTAAATAGTAACACCTGCCATAAATACTAGCTGCATCTTAATTTCTCTACCTCCCCCACACTCTGGCATGGCGGCCCTGTCGTTTTCTTGCT

gPAL2  CCATTTTTTTTTCTATTATCACACTTTTTCTCCTTTCTTCTAAGTAATTACTAAAACAAATGCTAAAGAAACACATTTATTTATTTATTTTTATCTTTCTAACCCTATTTAACCAGTTTA

gPAL2  GAAGCCAATTCCCAGAATCATAGTTCACTTTAACTATGTTTTTTTTTAGTGAGAAGAAGACAAAAGATGAATGATTGGTTGCGATTCCTTGCCCTTTTGTTCTTCTTATATATATATATAT
                                                                                                                    B

gPAL2  ATATATATATATATATATATATATATACCTAGAAAAGATAATATGTACGTTGAAATCTTTGTTAAAAGGAACCAGAAAATGTAAGGATTAGGAATTTAATTTCGTAGTTGGGTAACATTTAT
                B

gPAL2  TAATTAATCAATTAATTAATTAATTGATTGATTTATACGTCTATCTTCTATTCCACGTCCTGTGTTGGTAGGGAAGTACAGAGAAGTTAGGTTCTAGTCCACAAGGTGACATCGCACCCA
                    C
                                                                                                    A
gPAL2  GAGAAGGGGAGAAAAAATGCCACGTCGCGCAATGAGAGCCGCTGATGCAGGCTGGTAATCCTTCGCTTGTCATTATTTCTCCACCAACCCCCTTCACTTCCCCTTGTGCATCGTTACCAC
gPAL3                  *G TCCATGGAGAATCAGCCA A  AAAA TCAAGTT TCAAAGTAGT  TTT CT GT A GC   TT TTTTTGTTT AGTAC A TAA  ACA GGA  GTA G
              D                                                                                          TXN
gPAL2  CCTTTATACCCACCTACCAGACACCACGCTCCAGATTTGCTTCGGCCCTAACACTCTCCGTTATATATAACCCCTTCATGAACAAGGCTAATCATTCACCACATTACGCACTACTTTTCC
gPAL3   TGCC A ATG  TGGGTCAT TA  ATGAAT  CAA C  GAT TATC  AGT GAAGAGGTACCACA   A CCC G  CACA  GGT C        ACCA  T TGCC  C ATT
                                                                              A
gPAL2  TCTCTCCGTCTCCTCATTCCTTCATTTAAATCTGCTTCCCCACTCACCTACCCAAACGCGCCACTAACTCACATGGACGCAACGCCCAATGGAAAAGACGCTTTCGTCGTCACTGCCGCC
gPAL3  AAAG TTC AAA TTCCAT   T G  C A  TT GAT TCTC TA  G TAGCTTCACTGGT TGTAA     ...    CATT C CC C  A G G AG CAAAGTGAG TG GT
                                        TXN
gPAL2  AATGCCGCGGGGGACCCGCTCAACTGGGCCGCTGCGGCGGAGGCGCTCTCGGGCAGTCACCTCGACGAGGTCAAGCGCATGGTCGCCGAGTACCGCAAGCCCGCGGTGCGCCT...CGGC
gPAL3  GTGAG AACACA       C            GTAA   A T    T A  GAA    T  C   T G  G  A  A  AG G     G G            GG  G A T A C ATG GGGG  A

gPAL2  GGTCAGACGCTCACCATCGCTCAGGTAGCGGCCACTGCGGCGCACGACCAGGGTCTCAAGGTGGAGCTGGCGGAGTCCGCCAGGGCCTGCGTCAAGGCCATCAGTGACTGGGTGATGGAG
gPAL3   GG      A  T  TG AT AA     G  T  TGT   AAAC  GT  TT CAGGCT        T  C  CT T       T  A  A AAG     TG CT   G T CA G   A CG    T
                                                                                                                  INTRON
gPAL2  AGCATGGACAAGGGGACTGACAGTTACGGTATCACCACCGGGTTCGGTGCAACCTCCCACCGCCGTACCAAACAAGGTGGCGCCTTGCAGAAGGAGCTTATCAGGTCACTATAATCTACA
gPAL3   A   C     A  C T CC  T   T  G          T  T  T         A     A  A G AA  TC  G    CCTT  TC T              AA GG T   G GG  C CTC AC
                                                                          E
gPAL2  AAAGTCTCTCTCTCTCTCTCTTTCTGTATGCATTACATTATTACACATGCTCATCTGCAACTGGTGTTGCACCTCGTGAATCAACGCAAGAAAAATAGCATATTTTCTACCACATCTT
gPAL3  TCT CT TG  TGTGAAG  GTACA  CCT TT AATGTCGTG CTGT  CGA   A  T T G A A C ACAT    AG A  GT ATGTGTCT ATTT ATA   G TGA TT   G

gPAL2  TCAACATGACATGTATATATATAAACACAACAATTTGTGTATTACATATTAAATTTGTTTTTTTCTTGTCTTGCGATCTTACGAATGATATTAAAATTATTTCACTAGTAACATCTCTTT
gPAL3  GT  TT A A CAC  T T  A   A A TAC GGGT TGTA CAG G  TGTCAGTACCAG GT    T CTT..............................................

gPAL2  CTTGTCATGGTTCTATCGAACCTGTAACTGTAGGTGCAGCATCGTGAAGGAAAATTGCATGATCATGTATCATAAAATGAAAAAAAACATATTATTTTACATAAATTAATTTCACAACAA
gPAL3  ........................................................................................................................

gPAL2  GATTGTTTTTAACATCGTTTGTTGTAAATGTTTGAAAAAATAAAGAAAATGTGTGAAAAAAAATTGAAGGATGTTGTGTTTAATATATAATATATACTAATAGTAAAATTTGTTTTTTCT
gPAL3  ........................................................................................................................

gPAL2  ACTTTCTCTGAACTTTTCTTCCCTCTTTGCCATTGTACCCATCCCACACCCTCTTTCATTATGAGAAATAATTAAAGATGGACGATCTAATTAATATTATATGAACGTTTTTTTATTATT
gPAL3  ........................................................................................................................

gPAL2  CTTCTTTTATTCCCTTATATATTATTATTCCGTTGTTATTTTTCTGTTTCTGTTGCCCATTCATCATTTTCTATTGCCCAACATGCGTCTTATACTTTATTTTATTACTGCTTTCATTTC
gPAL3  ........................................................................................................................

gPAL2  TCTTAAGTAAATTATTGTTGCTCGGCTACCTAAAATATTCACCACACTAATTAACGAATTTAAATGAAATATTATTCATTGTTTCTCATAGTTGGAGAAAAGGAGGTGGATAAGTTGGCT
gPAL3  ........................................................................................................................

gPAL2  CACCTAACACAGGAAATTTGGTAATATTTTTCTTTTATTAGAGAGAGTTGTTTCTTATACTAAAAAAATTAGGTTTAGTATGCTATATTTATTTTTACTTTAATTCATGGTATATATGTT
gPAL3  ........................................................................................................................

gPAL2  TTTTTTTCCTAATATAGTAAGAATTTTGAAAAGGTTTCTGGGTTCAACTGTTCAAACCTTCAGACAATTTTAGATAAACTTTTTTATAAAAAATAAAAAACTCTTCAGACAAATAATTGA
gPAL3  ........................................................................................................................

