

The tomato polygalacturonase gene and ripening-specific expression in transgenic plants

C. R. Bird,¹ C. J. S. Smith,² J. A. Ray,¹ P. Moureau,^{1, 3} M. W. Bevan,⁴ A. S. Bird,¹ S. Hughes,² P. C. Morris,² D. Grierson^{2*} and W. Schuch^{1*}

¹ICI Seeds, Plant Biotechnology Section, PO Box 11, The Heath, Runcorn, Cheshire, WA7 4QE, England;

²Department of Physiology and Environmental Science, University of Nottingham, School of Agriculture, Sutton Bonington, Loughborough LE12 5RD, England; ³Present address: University of Louvain, Avenue Hippocrate 54, UCL 34.90, B-1200 Brussels, Belgium; ⁴Institute of Plant Science Research, Maris Lane, Trumpington, Cambridge CB2 2LQ, England (*author for correspondence)

Received 27 April 1988; accepted in revised form 25 August 1988

Key words: tomato fruit ripening, polygalacturonase, gene sequence, transgenic tomato, fruit-specific expression

Abstract

Polygalacturonase (PG) is the major cell wall degrading enzyme of tomato fruit. It is developmentally regulated and is synthesised *de novo* in ripening fruit. Genomic clones encoding a PG gene of tomato (*Lycopersicon esculentum* Mill cv. Ailsa Craig) have been isolated, mapped and sequenced. The sequence of the protein-coding region is identical to that of a PG cDNA [20]. Comparison of the cloned restriction fragments with genomic Southern data suggests that there may only be one gene for PG per haploid genome. The PG gene, which covers approximately 7 kb, is interrupted by 8 intervening sequences ranging in size from 99 bp to 953 bp. The transcription start point was identified by S1 mapping and primer extension analysis. About 1.4 kb of 5' flanking DNA has been sequenced. This contains putative TATA and CAAT boxes and also direct repeat sequences. A transcriptional fusion has been constructed between the putative 1.4 kb promoter fragment and the chloramphenicol acetyl transferase (CAT) gene. Constructs containing this gene have been transferred to tomato using binary vectors. Regenerated transgenic plants express CAT in ripe tomato fruit, but not in unripe tomatoes, leaves, or roots.

Introduction

Tomato ripening is associated with a number of physiological and biochemical changes. These include degradation of chlorophyll, synthesis of flavour compounds, a major increase in the synthesis of ethylene, cell wall solubilisation and fruit softening. The major cell wall hydrolase in ripening tomato believed to be responsible for fruit softening is polygalacturonase (PG) [18, 22]. PG protein and

enzyme activity have been shown to increase dramatically at the onset of ripening, reaching a maximum in the ripe fruit [39, 38]. PG mRNA also increases at the onset of ripening reaching a maximum in orange or firm red fruit [11, 28]. PG gene expression therefore appears to be regulated at the level of mRNA accumulation during fruit ripening.

PG protein has been isolated from ripe tomatoes and shown to be associated with the cell walls of the fruit [38]. There are three isoenzyme forms, (PG1,

PG2 (a and b)) which are believed to be structurally related [31, 39, 40, 43]. PG1 and a mixture of PG2a and 2b have molecular weights of 100 kDa and 42 kDa respectively when they are isolated under non-denaturing conditions [39]. A mixture of PG2a and 2b can be converted *in vitro* to a form resembling PG1 by a heat-stable non-dialysable factor [40], but the exact mechanism is unclear. On SDS-PAGE the subunit size of PG1 and PG2 was identical and was estimated to be 46 kDa. This compares with 41 828 Da for the unglycosylated protein deduced from the cDNA sequence [20]. The difference between the two isoenzymes of PG2 is thought to be the extent of glycosylation [43]. Different levels of activity during ripening have been observed for the PG1 and PG2 (a and b) isoenzymes [39]. The exact difference between the three isoenzymes is not clear. However, they may result from post-translational processing of a single polypeptide.

PG mRNA is produced in ripening fruit, but not in unripe tomatoes, leaves, or roots [28]. Thus, expression of this gene is developmentally regulated. PG enzyme activity is also found in flower abscission zones in tomato [41]. However, it is not known whether this enzyme is related to the PG isoenzymes present in fruit tissue.

We have previously isolated and sequenced a PG cDNA clone (pTOM6) [20] from a library prepared from ripe tomato mRNA [36]. The clone was identified by comparison of the amino acid sequence predicted from the cDNA with the amino acid sequence determined for the isolated PG2a isoenzyme. The sequence of a tomato PG cDNA has subsequently also been reported by Sheehy *et al.* [35]. The sequence of this clone is identical to pTOM6 apart from one base change in the 3' untranslated region.

In this paper we describe the isolation and characterisation of clones of tomato genomic DNA which encode PG. These data are consistent with the suggestion that there is only one PG gene in the tomato genome. A sequence from the 5' end of the PG gene has been used to construct a transcriptional fusion with the CAT reporter gene. Transgenic tomato plants containing this hybrid gene have been used to test the hypothesis that the 5' sequence is required for ripening-specific expression in tomato fruit.

Materials and methods

Plant material and extraction of plant nucleic acids

Tomato plants (*Lycopersicon esculentum* Mill cv Ailsa Craig) were grown in controlled environment chambers as described by Grierson *et al.* [19]. This cultivar was used for all experiments. Genomic DNA was extracted from leaves as described by Dellaportia *et al.* [13]. Tomato fruit RNA was isolated as previously described [19].

Construction and screening of genomic library

Tomato genomic DNA was partially digested with *Sau* 3A and size-fractionated on a 10–40% sucrose gradient. Fragments of between 10 and 15 kb were ligated into λ EMBL3 arms (Stratagene) as described in Frischauf *et al.* [17] and plated on *Escherichia coli* strain K803. Approximately 500 000 plaque-forming units (pfu) were screened with the radio-labelled *Pst* I insert of pTOM6 and the 5' *Pst* I–*Hind* III fragment of the cDNA [20]. Hybridising clones were purified to single plaques and phage DNA was prepared by the method of Maniatis *et al.* [27].

Restriction mapping and DNA sequencing

Restriction digests and Southern blotting of the phage DNA were performed as described in Maniatis *et al.* [27]. Fragments of the PG gene were cloned into M13 vectors and sequenced by the dideoxynucleotide chain termination method [34].

