Plant Molecular Biology **24**: 429–434, 1994.

Cloning and characterization of the gene for phytoene desaturase (*Pds*) from tomato (*Lycopersicon esculentum*)

Varda Mann, Iris Pecker and Joseph Hirschberg*
Department of Genetics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel (author for correspondence)*

Received 26 July 1993; accepted in revised form 12 November 1993

Key words: carotenoid, intron, splicing, tomato

Abstract

The gene *Pds* encodes phytoene desaturase, a key enzyme in carotenoid biosynthesis that converts phytoene to ζ-carotene. We have cloned and analyzed the genomic DNA sequence of *Pds* from tomato. In tomato *Pds* is comprised of 15 exons that, together with the introns occupy over 8 kb. A putative promoter sequence has been identified by comparison with the cDNA sequence of *Pds*. A consensus nucleotide sequence around intron splicing sites in tomato genes was determined by compiling data on 137 introns in 34 genes. This consensus sequence generally agrees with the consensus sequence of other higher plants with only minor differences that are unique to tomato.

Introduction

Carotenoids comprise the largest class of pigments in nature. They are synthesized *de novo* only in photosynthetic organisms and in several bacteria and fungi, but are essential cellular components in animals.

Carotenoids are produced in the general isoprenoid biosynthetic pathway [6]. The first step that is specific for carotenoid biosynthesis is the head-to-head condensation of two molecules of geranylgeranyl pyrophosphate (GGPP) to yield *cis*-phytoene, a C_{40} colorless hydrocarbon. Phytoene undergoes a series of desaturation and two cyclization reactions to yield ß-carotene. Most xanthophylls are oxygenated derivatives of ß-carotene.

Phytoene desaturase (PDS) is a key enzyme in the pathway. In cyanobacteria, algae and plants it is a single-gene product enzyme that carries out two dehydrogenation steps of phytoene to yield ζ-carotene [13, 16]. We have recently shown that in cyanobacteria PDS is a rate-limiting step in carotenogenesis [10]. Genes for enzymes involved in carotenogenesis have been cloned from *Rhodobacter capsulatus*, two *Erwinia* species and *Neurospora crassa* [1–4, 15]. However, due to lack of sequence conservation they could not be used

---

The nucleotide sequence data reported will appear in the EMBL and GenBank Nucleotide Sequence Databases under the accession number X71023 (LEPDSG).

as molecular probes for cloning the homologous plant genes.

The gene *Pds*, encoding a plant-type phytoene desaturase, was first cloned from *Synechococcus* sp. PCC 7942 [8, 9]. The cyanobacterial gene was used as a molecular probe for isolating the homologous cDNA clones from the alga *Dunaliella bardawil* [17] and from tomato [16] and soybean [5]. The nucleotide sequences of these clones have been determined. The nucleotide sequences of the *Pds* gene from the cyanobacterium *Synechocystis* 6803 [14] and the cDNA from bell pepper [12] have been recently published. Comparison of these sequences indicates a conservation of over 60% in the amino acid sequences of the various deduced PDS polypeptides [17].

Carotenoids are produced by all foliar tissues, however there is a large increase in carotenoids synthesis in certain tissues, such as flower petals and fruits. We have shown that this rise is accompanied by an elevation in the steady state levels of *Pds* transcripts in ripening tomato fruits [16] and in tomato leaves which are exposed to high photon flux density [17]. The increase in the amount of *Pds* mRNA has been recently found to be 10-fold in flowers and 10-fold in fruits [11]. Previously we have shown that *Pds* is a single gene in the tomato genome [17]. This observation has suggested that a transcriptional regulation of the same *Pds* sequence, which is likely to involve a single promoter, is responsible for the control of carotenogenesis. To study the molecular mechanisms of regulation of *Pds* transcription, we have cloned the genomic *Pds* sequence. We present here the organization and nucleotide sequence of the *Pds* gene from tomato.

Cloning and sequencing of *Pds* from tomato

A genomic DNA library from *Lycopersicon esculentum* VF36 (kindly provided by Dr D. Zamir, Faculty of Agriculture, The Hebrew University of Jerusalem) was screened with *Pds* cDNA [16]. This library was constructed by cloning genomic DNA fragments, resulting from a partial digestion with the restriction endonuclease *Sau* 3A, into the *Bam* HI site of $\phi\lambda$ EMBL3. *Escherichia coli* strain KM 392 was transfected with the recombinant phage.