gPAL2  TATAAAAATATATTATAATAATAATTATGTGTCATCTATAATAATATTTATGATAATTTCTGAGCTATTTTTAAAATATATTTGGTGAACAAATAATAAACGCGGTTTTAATAATTTGTT
gPAL3  ........................................................................................................................

gPAL2  GGGTGGTAAATTAAAAGAACACGTGCTGGCATTCAAAGAGAGTTGTATTGTCTGAATCGGATACTACCCTGATCTGGACTCGAGATCCTATTGAGTCGAAACGGCCGAGTGGCGAGTTAA
gPAL3  ........................................................................................................................

gPAL2  CTCATTAAATGATAACACGACGCACGTGCGCGAGGTTGGTGAGACTGCCAAGCCCTCCAACCTTTCCAAGTCCTAACACTCCATCTTTGACTTCTTAGCTTCTCAATTTTCTAATTTATT
gPAL3  ........................................................................................................................

gPAL2  TTTTGTCCTAATTTATTACACCCAATCAATTCAATTTTTTAATTTTTTGTATTTTTTTATGACATCAGCACACTAATAATGTCAATTATTTATTTTTCGTTTTTCCCCTATTTTATCATG
gPAL3  ..........................TG  TTTGC ACGT GA G AT T  C AG C  TAGGA TT TC A C TCATAGTGAA ACAC AA  AT C GAGATTATCAC T TC C