Genomic Southern blotting

5 μ g of digested genomic DNA were separated on agarose gels and blotted according to Chomczynski and Qasba [8]. The membranes were prehybridised in $5\times$ SSPE (0.9 M NaCl, 0.05 M NaH_2PO_4 pH 7.7, 0.5 mM $\text{Na}_2\text{-EDTA}$), 1% SDS, 100 μ g/ml denatured herring sperm DNA at 65 °C for 4 h and hybridised with the radio-labelled (Pharmacia

Oligo-labelling kit) [15] insert of the PG cDNA (pTOM6) in the same buffer for 18 h at 65 °C. The membranes were washed in $0.1 \times$ SSC (15 mM NaCl, 1.5 mM sodium citrate pH 7.0), 1% SDS at 65 °C and autoradiographed.

S1 mapping

Single-stranded probes were prepared by the 'prime-cut' method [14] as described by Biggin *et al.* [3]. The ethanol-precipitated probe was dried *in vacuo* and dissolved in 66% formamide, 0.5 M NaCl, 15 mM Pipes pH 6.4, 6 mM EDTA, 0.7 mg/ml calf liver tRNA. 10 μ l of probe was mixed with 3 μ l (1 μ g) poly(A)⁺ RNA from red tomato and placed at 95 °C for 60 s before being transferred to 52 °C for 18 h. Ice-cold S1 buffer (300 mM NaCl, 30 mM NaOAc pH 4.5, 10 mM ZnSO₄), was added together with 200 U S1 nuclease (Amersham) and incubated at 30 °C for 30 min. After phenol/chloroform extraction and ethanol precipitation, the precipitate was resuspended in 5 μ l TE and 4 μ l formamide containing 0.1% bromophenol blue and 0.1% xylene cyanol FF. After denaturing at 100 °C for 3 min the sample was electrophoresed on a 6% polyacrylamide/8 M urea sequencing gel. A sequence ladder of the 1.65 kb *Hind* III fragment from the 5' end of the gene in M13 mp19 was used as a size marker.

Primer extension

Oligonucleotide primer end-labelled with γ [³²P]ATP (6 ng) in 2 μ l 5 mM sodium phosphate pH 7.2, 5 mM EDTA was annealed to 1 μ g red tomato poly(A)⁺ RNA. Primer extension was performed according to Hudson *et al.* [25]. A sequence ladder of the *Hind* III fragment from the 5' end of the gene in M13 mp9 was used as a size marker.

Construction of pCB1

The construction of pCB1 is shown in Fig. 5. A *Spe* I restriction site was introduced between the transcription and translation start sites by changing the A at position 1446 to a T (see Figure 3). Site-directed

mutagenesis was performed on a clone of the 1.65 kb *Hind* III fragment in M13 mp9 by the double primer method [44] as described by Carter *et al.* [7]. After transformation into *E. coli* strain TG-2, mutagenised plaques were selected by hybridisation with the mutagenic oligonucleotide as described by Carter *et al.* [7]. The nucleotide sequence of the insert of the mutagenised clone was determined to check that only the required base change had occurred. Double-stranded DNA was prepared from single-stranded DNA *in vitro* as described in the M13 cloning and sequencing handbook (Amersham). The double-stranded DNA was digested with *Hind* III and *Spe* I and the 1.45 kb fragment was cloned into *Hind* III and *Xba* I cut pUC19. The 1.45 kb fragment was removed with *Hind* III and *Bam* HI and cloned into polyCATer [9]. The complete hybrid gene was sequenced to confirm the construct.

Tomato transformation

pCB1 was transferred from *E. coli* strain TG-2 to *Agrobacterium tumefaciens* LBA4404 by triparental mating with *E. coli* HB101 containing pRK2013 [2]. Stem internode explants excised from tomato seedlings (cv. Ailsa Craig), 4–6 weeks old, grown under sterile conditions, were incubated for 30 min in a suspension of *A. tumefaciens* in MS salts, 3.0% sucrose. These were transferred to feeder plates [16] of *Nicotiana tabacum* cv. Samsun on 0.8% shoot-inducing agar containing MS salts, 3.0% sucrose, 1.0 mg/l thiamine-HCl, 0.5 mg/l nicotinic acid, 0.5 mg/l pyridoxine HCl, 3.0 μ M IAA aspartic acid, 5.0 μ M zeatin riboside. After 2 days explants were transferred to the same media supplemented with 80 mg/l kanamycin, 500 mg/l carbenicillin. At intervals of 3 weeks, explants were subcultured onto shoot-inducing agar (0.8%) containing 80 mg/l kanamycin, 200 mg/l carbenicillin. Shoots were rooted on 0.8% agar containing MS salts, 3.0% sucrose, 1.0 mg/l thiamine-HCl, 0.5 mg/l nicotinic acid, 0.5 mg/l pyridoxine HCl, 30 mg/l kanamycin. After several cycles of rooting on agar containing kanamycin, to confirm transformation, plants were transferred to compost and grown to maturity in

growth rooms. The time interval from infection of tomato explants with *A. tumefaciens* to production of ripe fruit was approximately 6 months.

CAT assays

Root, stem, leaf, green and red fruit tissues were homogenised in a mortar and pestle (1 g tissue with 1 ml 250 mM Tris-HCl, pH 7.5 and 500 μ l glycerol). The homogenate was centrifuged at 13 000 *g* for 10 min and the supernatant passed through a Sephadex G-25 (coarse) column in 250 mM Tris-HCl, pH 7.5. Fractions eluting at the void volume were assayed for protein [4], and protein-containing fractions were pooled. 15 μ g protein aliquots were assayed for CAT activity by the addition of 10 μ l of 10 mM acetyl-CoA and 0.1 μ Ci ^{14} C-chloramphenicol (60 mCi/mMol; NEN) with 250 mM Tris-HCl, pH 7.5 to a final volume of 200 μ l. These samples were incubated at 37 °C for 30 minutes prior to extraction with two 300 μ l volumes of ethyl acetate. The combined ethyl acetate phases were dried, the residues taken up in 15 μ l ethyl acetate and subjected to TLC on silica gel plates with chloroform:methanol (95:5) as eluant prior to autoradiography. One unit of CAT activity (Pharmacia) is defined as catalysing the acetylation of one nanomole of chloramphenicol per minute at 37 °C and pH 7.8.

Results

Isolation of genomic clones encoding PG

Southern hybridisation of tomato genomic DNA to the insert from the PG cDNA (pTOM6) shows a simple restriction pattern when stringent hybridisation conditions are used (Figure 1). No additional bands are seen when less stringent ($2\times$ SSC, 1% SDS, 65 °C) hybridisation conditions are used (data not shown). A library of tomato genomic DNA was constructed by cloning size fractionated partial Sau3A digested fragments into λ EMBL3. Approximately 500 000 plaques were screened with the insert from pTOM6. Three hybridising clones were isolated and

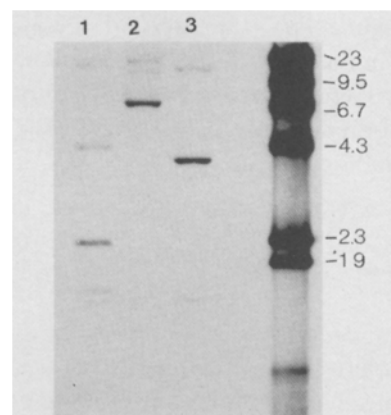


Fig. 1. Southern blot hybridisation of tomato genomic DNA probed with the PG cDNA clone, pTOM6. 5 μ g aliquots of DNA were digested with: Lane 1 *Hind* III and *Bam* HI; Lane 2 *Bam* HI; Lane 3 *Hind* III. The membrane was washed in $0.1\times$ SSC, 1% SDS at 65 °C.

purified to single plaques (GTOM6.2, GTOM6.4 and GTOM6.16). These clones hybridise to a 15 base oligonucleotide from the 3' end of the cDNA position 1205–1219 [20] but not to a 25 base oligonucleotide from the 5' end of the cDNA position 59–83 [20]. In order to isolate clones for the 5' end of the PG gene, the library was rescreened with the 200 bp 5' *Pst* I-*Hind* III fragment from pTOM6. Two additional clones were isolated (GTOM6.23 and GTOM6.30). Restriction analysis combined with Southern hybridisations using the 4 kb *Hind* III fragment from GTOM6.30 as a probe showed that these clones overlap with the initial three clones. Figure 2A shows the structure of these clones and the overlap between the restriction maps of the five clones. The combined restriction map of the PG gene represented by these five clones covers over 30 kb of tomato genomic DNA. A region of less than 10 kb hybridises to the PG cDNA (Figure 2B). GTOM6.23 contains a complete PG gene. All of the fragments that hybridise on the genomic Southern are represented in these clones.

Structure of the PG gene

The nucleotide sequence of 7.45 kb of the genomic clones has been determined (Fig. 3). The structure