gPAL2  GTGTCTGTTCCAAAAGAATTATGTCTTTTCTTGGTTTTGTATGTTGATGTCCTCACCTAATTCCCCACCCAACTACATCATATCTTTTATTAATTTATAAAAAAGTTCATAATTCTTAAT
gPAL3  T TGTCT GTA GTT T  C  CAAC  ATC CT  CAACTGA AA A AATCACTTC TA A AACA G  ACTGCATTG  TA A   A T C TCTTTTTG   G  GTA TCATGG
                      INTRON
gPAL2  CTTTTTTTTTGTTCTTCTTTCAGGTTTTTGAATGCTGGGATATTTGGCAATGGTACAGAGTCCAACTGCACCCTACCCCACACTGCAACTAGAGCAGCTATGCTTGTGAGAGTGAACACT
gPAL3  A   ACC C T   C TCCCA        A  CTGC CC           T CCAA  G    TATCTCAT  A  G  AA AT A     C           A            G  T  T  C
```

Fig 2. See p. 372.

Fig. 2. A. Comparison of the nucleotide sequence of gPAL2, gPAL3 and cDNA pPAL5 (gPAL1). Transcription start sites (TXN) and intron-exon boundaries are denoted by horizontal arrows. DNA sequences 5′ to the transcription start sites, which resemble signals for transcriptional regulation of other eukaryotic genes, are boxed (TATA, CAAT and SV40 enhancer core), and elements A–D putatively involved in the specific regulation of PAL transcription are denoted by solid lines. The underlined region E adjacent to the 3′ end of the first exon is the 17 bp insert present in the gPAL3 clones 3 and 12, but absent from clone 14. Boxed regions at the 3′ end of the coding regions are the translation termination codons and those sequences used for the construction of gene-specific oligonucleotides complementary to PAL mRNA. (B) Comparison of the deduced amino acid sequences of the three classes of bean PAL. The vertical arrow denotes the residue encoded by the codon which contains the intron. Asterisks denote the start point of specific sequences.

nucleotides 1660 and 1661, which are present within the consensus junction sequence AGGT (Fig. 2A). The 149 bp fragment that is more weakly protected in this experiment, can be ascribed to the effects of the 5 T residues immediately following the boundary G residues on the stability of this terminus of the duplex. Two regions of the 3.5 kb *Eco* RI-*Hin*d III fragment of gPAL2 (Fig. 1) are likewise protected from S1 nuclease digestion. In addition to the 155 bp frag-

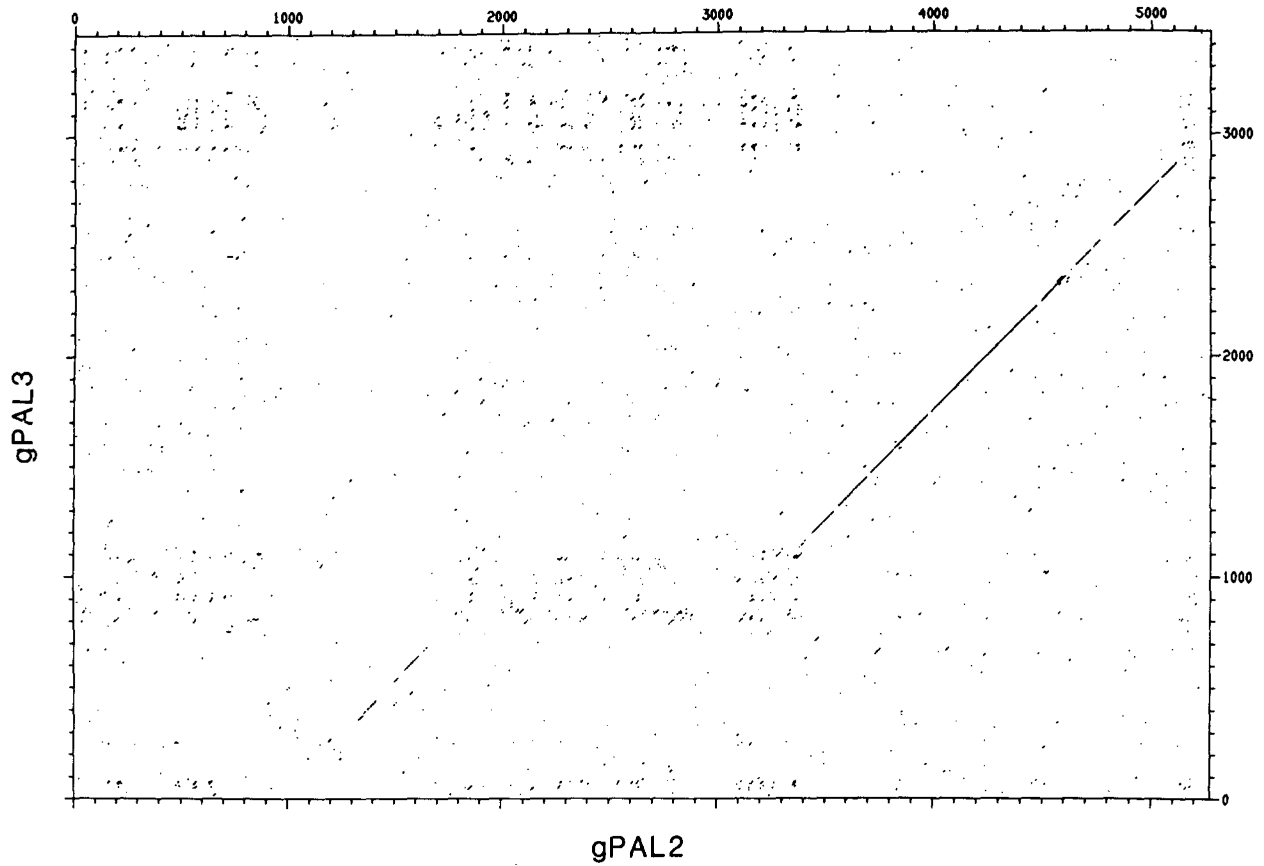

Fig. 3. Dot matrix comparison of the nucleotide sequences of gPAL2 and gPAL3. Window = 21 bp; stringency = 14 matches.

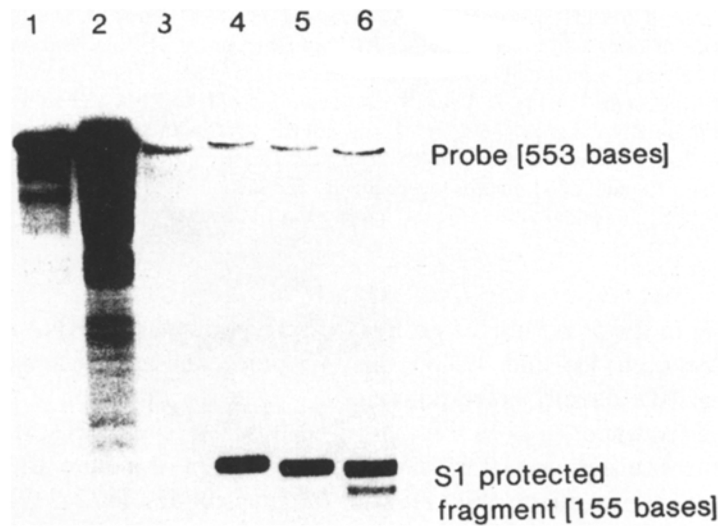

Fig. 4a. See p. 374.

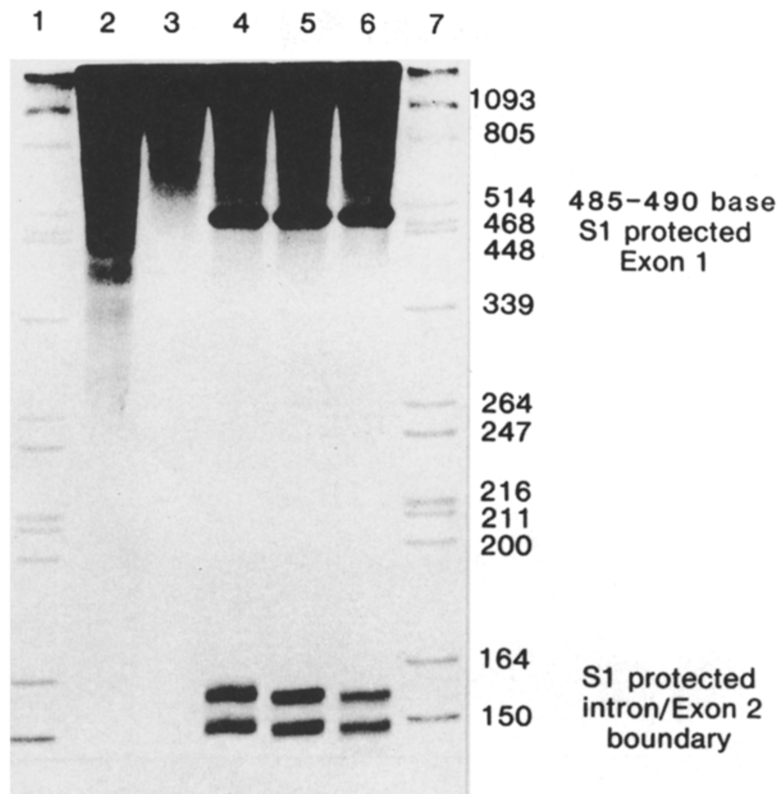

Fig. 4b.

Fig. 4. Identification of intron-exon junctions and determination of the size of exon I in gPAL2 by S1 nuclease protection. A. The boundary between the intron and exon II was located by partial protection of a single-stranded DNA fragment derived by subcloning the 536 bp *Hinc* II-*Hin*d III fragment of gPAL2 (see Fig. 1) into M13mp8. B. The size of exon I was determined by the S1 nuclease protection of a single-stranded 3.5 kb *Eco* RI-*Hin*d III fragment. [^{32}P]-labeled single-stranded DNA probes were synthesized using M13 universal primer and strand-purified on denaturing gels [3] prior to hybridization to total cellular RNA isolated from wounded hypocotyls. In panel A, Lanes 1 and 2: controls with no RNA and no S1 nuclease; Lane 3: control with 1 unit of S1 nuclease but no RNA; Lanes 4–6: 5 μg of total cellular RNA with 1, 5 and 10 units of S1 nuclease respectively. In panel B, Lanes 1 and 7: size markers; Lane 2: no RNA, no S1 nuclease controls; Lane 3: no RNA, 5 units of S1; Lanes 4–6: 5 μg of RNA with 1, 5 and 10 units of S1 nuclease respectively. The second band of 149 bases may reflect the presence of 5 T residues following the intron-exon I boundary nucleotide G.

ment corresponding to the 5′ portion of exon II, a fragment sized between 485 and 490 bp was also protected (Fig. 4B), directly corresponding to the entire exon I sequence.

To determine the transcription start site of gPAL2, a 494 bp *Rsa* I fragment (Fig. 1) was inserted into M13 and DNA probes for both S1 nuclease protection and sequence reaction marker lanes were generated using an internal synthetic oligonucleotide primer (see legend, Fig. 5). Hybridization to RNA from wounded hypocotyls or from elicitor-treated cell cultures resulted in S1 nuclease protection of several fragments 187–190 bases in length (Fig. 5) and established the transcription start site within the sequence CATT (nucleotides 1172–1175), as indicated in Fig. 2. Comparison with the eukaryotic transcription start site consensus sequence Py--Py A(+1) Py Py Py Py [7], suggests that the A at nucleotide 1173 is the first transcribed base. The

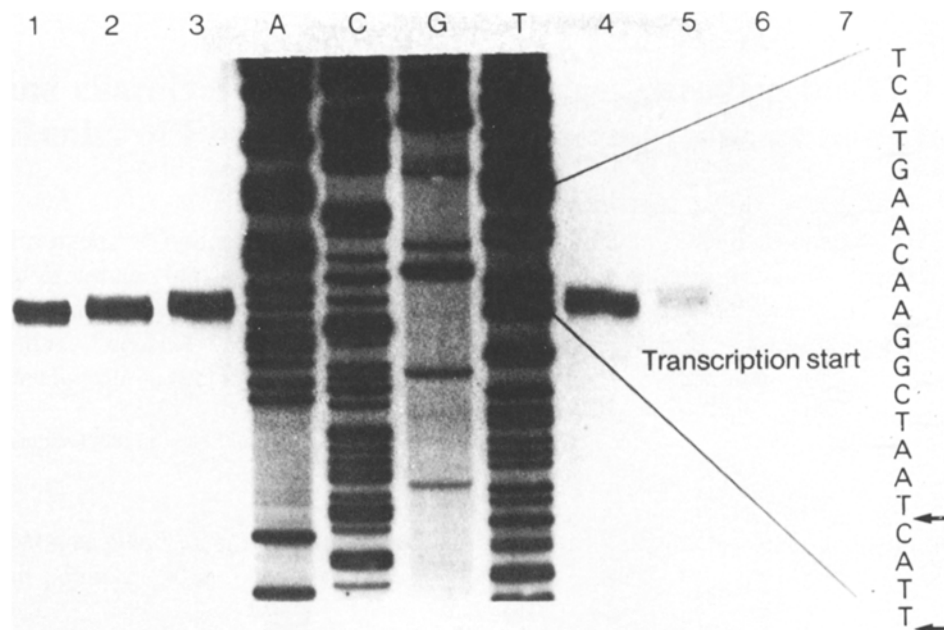

Fig. 5. Identification of the transcription initiation site of gPAL2 by S1 nuclease protection. The hybridization probe was a single-stranded DNA fragment prepared as follows. A 494 *Rsa* I fragment (nucleotides 918–1412 as shown in Fig. 1) was cloned into the *Sma* I site of M13mp8. A 21 base oligonucleotide (5′-CCGAGAGCGCCTCCGCCGCAG-3′, complementary to gPAL2 nucleotides 1352–1373) was synthesized and used to prime DNA synthesis from this subclone to generate DNA probes for S1 nuclease protection experiments [3]. The same primer was used for dideoxy sequencing reactions to generate the marker lanes A, C, G, and T, allowing direct identification of the limit of the protected sequences. The DNA probe was cut within vector sequences by *Eco* RI and labeled single-stranded DNA was purified on denaturing gels prior to hybridization with 25 μg RNA. Lanes 1–4: RNA from wounded hypocotyls; Lane 5: RNA from elicitor-treated cell cultures; Lane 6: RNA from control, uninduced cell cultures. Samples were treated with 5 units of S1 nuclease in lane 1; 2.5 units in lane 2; 1 unit in all other lanes.