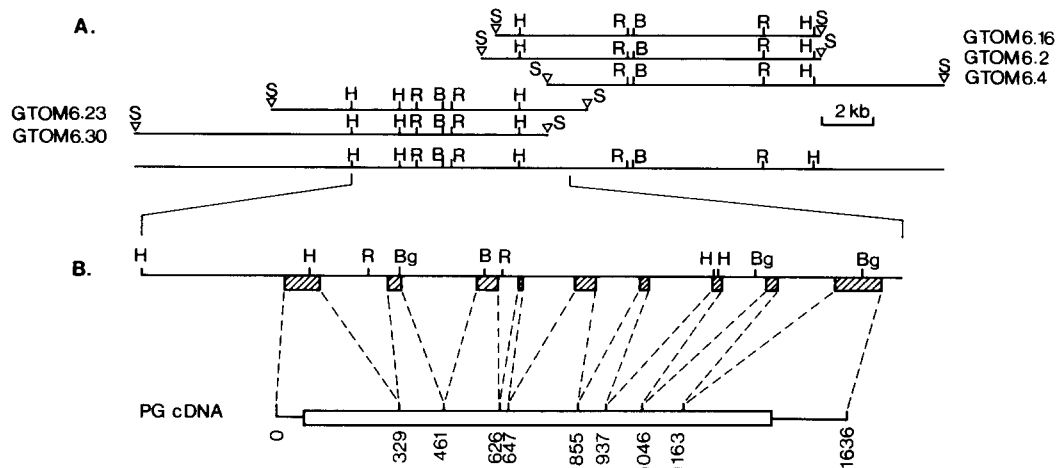


Fig. 2. Structure of the PG gene. A. Restriction map of the PG gene and its surrounding regions derived from five overlapping clones. B. Structure of the PG gene as determined by comparison to the sequence of pTOM6, a PG cDNA. Shaded boxes indicate exons. Open box indicates the open reading frame of the PG cDNA, pTOM6; S = *Sal* I (cloning site of λ EMBL3). H = *Hind* III, R = *Eco* RI, B = *Bam* HI, Bg = *Bgl* II.

```

AAGCTTCTTAAAAAGCAATTGATTAATTTGAAGTCAAAATAATTATTAACAATGTAAGACACTTAAGAACCATAGTTTGAAGGTTACCAATGCGCTATATTAATCAACT 120
TGATAATATAAAAAATTTCAATTCGAAAAAGGCTAAATATTTCTCAAGTATTCGAAATGTAACAACATACCATCCGCCACTATTGACTCCAAAAATAAATTTATCCACCTT 240
TGAGTTTAAATTTGACTACTTATATAACAATTCTAAATTTAACTATTTTAACTCTTTTAAATACATGAGCTTCAAAATTTAATATAATTTAATTTATBAATATCATTATAACCA 360
ACCACTACCACTCATTAAATCAATTAATCCACCAATTTCTACTATCAAAATTTGCTTAACACTACTAATAAGACGAAATTTGTCBAGTCCGAATCGAAGCACCATCTAATTTA 480
GGTTGAGCCCATATTTAGGAGGACCTTTCAATAGTATTTTCAAGCATGAATTTGAATTTAAGATTAAATGTAAGAGAGTATGACCCGAAATTAATTCATGCTTTTAAATA 600
TAATTATATAATATTTATGATTTGTTTAAATATAAACTTGAATATATTTTAAAAAATATCTATTAGTACCATCACATAATTGAGACGAGGAATAATTAAGATGAACATA 720
GTGTTAATAGTAAATGATGAGTGGTAGTAATTTATTTATAAATATATCAATAGCTTAATTTATACAAATATTTGAGCCCATGATTTTAAAAAATATTAATAGTTTGAATTTAA 840

ACCCTTAGATAATGTCATTTTGAACCCAAAAATGATGAGAGGATATTTAGAGCCAAATAGGAGGATGAGAGGATATTTGAAGCCAAATATGATGATGATGAGGATATTTGT 960
ATCAATTTCTAACTTTAAABATATTTAGGTCAATTTCCCTCTTTAGTTTATAGACTATAGTTAGTTGATCGAATATCATCTATTATTTCCGCTTAAATATTTTATTTTATA 1080
AATTTTAAAAAATAAATATTTTCCATTTAACTTTGATTTGAATTAATTTTAAAAATACCAATATATAAATTAATTTTAAACAAAGATTTGAACATATAATTTTAA 1200
TTATTCAAAATAAATATTTTAAACATCATATAAAGAAATACBACAAAAAATGAGACGGGAGAGCAAGCCAGCAAAAAATGTCGAAGAACTCTTTCTCTAAATATCTCTCATC 1320

CAAACTAATATAATACCATTAACCATATTGACCAACTCAACCCCTTAAATCTATAATAGCAAAACCTTCCCATACCTCTTATCATAAAAAATAAATATCTTTTCAA 1440
H V I Q R N S I L L I I F A S S I S T C R S N V I
TAGCAAGTTTAAAAACCATACATATATCATGTTTCCAAAGGAATAGTTTCTCTCTCATATTATTTTGTCTCATCAATTTCAACTTTGTAAGAGCAATGTTATG 1560
D D N L F K Q V Y D N I L E Q E F A H D F Q A Y L S Y L S K N I E S N N I D K
ATGACATTTTATCAACAGGTTTATGATAATTTCTTGAACAGGATTTGCTCATGATTTTCAAGCTATCTTTCTATTGAGCAAAAAATTTGAAGCAACAAATATTTGACAAAG 1680
V D K N G I K V I N V L S F G A K G D G K T Y D N I
TTGATAAAATGGGATTAAGTGAATGATGACTAGCTTTGGAGCTAAGGATGATGAAAAACATATGATAATTTGTAAGTATTTAAATATTGGAATATATTTGTTGGGATGAAATG 1800
ATAGAGATATAAGAAATATTTTGAAGGATGAAAGTTATATTTTAAAGTAGAAATATTTTCTCGTTTTAGTATTAGGTGAAATGAGTTTCTCGTTAAGCGAGGAAAGCTAT 1920
TTTCCATGTAAGTCTGATTTTCTTTTAAACGTCATAGTATTTGCTATACCTAAGAAATAGACACTTATTTATGATGATTTAGTCTCGAAGAGAAATGATAGTAATTTG 2040
CTTAATATAACTATCAATTTCTTATATGATATTTTCAACCAAAATACCAAGGCTAATCCAATAGTGGGCTCTAGAAATAGAGTAAGTTCTATTCAATTTCTAACCTTATTTAAT 2160
TTTAGTGGAAACCTCBACAAAAACGAACAACTATTTCAACCTTTTATATTCGAAATTCGAGACCAACCATATGAACAACTCAGACATGATATGCTTAAATATATATTTTCTA 2280
AAAAATATCTTCAATCTACCATATTGAATATTGAAAAATGACTTTTATCTATCGAACACATATCAAGAGTTTCTTTAAGAAATTTACCACATCATTTGGTATGTTTCTTATCGTGT 2400
A F E Q A W N E A C S S R T P V Q F V U P K N K N Y L L K Q I T F S G
AAATATCTTTCAAGCATTGAGCAAGCATGGAATGAAGCATGTTCTATGAAACCTGTTCAATTTGTTGTTTCTTAAAAACAAATATCTCTCAAGCAATCACCTTTTCAGGT 2520
P C R S S I S V K
CCATGCAATCTTCTATTTCAGTAAGTTAGCAJATGATTTATATCTCTTTGTTAGCAATATATCTGTTTATGACAAAAATTAAGAAATATCAAGATAGATAAACAA 2640
TGATTTTCTGCTCAATTTAGCGATTAAGTGAAGATATCAAAATGTTATGTTAGCTATGAGCAACTAGCTATGATTAAGTATGAGTGAAGAGTTTGAAGTCAATTTCTATTTTCT 2760
TGTAAGTAAAGATTTTGAACACATGATTAATTAATTAATGTTCTTAATTAATGTTCAATGATATGTTCAACTAAGAACTGTCAAGAAAAATAGAGAAATATTTATTT 2880
TAAATTAATTTAAAAAGAAATATGAGAAATTAATCAAGCGAGAGGATATACATAATCTTGGGATAAAGGATATATATATGTAAGAAAAACGACTACACATATCTAATAAA 3000
GTCTCATAAATGATATAAAAAATAGTGTGTAAGCAACAGTTATCCCTACAAAACTTTTGGGATGATGATCCAGAGGTTGTTTCCAGACTCTTGTCTAAAAAATATTTTCT 3120
AAATAGTTTGAAGAAATGTTATATGATGAAATATGAGAAAAACATATCAATTTAAAAATATAAGTAAATCAAGTAACCAATATAGGATAATCTCAATATGAAA 3240
I F G S L E A S S K I
TTTAGTGGCTTTCTTAAACATACTTATGTTTATCATGTTTCTTAAATTTCCCTCTTATTTTGTAAATTAATGCAAGATTTTGGATCTTGAAGCATCTAGTAAATTT 3360
S D Y K D R R L W I A F D S V Q N L V V G G G G T I N G N G Q V W H P S S C K I
CAGACTACAAATAGAGGCTTTGATGTTGTTGATAGTGTCAAAATTTAGTGTGTTGAGAGGAGGAACTATCAATGCAATGGACAGTATGTTGCAAGTTCTTGCAAAAATA 3480
N K S L
ATAAATCACTGTAATTTTAAACCTGCTTATAGTTTACGCTATGTTGCTCAATCTTTAACTTTGTTCAAGATATATATATTTGAAGAGGTTGACAAATGCATCACATTT 3600

```

of the PG gene (Fig. 2B) was deduced by comparing it to the PG cDNA sequence [20]. The sequence of the protein coding region is identical to that of the cDNA. It is apparently divided into 9 exons ranging in size from 21 to 460 bp separated by 8 A/T-rich intervening sequences ranging in size from 99 to 953 bp. The boundary sequences between the introns and the exons are similar to the consensus sequences for plant genes [6] and animal genes [30].

The transcription start site of the PG gene was determined by S1 mapping and primer extension experi-

ments (Fig. 4). A ^{32}P -labelled single-stranded 'prime cut' probe was prepared from the 5' 1.6 kb *Hind* III fragment cloned into M13 mp19. After synthesis of a labelled second strand, the DNA was digested with *Eco* 0109, which cuts at position 156, and a 1541 base single-stranded fragment which includes the universal sequencing primer was reco-

vered from a preparative denaturing acrylamide gel. This probe was hybridised to poly(A)⁺ RNA from ripe tomato and digested with S1 nuclease. A fragment of 218/219 bases was the most intense band on a denaturing polyacrylamide gel (Fig. 4A). This corresponds to a transcription start site at position 1411/1412 (Fig. 3). Three fainter bands represent

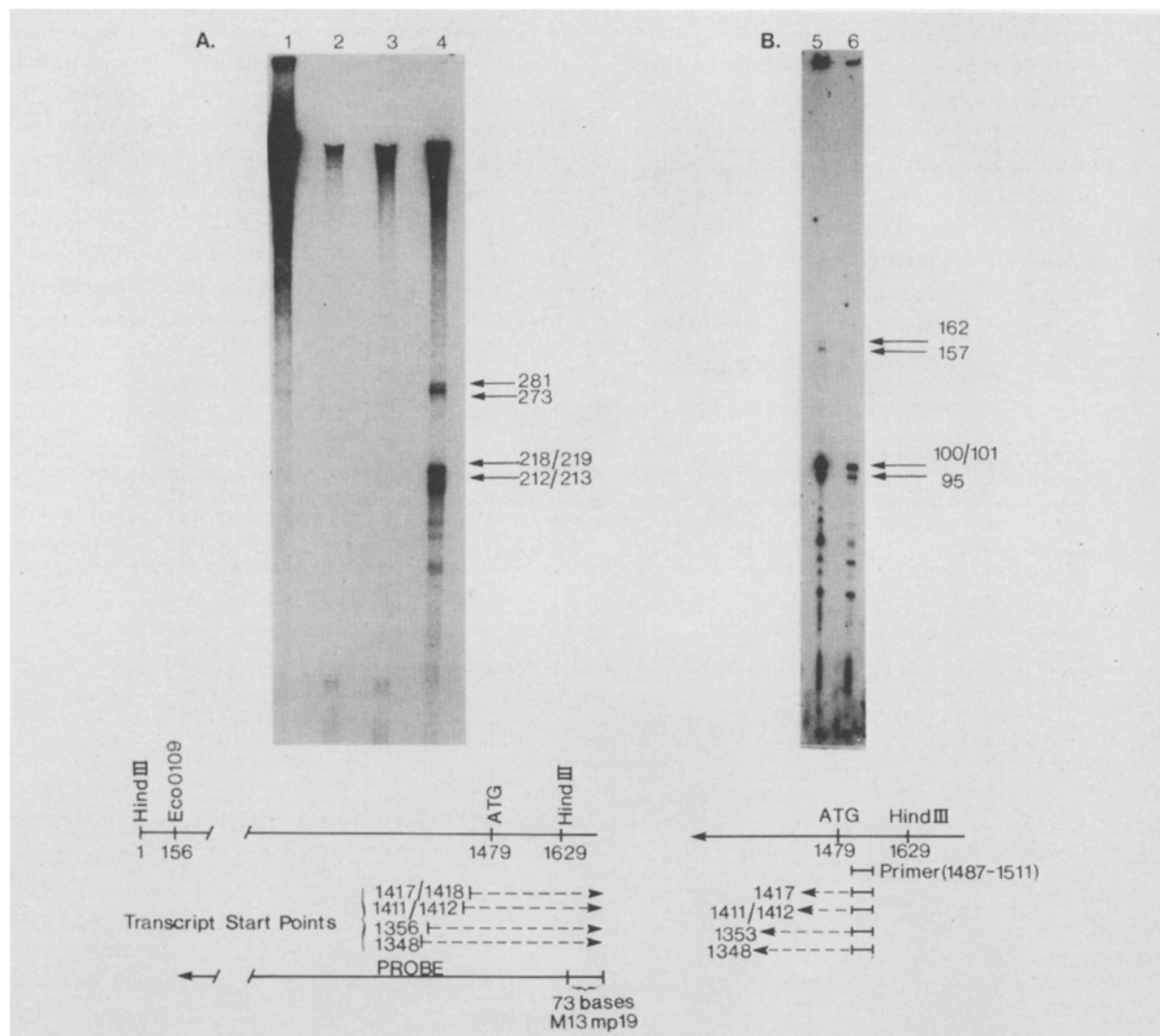


Fig. 4. Determination of the transcription start site of the PG gene. **A.** S1 mapping. The preparation of the 1541 bp probe is described in the text. Lane 1, probe without treatment with S1 nuclease. Lane 2, digestion of probe with 200 U S1 nuclease without hybridisation to RNA. Lane 3, digestion of probe with 100 U S1 nuclease without hybridisation to RNA. Lane 4, poly(A)⁺ RNA (1 µg) from orange tomato fruit hybridized to the probe and subsequently digested with 200 U S1 nuclease. The length of the labelled fragments is indicated adjacent to the autoradiograph of the gel. The position of the transcription start sites on the gene sequence are indicated in the diagram. **B.** Primer extension. Lanes 5 and 6, 25 bp oligonucleotide primer (see text for sequence) was annealed to 1 µg orange tomato fruit poly(A)⁺ RNA and extended with reverse transcriptase. The length of the labelled fragments is indicated adjacent to the autoradiograph of the gel. The determined position of the transcription start sites on the gene sequence are indicated in the diagram.

fragments of approximately 212/213, 273 and 281 bases. These correspond to potential transcription start points close to positions 1417/1418, 1356 and 1348. A synthetic oligonucleotide primer 5'-GAGAAGGAGAATACTATTCCTTTGG-3' complementary to the PG mRNA at a position corresponding to 1487 and 1511 on the gene sequence was hybridised to poly(A)⁺ RNA from ripe tomato and extended with reverse transcriptase. A fragment of 100 or 101 bases was the longest heavily labelled band detected on a denaturing polyacrylamide gel (Fig. 4B). This corresponds to a transcription start site at position 1411/1412. This is the same position as that predicted from the most intense band obtained by the S1 mapping. Three fainter bands represent fragments of approximately 95, 157 and 162 bases. These correspond to transcription start points close to position 1417, 1353 and 1349 which are similar positions to those predicted by S1 mapping.