first ATG is 99 bp downstream from the transcription start and initiates the open reading frame of exon I (Fig. 2).

The S1 nuclease protection studies locate the exon I-intron junction between 485 and 490 bp downstream from the transcription start site in the sequence TCAGGT (Fig. 2A). Utilization of the consensus AGGT junction site at this location establishes sizes of 488 bp for exon I, and 1720 bp for the intron. Thus gPAL2 contains an open reading frame consisting of 389 bp of exon I and 1747 bp of exon II, which encodes a polypeptide of 712 amino acids (M_r = 77849 Da; Fig. 2).

The transcription start site of gPAL3 was established by S1 nuclease protection experiments using a 436 bp *Bam* HI-*Hpa* II fragment of clone 14, that contains sequences from exon I and the 5′ flanking region (Figs. 1 and 6). RNA from wounded hypocotyls protected a 173 bp portion of the *Bam* HI-*Hpa* II fragment (Fig. 3), defining a transcription start site 35 bp upstream from the translation start site of the exon I open reading frame, which is conserved between gPAL2 and gPAL3. The precise size of the 173 bp protected fragment was confirmed on gels with markers generated by Maxam and Gilbert sequencing reactions of the *Hpa* II-*Hin*d III fragment of gPAL3 (data not shown). In contrast to gPAL2, gPAL3 sequences are not protected by RNA from elicitor-treated cell cultures (Fig. 6), indicating differential regulation of gPAL3 compared to gPAL1 (pPAL5 cDNA) and gPAL2.

Sequence comparisons between gPAL2 clone 2 and gPAL3 clone 14 showed conservation of the intron-exon II junction, but not the exon I-intron junction. However, sequence analysis of two other independently isolated Class III

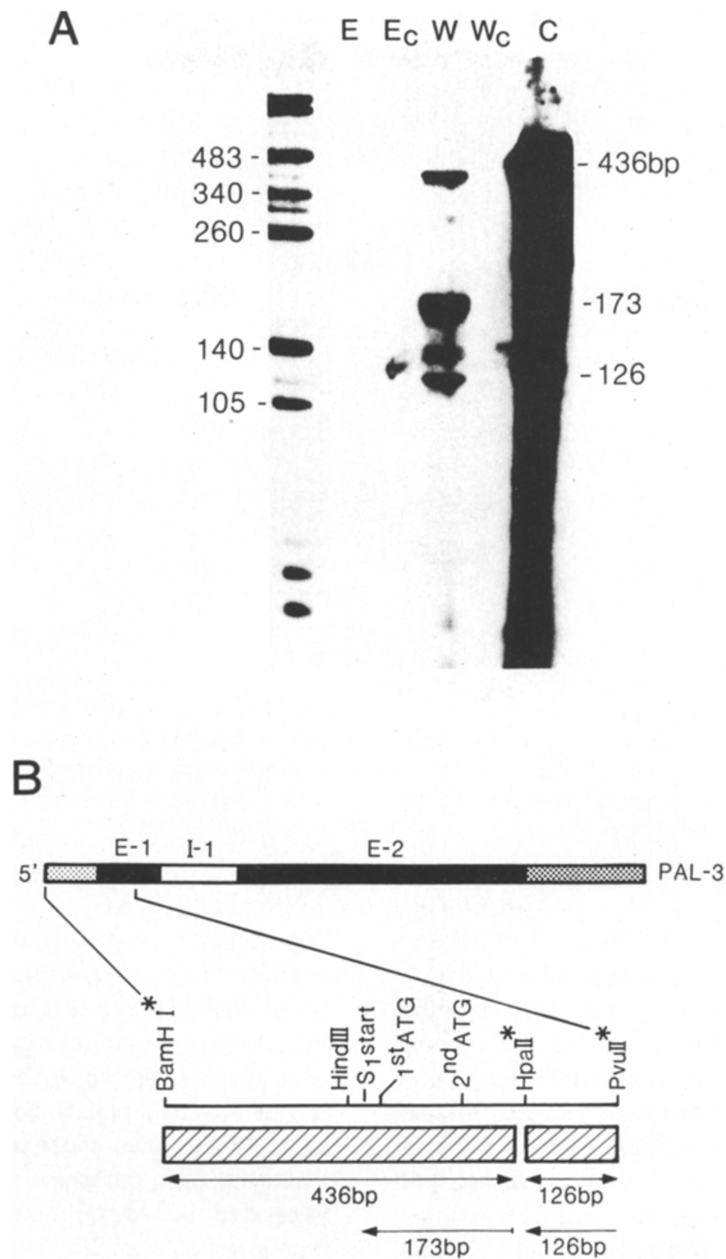

Fig. 6. Identification of the transcription initiation site of gPAL3 by S1 nuclease protection. A. A mixture of *Bam* HI-*Hpa* II (436 bp) and *Hpa* II-*Pvu* II (126 bp) fragments [^{32}P]-labeled at the 5′ end with T4 polynucleotide kinase, were digested with S1 nuclease after hybridization to 100 μg of total RNA from: control suspension cultured cells (Ec); suspension cultured cells 4 h after addition of fungal elicitor (E); hypocotyls 7 h after wounding (W); RNA from untreated hypocotyls (Wc); and probe without RNA or S1 nuclease (C). B. Map of the probe fragments. The 126 bp internal fragment serves as a full length protected control.

clones (Table 1: clone 3, cv. Canadian Wonder, and clone 12, cv. Tendergreen), which did not contain full-length gPAL3 genes, revealed the presence of an additional 17 bp at the 3′ end of exon I (region E, Fig. 2A), which restores the reading frame and intron junction with respect to gPAL2. Clone 3, a 3′ truncated gene, showed complete sequence identity to clone 14 in all other regions analyzed, including all of the 5′ untranslated sequence, exon I, the intron, and the 5′ end of exon II. Clone 12, a 5′ truncated gene, also showed complete sequence identity (with the exception of the 17 bp insert) in all regions analyzed, including most of exon I, the intron, exon II, and 3′ untranslated regions. gPAL3 clone 14 may therefore contain a recent deletion and represent an inactive allele of Class III PAL genes (discussed further below). Class III genes (including the 17 bp insertion in clones 3 and 12) have an open reading frame which contains 389 bp of exon I and 1741 bp of exon II, encoding a polypeptide of 710 amino acids with a predicted M_r of 77364 Da (Fig. 2). The intron is 447 bp and is located at the same site as in gPAL2.

Exon I is not highly conserved between gPAL2 and gPAL3, with only 59% and 50% sequence similarity at the nucleic acid and protein levels respectively. Adjacent to the *N*-terminus there is a stretch of 16 amino acid residues that are completely divergent between the two deduced polypeptides (Fig. 2A). The introns do not show sequence conservation, although both are AT-rich. Exon II of gPAL2 and gPAL3 show 74% and 77% similarity at the nucleic acid and protein levels respectively. The partial sequence of exon II of gPAL1 (equivalent to cDNA pPAL5) is more similar to the corresponding region of gPAL2 than gPAL3 (90% and 76% similarity at the amino acid level, respectively). Within exon II, the 5′ half is more highly conserved at both the nucleic acid and protein levels than the downstream regions (Figs. 2 and 3). The 5′ untranslated regions of gPAL2 and gPAL3 show little sequence similarity (Figs. 2 and 3). Likewise the nucleotide sequences 3′ of the translation stop codons of gPAL1 (pPAL5 cDNA), gPAL2, and gPAL3 are highly divergent.

Promoter sequences