Ripening specific expression in transgenic tomatoes

A transcriptional fusion, pCB1, of 1.45 kb of DNA

from the 5' end of the PG gene plus the CAT reporter gene was constructed in the binary plant transformation vector Bin19 (Bevan, 1984) (Fig. 5). Tomato stem segments were transformed with *A. tumefaciens* containing pCB1. Transformed shoots were selected by their ability to grow and root on kanamycin-containing media. After several rounds of selection in tissue culture, plants were transferred to compost and grown to maturity in the growth room. Eleven independently transformed kanamycin-resistant plants were regenerated. Genomic Southern analysis of *Eco* RI/*Hind* III-digested DNA indicated that the intact chimaeric CAT gene was present in all of these plants (data not shown). CAT activity was detected in ripening fruit from three of these plants, two of which (GR19 and GR87) were selected for more detailed analysis. Tomato fruit, leaf, stem and root material from these transformed and control untransformed plants were assayed for CAT enzyme activity. The results from GR19 are presented in Fig. 6. Similar results were obtained for GR87 (data not shown). Reconstruction experiments using CAT enzyme and extracts from untransformed plant material were also carried out. Exogenous CAT enzyme was partially inhibited by

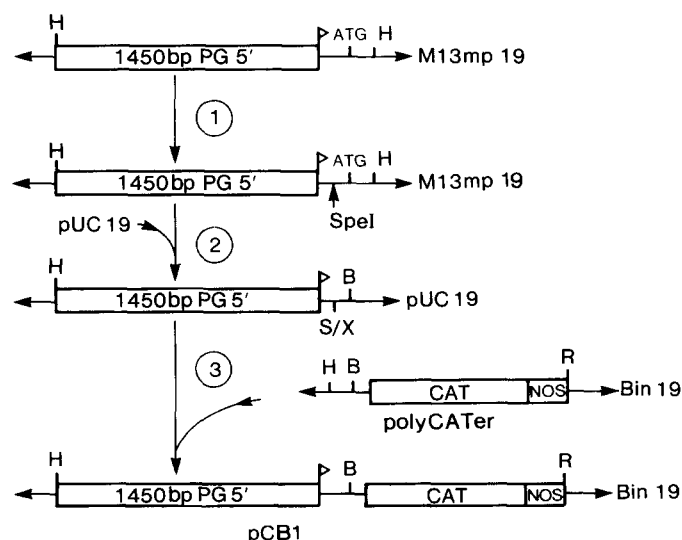


Fig. 5. Construction of the transcriptional fusion vector pCB1. Step 1: site-directed mutagenesis to introduce a *Spe* I site into a clone of the 1.65 kb *Hind* III fragment in M13 mp19 (see Materials and methods). Step 2: cloning of the 1.45 kb *Hind* III/*Spe* I fragment into pUC19 cut with *Hind* III and *Xba* I. Step 3: Cloning of the 1.45 kb *Hind* III/*Bam* HI fragment into polyCATer [9] cut with *Hind* III and *Bam* HI, containing the CAT gene and the nopaline synthase (NOS) polyadenylation sequence in Bin19. ATG: translation start; P: transcription start point; H: *Hind* III; S/X: *Spe* I ends cloned into *Xba* I site; b: *Bam* HI; R: *Eco* RI.

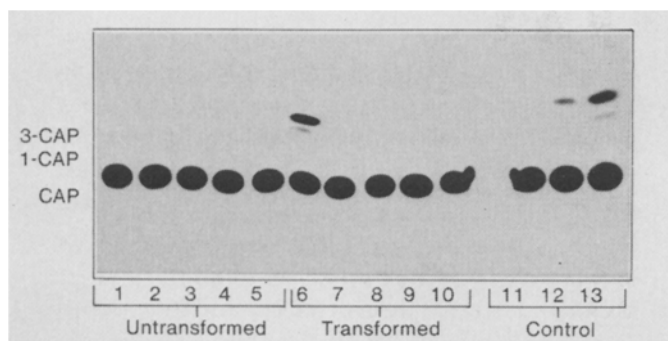


Fig. 6. Chloramphenicol acetyl transferase activity in different organs of the transgenic tomato plant (GR19). CAT assays were carried out on 15 μ g aliquots of protein extracted as described in Materials and methods, the reaction products separated by thin-layer chromatography and detected by autoradiography. Protein extracts were from roots (Lanes 1, 10), stems (Lanes 2, 9), leaves (Lanes 3, 8), green fruit (Lanes 4, 7) and red fruit (Lanes 5, 6). The activity of pure CAT enzyme (Pharmacia) is also shown. (Lane 11, 0 units; Lane 12, 0.01 units; Lane 13, 0.02 units) CAP: ^{14} chloramphenicol; 1-CAP: 1-acetyl chloramphenicol; 3-CAP: 3-acetyl chloramphenicol. Similar results were obtained from plant GR87.

tomato extracts. The degree of inhibition varied depending on the tissue, but was greatly reduced by passing extracts through a Sephadex G-25 column before assay. In plants GR19 and GR87, high levels of CAT activity (greater than 66 units/mg protein) were detectable in the red tomato fruit extracts. Red fruit from untransformed plants showed negligible levels of CAT activity. Similarly, no significant CAT activity was detectable in green fruit, root, leaf or stem material from either GR19, GR87 or untransformed plants (Fig. 6).

Discussion

A tomato polygalacturonase gene was identified on a group of overlapping clones of genomic DNA. The structure of the gene has been determined by DNA sequence analysis and comparison with the sequence of a near-full-length cDNA. The regions of the gene that correspond to the cDNA have identical sequence. All of the fragments hybridising to the PG cDNA on a genomic Southern are attributable to the tomato DNA that is represented in the overlapping genomic clones. This data is consistent with the sug-

gestion that there may be only one PG gene in tomato per haploid genome. This conclusion is supported by the fact that only one gene has been isolated and that the PG gene has been mapped to a single locus on chromosome 10 (M. Mutschler *et al.*, unpublished). Bernatzky and Tanksley [1] have found that the majority of a selection of unique random cDNA clones correspond to single loci in the tomato genome.

Another tomato ripening-related cDNA, pTOM13, has at least three genes per haploid genome [23]. These genes are differentially expressed in ripening fruit and wounded tissue. A third ripening-related cDNA, pTOM5, has at least two genes per haploid genome, one of which is a pseudogene (P. Moureau *et al.*, unpublished).

There are several implications if there is only one PG gene. Firstly, the three isoforms of PG (PG-1, PG-2a, PG-2b) have been shown to be structurally and immunologically related [39, 43]. All isoforms are glycosylated [43] and the mixture of PG2a and 2b can be converted into a form resembling PG1 *in vitro* [40]. It has not been clear whether the isoenzymes are products of separate genes or are formed by post-translational modification of a single polypeptide [18]. The evidence presented here suggests that there is only one PG gene, which implies that the isoenzymes are the result of post-translational modification of the product of this gene. Thus, the sequential appearance of PG-1 and PG-2 during ripening [39] is unlikely to be due to the differential expression of two PG genes. Secondly, PG activity is found in flower abscission zones as well as ripening tomatoes [41]. There may be two PG genes that are so divergent in sequence that the one expressed in abscission zones does not cross-hybridise with the one expressed in fruit. Alternatively, the enzyme activities found in ripening fruit and abscission zones could also be the products of a single gene, which is regulated by tissue-specific control systems in each developmental situation. This possibility is particularly intriguing in view of the evidence from studies on the *rin* ripening mutant of tomato. This mutant contains the PG gene described in this paper but fruit fail to ripen and the PG mRNA only accumulates to about 0.5% of normal levels (J. Knapp *et al.*, unpublished) [12]. In contrast, normal levels of PG

activity are detectable in the flower abscission zones of the *rin* mutant [33].

The coding region of the PG gene is interrupted by eight introns all of which have boundary sequences that conform to consensus sequences [6]. The exons range in size from 21 to 460 bp. S1 mapping and primer extension experiments indicated several transcriptional start sites. The site at position 1411 or 1412 corresponds to the most intense bands from both the S1 mapping and primer extension experiments. This is probably the major transcription start site in ripening fruit. It is possible that the multiple transcription start sites are utilised differentially in different developmental situations. At the end of the gene there is a potential polyadenylation signal sequence AATAATAAA [32, 10] 172 nucleotides downstream of the UAA stop codon and 14 bp 5' of the polyadenylation site. This sequence is identical to that found in another tomato ripening gene, pTOM13 [24].

1.4 kb of DNA 5' of the transcription start was sequenced as a prelude to further work on the regulation of expression of the PG gene. The sequence is very A/T-rich (75%) and contains several sequences that may have a role in controlling expression. At position 1381 the sequence TATAA resembles a TATA box, 31 bases 5' of the major transcription start. Two sequences CAATT (1341) and CCAT (1350) are putative CAAT boxes [5] 69 and 62 bases respectively 5' of the major transcription start. The sequence TAATATAATA (1326) may include TATA boxes for the minor transcription start points. Between positions 877 and 935 (535 bases 5' of the major transcription start) there is a 28 bp almost perfect direct repeat which is separated by 3 nucleotides. Sequences either side of this direct repeat have some limited similarity with the repeat sequence. There is little obvious homology between these direct repeats and those preceeding other developmentally regulated plant genes (e.g. Rubisco small subunit [29]; sucrose synthase [42]; zein [26]; and chalcone synthase [37]). The possible role of the 5' sequence as a promoter for regulating polygalacturonase gene expression in ripening tomatoes was investigated in transgenic tomato plants using chimaeric gene constructs containing a CAT reporter gene.

There is now considerable evidence that ripening of tomatoes requires the expression of new genes

that encode enzymes catalysing different aspects of ripening. Furthermore, the existence of a group of developmentally regulated genes may well be a general feature of the ripening process in all fruits [18]. The gene for PG is the first ripening-related genomic sequence to be characterised. Expression of the PG gene is stimulated by C₂H₄ [20, 28] but the mechanism is unknown. We have begun a programme to analyse in detail *cis*- and *trans*-acting factors governing the regulation of its expression using a homologous transformation system. This may allow a better quantitative assessment of the relative importance of specific control regions than a heterologous transformation system. The results obtained with transgenic tomato plants containing the pCBI construct show that the 1.4 kb 5' fragment of the PG gene (Fig. 5) is sufficient to cause the bacterial CAT gene to be expressed in an organ-specific and developmentally regulated manner. Deletion analysis will be required in order to locate precisely the sequences conferring ripening-specific gene expression. In addition, further work will be necessary in order to establish whether there are other *cis*-acting control signals, located elsewhere, that are also involved in qualitative or quantitative aspects of gene expression during ripening, or other developmental situations. The ability to obtain specific expression of foreign genes in ripening tomatoes opens up the possibility of genetic manipulation of this important crop.

Acknowledgements

We thank C. Wighton, N. Halstead, and S. Sheerman for technical assistance. This work was supported by a SERC-Cooperative Award (UK) to D. Grierison. P. Moureau was a recipient of an EEC Training Fellowship in Biomolecular Engineering. This work was done under MAFF licences PHF 77/132 (67) and PHF 1039/9 (77).