Presumptive TATA and CAAT boxes [7] are located 32 and 74 bp respectively upstream from the transcription start site of gPAL2, and 30 and 73 bp upstream from the transcription start site of gPAL3 (Fig. 2). The 5′ flanking sequences of gPAL2 and gPAL3 contain copies of the element $^{\mathrm{TTT}}_{\mathrm{AAA}}$CCAC, which is an inverted version of the GTGG$^{\mathrm{AAA}}_{\mathrm{TTT}}$ element that comprises the SV40 enhancer core sequence [39]. There is a 15 bp AC-rich element (A) 78 bp upstream of the transcription start site of gPAL3 that is almost identical to a sequence motif found in the promoters of coordinately expressed bean genes encoding the phenylpropanoid biosynthetic enzyme chalcone synthase, which catalyzes the first reaction in a branch pathway specific for the synthesis of flavonoid pigments and isoflavonoid phytoalexins [18]. A 13 bp variant of this sequence is also found 121 bp upstream of the transcription start site of gPAL2. Otherwise the 5′ flanking regions of gPAL2 and gPAL3 sequenced to date exhibit little similarity. gPAL2 contains an AT repetitive sequence (B) and a 40 bp AT-rich palindromic sequence (C) 297 bp upstream of the transcription start site. The latter is part of a region containing a series of overlapping motifs that are also found in a similar relative position in the 5′ flanking regions of bean chalcone synthase genes [18]. A second region (D), 73 bp upstream from the gPAL2 transcription start site, also resembles a sequence at a similar relative position in the chalcone synthase genes.

Gene organization

To analyze the organization of PAL genes within the bean genome, Southern blots of genomic DNA from cv. Tendergreen and cv. Canadian Wonder were hybridized with various PAL sequences including gene-specific cloned fragments and synthetic oligonucleotides. Under conditions of relatively low stringency, the insert of cDNA clone pPAL5 hybridized to several fragments of cv. Canadian Wonder (Fig. 7) or cv. Tendergreen

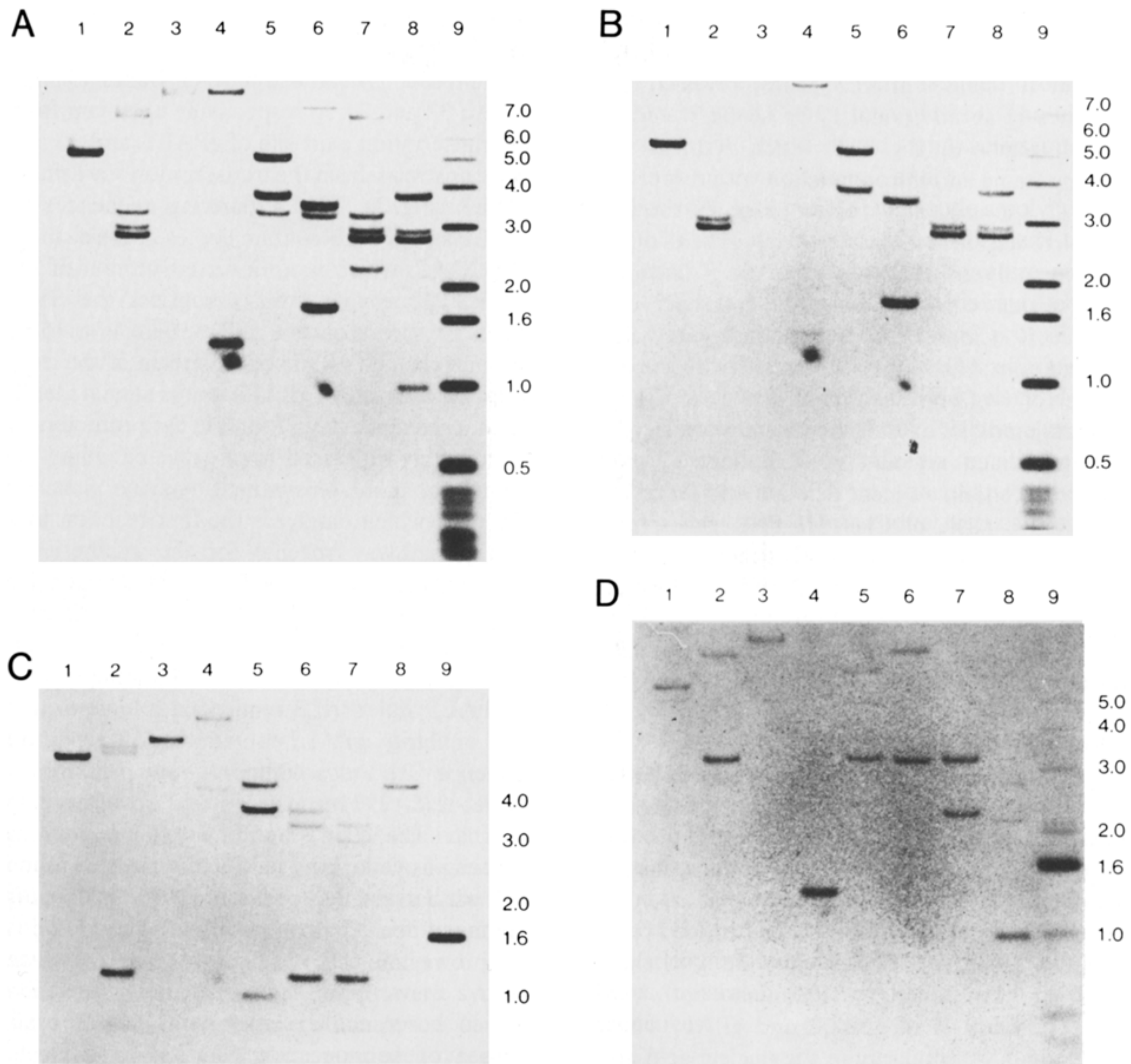

Fig. 7. Organization of PAL genes in the genome of bean cv. Canadian Wonder. Total genomic DNA from seedlings of cv. Canadian Wonder (5 μg/lane) was digested with restriction endonucleases prior to agarose gel electrophoresis, blotting and hybridization with nick-translated, gel-purified probes. A. Low-stringency hybridization with the 5.3 kb *Eco* RI fragment containing gPAL2. B. High-stringency hybridization with the 5.3 kb *Eco* RI gPAL2 fragment. C. High-stringency hybridization with the 5.8 kb *Eco* RI gPAL3 fragment. D. High-stringency hybridization with the 1.7 kb insert of pPAL5, a cDNA clone corresponding to gPAL1. Restriction endonucleases were: Lane 1: *Eco* RI; Lane 2: *Hind* III; Lane 3: *Bam* HI; Lane 4: *Bgl* II; Lane 5: *Hinc* II; Lane 6: *Eco* RI plus *Hind* III; Lane 7: *Hind* III plus *Bam* HI; Lane 8: *Hind* III plus *Hinc* II; Lane 9: molecular weight markers.

(data not shown) genomic DNA digested with *Eco* RI or several other restriction endonucleases. Hybridization to cv. Tendergreen genomic DNA under conditions of high stringency with cloned fragments from exon II of gPAL1, gPAL2 and gPAL3 showed that the 5.3 kb *Eco* RI fragment contained gPAL2 sequences (Fig. 8). Sequences from exon II of gPAL3 hybridized specifically to a 6.0 kb *Eco* RI fragment that was not hybridized by pPAL5 cDNA sequences, consistent with the

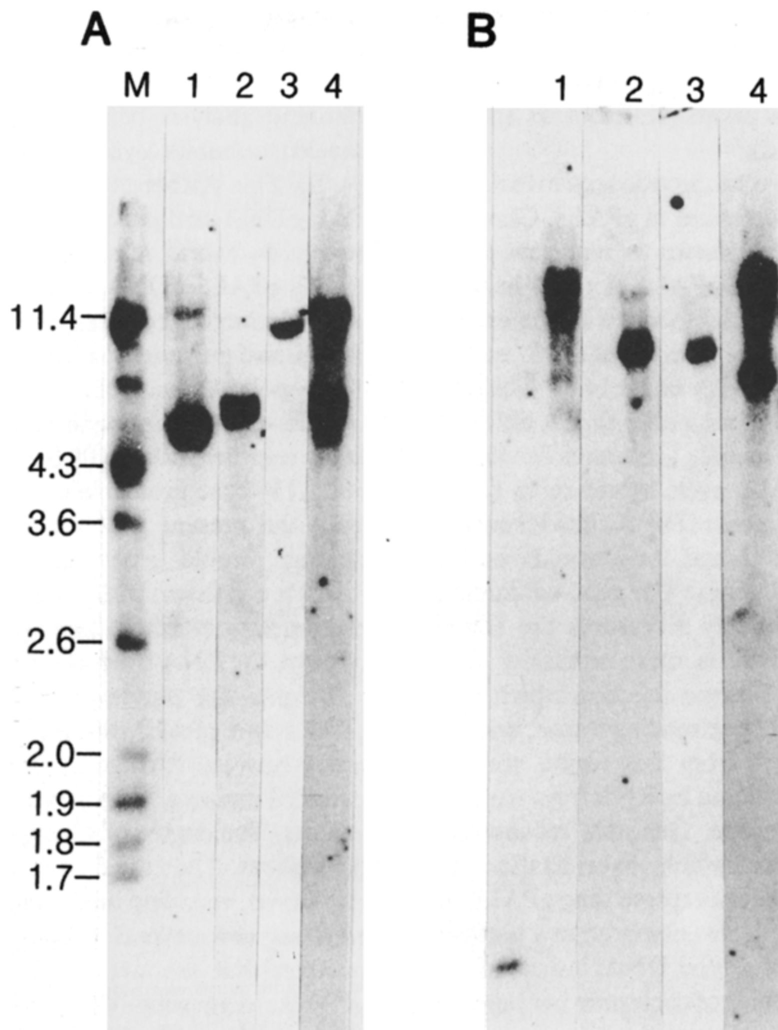

Fig. 8. Organization of PAL genes in the genome of bean cv. Tendergreen. Total DNA from bean cv. Tendergreen (5 μg DNA/lane) was digested with *Eco* RI (A) or *Ava* I (B) prior to agarose gel electrophoresis, blotting and hybridization with nick-translated, gel-purified PAL sequences. Lane 1: gPAL2 (*Eco* RI-*Ava* II 1.33 kb fragment); Lane 2: gPAL3-5′ (562 bp *Bam* HI-*Pvu* II fragment); Lane 3: gPAL3-3′ (1.26 kb *Eco* RI-*Bam* HI fragment); Lane 4: pPAL5 (1.7 kb cDNA insert).

greater sequence divergence of gPAL3 (see above). Sequences from exon II of gPAL1 hybridized to both the 6.5 and 11.4 kb *Eco* RI fragments. This indicates the presence of a second class I gene, gPAL1′, located on the 11.4 kb *Eco* RI fragment, that is closely related to gPAL1, known from nucleotide sequence analysis of genomic clones to reside on the 6.5 kb *Eco* RI fragment (see above). These assignments were confirmed by hybridization with labeled, gene-specific synthetic oligonucleotides (Fig. 2) from a region of exon II that is not highly conserved between the 3 classes of PAL genes (data not shown).

A very similar overall organization of PAL genes was found in the genome of cv. Canadian Wonder, although only weak hybridization was observed to the 11.4 kb *Eco* RI fragment representing the class I polymorphic form containing gPAL1′ sequences (Fig. 7). Digestion of bean genomic DNA with other enzymes, or by combinations of enzymes, revealed additional fragments for each of the PAL gene classes, indicating

that the PAL gene family is polymorphic for all representative genes (Fig. 7). This was confirmed by Southern blot analysis of genomic DNA isolated from individual plants (K. Edwards and W. Schuch, unpublished).

As described above, nucleotide sequence analysis revealed an allelic variant of gPAL3. Class III clones 3 and 12 were shown to represent truncated genes identical to gPAL3 clone 14 in both coding and untranslated regions, with the exception of a 17 bp sequence inserted at the 3′ end of exon II (Fig. 2A). gPAL3 clone 14 is likely to represent an allelic variant rather than a different locus because gene-specific sequences for the 3′ and 5′ ends of gPAL3 each hybridize to single *Eco* RI genomic fragments (Fig. 8). *Eco* RI cuts in the middle of exon II and thus would resolve tandemly duplicated genes. The gene containing the 17 bp insert probably represents the active allele because (a) there is close similarity with gPAL2 at the exon I-intron junction which preserves the conserved open reading frame; and (b) fragments of clone 14 from this region are not protected from S1 nuclease by RNA from wound-induced hypocotyl tissue. Genomic reconstruction experiments, in which hybridization to genomic DNA sequences representing gPAL3 or gPAL1 was quantitatively compared to a haploid genome equivalent of cloned DNA, indicated the presence of 1 to 2 copies of each gene per haploid genome (data not shown).

Discussion

The present data demonstrate that PAL is encoded by a small gene family in the bean genome comprising: gPAL1 which corresponds to the previously characterized cDNA sequences [20]; gPAL2 and gPAL3, the complete nucleotide sequences of which have been determined; and gPAL1′ which is closely related to gPAL1, but which has not been characterized further. gPAL2 and gPAL3 exhibit very similar organization in terms of the number and size of exons, location of the single intron, and size of the polypeptides encoded by the respective open reading frames. Deduced M_rs of 77364 and 77849 respectively are consistent with a size for PAL polypeptides of 77–78 kDa as previously determined by electrophoretic analysis of immunoprecipitable [^{35}S]-labeled subunits synthesized *in vivo* or *in vitro* [4, 5]. This further confirms the identity of bean PAL cDNA and genomic clones initially characterized by hybrid selected translation [20].

The pPAL5 cDNA contains a 192 bp nucleotide sequence between the translation termination and polyadenylation sites, together with a 37 bp poly(A) tail [20]. A 3′ untranslated sequence of the same size in gPAL2 transcripts, taken together with the 99 base 5′ leader sequence and 2136 base protein coding sequence deduced from the present genomic nucleotide sequence analysis, would generate a gPAL2 mRNA of 2.46 kb, consistent with a size of 2.5 kb previously estimated by Northern blot hybridization analysis of poly(A)$^+$ RNA from elicitor-treated cells [20].

Despite the similar overall organization of gPAL2 and gPAL3, there is considerable divergence between PAL genes in the nucleotide and deduced amino acid sequences within the coding regions. The degree of divergence is not uniform throughout. Thus the 5′ half of exon II is highly conserved, encoding one stretch of 33 amino acid residues and several stretches greater than 20 residues that are identical between gPAL2 and gPAL3, suggesting this region is critical for catalysis. In marked contrast, exon I of gPAL2 and gPAL3 and the 3′ half of exon II of gPAL1 (cDNA), gPAL2 and gPAL3 are more divergent. This may provide the basis for regulatory and functional variants of the gene product. PAL isozymes with different patterns of inhibition by phenylpropanoid intermediates and end products have been described [1, 6, 23]. Several isoforms of PAL polypeptide subunits and native tetrameric enzyme have been characterized in bean. The isoforms of the native enzyme exhibit different K_ms for phenylalanine and the forms with low K_m are preferentially induced by fungal elicitor, thereby exerting a metabolic priority for phenylpropanoid synthesis in the cellular economy of phenylalanine specifically under conditions of stress [4].

The high degree of sequence divergence between classes of PAL genes, especially in non-coding regions, allows application of S1 nuclease protection analysis to the detection of isogene-specific PAL mRNAs. The present data (Figs. 5 and 6) provide evidence for differential regulation of PAL at the gene level, since both gPAL1 (cDNA) and gPAL2 are induced by elicitor treatment of cv. Canadian Wonder cell suspension cultures, whereas gPAL3 is not responsive to elicitor treatment of these cells. Moreover, RNase protection experiments have shown that while all these genes are induced by wounding of hypocotyls, there are marked differences between members of the gene family in their response to other environmental stimuli such as irradiation or infection, and in their organ-specific patterns of regulation during development (X. Liang, R. A. Dixon and C. J. Lamb, unpublished). Temporal differences in the expression of gPAL1 and gPAL2 in elicitor-treated cells have also been observed (K. Edwards and W. Schuch, unpublished).

The present data provide no evidence for clustering of members of the PAL gene family within the bean genome. This is in marked contrast to the extensive clustering within the family of 6–8 bean chalcone synthase genes [34]. Moreover, members of the bean chalcone synthase gene family show a very high degree of conservation of nucleotide sequence both within coding regions and the single intron (S. D. Clouse, J. Schmid, T. B. Ryder, R. A. Dixon and C. J. Lamb, unpublished), again in marked contrast to the quite substantial divergence between the coding regions of different PAL genes, and the extensive divergence of PAL intron sequences. Taken together these data suggest that gene duplication and familial divergence started much earlier in evolution for PAL compared to chalcone synthase. The ancient divergence of PAL genes, the emerging complex patterns of PAL gene regulation, and the apparent biochemical specialization of the encoded isopolypeptides, may be related to the highly diverse biological functions of phenylpropanoid natural products elaborated from the PAL reaction product cinnamic acid.