References

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

2. Bevan MW: Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Res* 12: 8711–8721 (1984).
3. Biggin M, Farrell PJ, Barrell BG: Transcription and DNA sequence of the *Bam* HI L fragment of B95-8 Epstein-Barr Virus. *EMBO J* 3: 1083–1090 (1984).
4. Bradford MM: A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding. *Anal Biochem* 72: 248–254 (1976).
5. Breathnach R, Chambon P: Organisation and expression of eukaryotic split genes coding for proteins. *Ann Rev Biochem* 50: 349–383 (1987).
6. Brown JWS: A catalogue of splice junction and putative branch point sequences from plant introns. *Nucleic Acids Res* 14: 9549–9559 (1986).
7. Carter P, Bedouelle H, Wayne MMY, Winter G: Oligonucleotide site-directed mutagenesis in M13. Anglian Biotechnology Ltd., Colchester (1985).
8. Chomczynski P, Qasba PK: Alkaline transfer of DNA to plastic membranes. *Biochem Biophys Res Comm* 122: 340–344 (1984).
9. Colot V, Robert LS, Kavanagh TA, Bevan MW, Thompson RD: Localisation of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. *EMBO J* 6: 3559–3564 (1987).
10. Dean C, Tamaki S, Dunsmuir P, Faureau M, Katayama C, Dooner H, Bedbrook J: mRNA transcripts of several plant genes are polyadenylated at multiple sites *in vivo*. *Nucleic Acids Res* 14: 2229–2240 (1986).
11. Della Penna D, Alexander DC, Bennett AB: Molecular cloning of tomato fruit polygalacturonase: Analysis of polygalacturonase mRNA levels during ripening. *Proc Natl Acad Sci USA* 83: 6420–6424 (1986).
12. Della Penna D, Kates DS, Bennett AB: Polygalacturonase gene expression in Rutgers, *rin*, *nor* and *Nr* tomato fruit. *Plant Physiol* 85: 502–607 (1987).
13. Dellaporta SL, Wood J, Hicks JB: A plant DNA miniprep: Version 11. *Plant Mol Biol Rep* 1: 19–21 (1983).
14. Farrell PJ, Deininger P, Bankier A, Barrell BG: Homologous upstream sequences near Epstein-Barr virus promoters. *Proc Natl Acad Sci USA* 80: 1565–1569 (1983).
15. Feinberg A, Vogelstein B: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 137: 266–267 (1984).
16. Fraley RT, Rogers SG, Horsch RB, Sanders PR, Flick JS, Adams SP, Bittner ML, Brand LA, Kink CL, Fry JS, Galluppi GR, Goldberg SB, Hoffmann NL, Woo SC: Expression of bacterial genes in plant cells. *Proc Natl Acad Sci USA* 80: 4803–4807 (1983).
17. Frischauf AM, Lehrach H, Poustka A, Murray N: Lambda replacement vectors carrying polylinker sequences. *J Mol Biol* 170: 827–842 (1983).
18. Grierson D: Gene expression in ripening tomato fruit. *CRC Crit Rev Plant Sci* 3: 113–132 (1985).
19. Grierson D, Slater A, Speirs J, Tucker GA: The appearance of polygalacturonase mRNA in tomatoes: one of a series of changes in gene expression during development and ripening. *Planta* 163: 263–271 (1985).
20. Grierson D, Tucker GA, Keen J, Ray J, Bird CR, Schuch W: Sequencing and identification of a cDNA clone for tomato polygalacturonase. *Nucleic Acids Res* 14: 8595–8603 (1986).
21. Heidecker G, Messing J: Structural analysis of plant genes. *Ann Rev Plant Physiol* 37: 439–466 (1986).
22. Hobson GE: The firmness of tomato fruit in relation to polygalacturonase activity. *J Hort Sci* 40: 66–72 (1965).
23. Holdsworth MJ: Characterisation of an ethylene related small multigene family from *Lycopersion esculentum*. Ph.D. Thesis, University of Nottingham (1987).
24. Holdsworth MJ, Bird CR, Ray J, Schuch W, Grierson D: Structure and expression of an ethylene related mRNA from tomato. *Nucleic Acids Res* 15: 731–739 (1987).
25. Hudson GS, Farrell PJ, Barrell BG: Two related but differentially expressed potential membrane proteins encoded by the *Eco* RI Dhet region of Epstein-Barr virus B95-8. *J Virol* 53: 528–535 (1985).
26. Langridge P, Feix G: A zein gene of maize is transcribed from two widely separated promoter regions. *Cell* 34: 1015–1022 (1983).
27. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
28. Maunders MJ, Holdsworth MJ, Slater A, Knapp JE, Bird CR, Schuch W, Grierson D: Ethylene stimulates the accumulation of ripening-related mRNAs in tomatoes. *Plant Cell Envir* 10: 177–184 (1987).
29. Morelli G, Nagy F, Fraley RT, Rogers SG, Chua N-H: A short conserved sequence is involved in the light-inducibility of a gene encoding ribulose 1,5-bisphosphate carboxylase small subunit of pea. *Nature* 315: 200–204 (1985).
30. Mount SM: A catalogue of splice junction sequences. *Nucleic Acids Res* 10: 459–472 (1982).
31. Pressey R, Avants JK: Two forms of polygalacturonase in tomatoes. *Biochim Biophys Acta* 309: 363–369 (1973).
32. Proudfoot NJ, Brownlee GG: 3' Non-coding region sequences in eukaryotic messenger RNA. *Nature* 263: 211–214 (1976).
33. Roberts JA, Schindler CB, Tucker GA: Ethylene-promoted tomato flower abscission and the possible involvement of an inhibitor. *Planta* 160: 159–163 (1984).
34. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467 (1977).
35. Sheehy RE, Pearson J, Brady CJ, Hiatt WR: Molecular characterisation of tomato fruit polygalacturonase. *Mol Gen Genet* 208: 30–36 (1987).
36. Slater A, Maunders MJ, Edwards K, Schuch W, Grierson D: Isolation and characterisation of cDNA clones for tomato polygalacturonase and other ripening-related proteins. *Plant Mol Biol* 5: 137–147 (1985).
37. Sommer H, Saedler H: Structure of the chalcone synthase gene of *Antirrhinum majus*. *Mol Gen Genet* 202: 429–434 (1986).
38. Tucker GA, Grierson D: Synthesis of polygalacturonase during tomato fruit ripening. *Planta* 155: 64–67 (1982).
39. Tucker GA, Robertson NG, Grierson D: Changes in poly-

- galacturonase isoenzymes during the ripening of normal and mutant tomato fruit. *Eur J Biochem* 112: 119–124 (1980).
40. Tucker GA, Robertson NG, Grierson D: The conversion of tomato-fruit polygalacturonase isoenzyme 2 into isoenzyme 1 *in vitro*. *Eur J Biochem* 115: 87–90 (1981).
 41. Tucker GA, Schindler CB, Roberts JA: Flower abscission in mutant tomato plants. *Planta* 160: 164–167 (1984).
 42. Werr W, Frommer W-B, Maas C, Starlinger P: Structure of the sucrose synthase gene on chromosome 9 of *Zea mays* L. *EMBO J* 4: 1373–1380 (1985).
 43. Zainon MA, Brady CJ: Purification and characterisation of the polygalacturonase of tomato fruit. *Aust J Plant Physiol* 9: 155–169 (1982).
 44. Zoller MJ, Smith M: Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. *DNA* 3: 479–488 (1984).