By analogy with other inducible plant and animal genes, the induction of PAL genes during development and in response to environmental stimuli is likely to be controlled, at least in part, by *cis*-acting regulatory elements located 5′ of the transcription initiation sites [7]. The 5′ flanking regions of both gPAL2 and gPAL3 contain features common to the promoters of other eukaryotic genes, including TATA and CAAT boxes about 30 and 74 bases upstream from the transcription start site [7]. In addition, both genes contain elements resembling the SV40 viral enhancer core [39], which is important in the transcriptional regulation of a number of eukaryotic genes [7, 39]. Similar sequences have been identified in the promoters of several plant genes encoding ribulose 1,5-bisphosphate carboxylase [2, 8, 11], proteinase inhibitors [29], chalcone synthase [18] (S. D. Clouse, R. A. Dixon and C. J. Lamb, unpublished), zein [29], octopine synthase [28] and alcohol dehydrogenase [21]. Moreover, the AT-repetitive element (B) in gPAL2 is reminiscent of certain upstream enhancer elements in yeast [37] and nuclear scaffold attachments in *Drosophila* [22]. The functional significance of these elements in plants has not yet been determined.

Coordinately regulated genes often exhibit common sequence motifs in their respective promoters that confer a specific pattern of regulation during development or in response to environmental stimuli [7, 29]. Overall, there is extensive divergence between the 5′ flanking sequences of gPAL2 and gPAL3 determined to date. This is consistent with the emerging evidence that these genes exhibit substantially different patterns of expression. However, an AC-rich motif (A) has been identified that is present in the promoters of gPAL2, gPAL3 and bean chalcone synthase genes, and hence is likely to be involved in placing this set of genes in a specific regulatory network. Likewise, sequences similar to the elements (C) and (D) in gPAL2 are also found, organized in like manner, in the promoters of coordinately regulated chalcone synthase genes [18] and hence may represent putative *cis*-acting regulatory sequences. Analysis of the effects of deletions in

the 5′ flanking region of chalcone synthase genes has shown that a 120 bp region containing the sequence similar to the 40 bp AT-rich palindrome present in gPAL2 (C) binds *trans*-acting factors involved in regulation by elicitor [18] (M. A. Lawton and C. J. Lamb, unpublished). Further gene transfer experiments will be required to define precisely the role of the palindromic sequence and other putative *cis*-acting elements in the regulation of genes encoding PAL and other enzymes of phenylpropanoid biosynthesis.

Acknowledgements

This work was supported by grants from Imperial Chemical Industries plc (R.A.D., C.J.L.), the Science and Engineering Research Council (R.A.D.), the Samuel Roberts Noble Foundation (C.J.L.) and Tan Sri Tan (C.J.L.). C.L.C. was an NSF Fellow in Plant Biology and M.D. was a NATO Fellow. We thank Valerie Zatorski and Cynthia Doane for preparation of the manuscript.

References

1. Alibert G, Ranjeva R, Boudet A: Recherches sur les enzymes catalysant la formation des acides phénoliques chez *Quercus pedunculata* (Ehrh.) II. Localisation intracellulaire de la phenylalanine ammoniaque-lyase, de la cinnamate 4-hydroxylase, et de la 'benzoate synthase'. Biochim Biophys Acta 279: 282–289 (1972).
2. Berry-Lowe SL, McKnight TD, Shah DM, Meager RB: The nucleotide sequence, expression, and evolution of one member of a multigene family encoding the small subunit of ribulose 1,5-bisphosphate carboxylase in soybean. J Mol Appl Genet 1: 483–498 (1982).
3. Biggin M, Farrell PJ, Barrel BG: Transcription and DNA sequence of the *Bam* HI L fragment of B95-8 Epstein-Barr virus. EMBO J 3: 1083–1090 (1984).
4. Bolwell GP, Bell JN, Cramer CL, Schuch W, Lamb CJ, Dixon RA: L-Phenylalanine ammonia-lyase from *Phaseolus vulgaris*: Characterization and differential induction of multiple forms from elicitor-treated cell suspension cultures. Eur J Biochem 149: 411–419 (1985).
5. Bolwell GP, Sap J, Cramer CL, Lamb CJ, Schuch W, Dixon RA: L-Phenylalanine ammonia-lyase from *Phaseolus vulgaris*: Partial degradation of enzyme subunits *in vitro* and *in vivo*. Biochim Biophys Acta 881: 210–221 (1986).
6. Boudet A, Ranjeva R, Gadal P: Proprietés allostériques spécifiques des deux isoenzymes de la phénylalanine ammoniaque-lyase chez *Quercus pedonculata*. Phytochemistry 10: 997–1005 (1971).
7. Breathnach U, Chambon P: Organization and expression of eukaryotic split genes coding for proteins. Annu Rev Biochem 50: 349–384 (1981).
8. Broglie R, Coruzzi G, Lamppa G, Keith B, Chua N-H: Structural analysis of nuclear genes coding for the precursor to the small subunit of wheat ribulose 1,5-bisphosphate carboxylase. Biotechnology 1: 55–61 (1983).
9. Chappell J, Hahlbrock K: Transcription of plant defense genes in response to UV light or fungal elicitor. Nature 311: 76–78 (1984).
10. Chomczynski P, Qasba PK: Alkaline transfer of DNA to plastic membranes. Biochem Biophys Res Comm 122: 340–344 (1984).
11. Coruzzi G, Broglie R, Edwards C, Chua N-H: Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose 1,5-bisphosphate carboxylase. EMBO J 3: 1671–1679 (1984).
12. Cramer CL, Bell JN, Ryder TB, Bailey JA, Schuch W, Bolwell GP, Robbins MP, Dixon RA, Lamb CJ: Coordinated synthesis of phytoalexin biosynthetic enzymes in biologically-stressed cells of bean (*Phaseolus vulgaris* L.). EMBO J 4: 285–289 (1985).
13. Cramer CL, Ryder TB, Bell JN, Lamb CJ: Rapid switching of plant gene expression by fungal elicitor. Science 227: 1240–1243 (1985).
14. Dale RMK, McClure BA, Houchins JP: A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: Application to sequencing the corn mitochondrial 18S rDNA. Plasmid 13: 31–40 (1985).
15. Dixon RA, Dey PM, Lamb CJ: Phytoalexins: Molecular biology and enzymology. Adv Enzymol Relat Areas Mol Biol 55: 1–135 (1983).
16. Dixon RA, Bailey JA, Bell JN, Bolwell GP, Cramer CL, Edwards K, Hamdan MAMS, Lamb CJ, Robbins MP, Ryder TB, Schuch W: Rapid changes in gene expression in response to microbial elicitation. Phil Trans Royal Soc B 314: 411–426 (1986).
17. Downie JA, Johnston AWB: Nodulation of legumes by *Rhizobium*: The recognized root? Cell 47: 153–154 (1986).
18. Dron M, Clouse SD, Dixon RA, Lawton MA, Lamb CJ: Glutathione and fungal elicitor regulation of a plant defense gene promoter in electroporated protoplasts. Proc Natl Acad Sci USA 85: 6738–6742 (1988).
19. Ebel J: Phytoalexin synthesis: The biochemical analysis of the induction process. Annu Rev Phytopath 24: 235–264 (1986).
20. Edwards K, Cramer CL, Bolwell GP, Dixon RA, Schuch W, Lamb CJ: Rapid transient induction of phenylalanine ammonia-lyase mRNA in elicitor-treated bean cells. Proc Natl Acad Sci USA 82: 6731–6735 (1985).

21. Ellis JG, Llewellyn DJ, Dennis ES, Peacock WJ: Maize *Adh-1* promoter sequences control anaerobic regulation: Addition of upstream promoter elements from constitutive genes is necessary for expression in tobacco. EMBO J 6: 11–16 (1987).
22. Gasser SM, Laemmli UK: A glimpse at chromosomal order. Trends Genet 3: 16–22 (1987).
23. Jones DH: Phenylalanine ammonia-lyase: Regulation of its induction, and its role in plant development. Phytochemistry 23: 1349–1359 (1984).
24. Kuhn DH, Chappell J, Boudet A, Hahlbrock K: Induction of phenylalanine ammonia-lyase and 4-coumarate: CoA ligase mRNAs in cultured plant cells by UV light or fungal elicitor. Proc Natl Acad Sci USA 81: 1102–1106 (1984).
25. Lawton MA, Lamb CJ: Transcriptional activation of plant defense genes by fungal elicitor, wounding and infection. Mol Cell Biol 7: 335–341 (1987).
26. Leder P, Tiemeier D, Enquist L: EK2 derivatives of bacteriophage lambda useful in cloning of DNA from higher organisms: The λgtWES system. Science 196: 175 (1977).
27. Lee JS, Brown WE, Graham JS, Pearce G, Fox EA, Dreher TW, Aher, KG, Pearson GD, Ryan CA: Molecular characterization and phylogenetic studies of a wound-inducible proteinase inhibitor 1 gene in *Lycopersicon* species. Proc Natl Acad Sci USA 83: 7277–7281 (1986).
28. Leisner SM, Gelvin SB: Structure of the octopine synthase upstream activator sequence. Proc Natl Acad Sci USA 85: 2553–2557 (1988).
29. Maier V-G, Brown JWS, Toloczyki C, Feix G: Binding of a nuclear factor to a consensus sequence in the 5′-flanking region of zein genes from maize. EMBO J 6: 17–22 (1987).
30. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
31. Maxam AM, Gilbert W: Sequencing end-labeled DNA with base specific chemical cleavages. Methods Enzymol 65: 499–560 (1980).
32. Melton DA, Krieg PA, Rebagliati MR, Maniatis T, Zinn K, Green MR: 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 (1984).
33. Murray MG, Thompson WF: Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8: 4321–4325 (1980).
34. Ryder TB, Hedrick SA, Bell JN, Liang X, Clouse SD, Lamb CJ: Organization and differential activation of a gene family encoding the plant defense enzyme chalcone synthase. Mol Gen Genet 210: 219–233 (1987).
35. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
36. Stachel SE, Zambryski PC: *Agrobacterium tumefaciens* and the susceptible plant cell: A novel adaptation of extracellular recognition and DNA conjugation. Cell 47: 155–157 (1986).
37. Struhl K: Genetic properties and chromatin structure of the yeast *gal* regulatory element: An enhancer-like sequence. Proc Natl Acad Sci USA 81: 7865–7869 (1984).
38. Sun SM, Slightom JL, Hall TC: Intervening sequences in a plant gene: Comparison of the partial sequence of cDNA and genomic DNA of French bean phaseolin. Nature 289: 37–40 (1981).
39. Weiher H, König M, Gruss P: Multiple point mutations affecting the simian virus 40 enhancer. Science 219: 626–631 (1983).
40. Wood WI, Gitschier J, Lasky LA, Lawn RM: Base composition-independent hybridization in tetramethylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries. Proc Natl Acad Sci USA 82: 1585–1588 (1985).