The genomic library was screened by plaque hybridization using as probe the ^{32}P-dCTP-labelled cDNA from *L. esculentum* encoding phytoene desaturase [16]. A positive λ phage clone was isolated. The restriction map of the genomic DNA insert surrounding the *Pds* sequence is shown in Fig. 1.

Five *Hind* III fragments, which were derived from the original genomic insert in the phage vector carrying the gene *Pds*, were subcloned into the plasmid pUC118 and deletion derivative were generated. The double-stranded DNA templates were sequenced by the chain termination method [20] using the Sequenase kit (USB). The nucleotide sequence is given in Fig. 2.

The coding sequences of *Pds* and the introns were identified by comparison with the cDNA sequence [16]. This comparison revealed that *Pds* in tomato is comprised of 15 exons interrupted by 14 introns. One of the introns is located in the 5′-untranslated region. The exon-intron organization of *Pds* is schematically illustrated in Fig. 1.

According to the cDNA sequence [16] the transcription starts at nucleotide 561 (Fig. 2). Putative promoter 'TATA box' consensus sequences are found upstream to this site. We presume that this promoter, together with the nucleotide sequence upstream to it, are involved in the regulation of *Pds* transcription during fruit ripening [16] and following oxidative stress [17].

The tomato gene that encodes the chromoplast-type of phytoene synthase, *Psy1*, has recently been reported [18]. Though its expression during fruit ripening is up regulated at the transcriptional level,

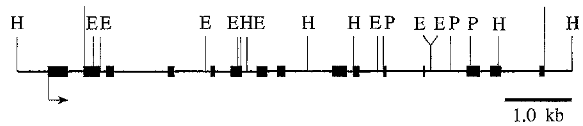

Fig. 1. Restriction map and structure of the tomato genomic *Pds* gene. H, *Hind* III; E, *Eco* RI; P, *Pst* I. Boxes represent exons; NH_2 and COOH, represent the amino and carboxy termini of the protein, respectively; arrow, beginning of cDNA.

Table 1. Consensus sequence for RNA splicing in tomato and other plants. Nucleotide frequencies are given as a percentage. A total of 137 introns were examined. *L.e.*, consensus for tomato; *A.t.*, consensus for *Arabidopsis thaliana* (data from the *Arabidopsis* Data Base); Plant, consensus for higher plants (data from Brown [7]).

A. Nucleotide frequencies at the 5′ exon-intron splice junction

	Exon							Intron						
	−7	−6	−5	−4	−3	−2	−1	+1	+2	+3	+4	+5	+6	+7
A	30.7	31.4	31.4	27.7	32.8	77.4	5.1	0.7	1.5	74.5	61.3	24.8	17.5	43.0
C	21.1	19.0	21.9	16.1	38.0	3.6	2.2	0	6.6	5.8	16.8	11.7	16.8	8.8
G	16.8	26.3	13.9	22.6	19.7	5.1	85.4	97.8	0.7	2.9	2.2	43.1	9.5	8.8
T	31.4	23.3	32.8	33.6	9.5	13.9	7.3	1.5	91.2	16.8	19.7	20.4	56.2	39.4
L.e.	t/a	N	t/a	t/a	c/a	A	G	G	T	A	A	G	T	A/T
A.t.				t/a/c	a/c	A	G	G	T	A	A	G	T	T/A
Plant					C/A	A	G	G	T	A	A	G	T	

B. Nucleotide frequencies at the 3′ exon-intron splice junction

	Intron															Exon						
	−15	−14	−13	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	+1	+2	+3	+4	+5	+6	+7
A	29.2	23.4	23.4	17.6	18.2	24.1	33.7	31.4	22.6	32.1	17.5	29.9	4.4	94.9	2.9	8.0	19.7	26.3	31.4	15.3	31.4	27.0
C	10.2	7.3	8.8	10.9	9.5	14.6	9.5	13.1	16.1	10.2	2.9	11.7	65.7	2.9	0	13.9	12.4	20.4	14.6	13.9	16.0	16.1
G	17.2	16.8	15.3	13.1	19.7	11.7	10.9	8.0	11.7	18.3	13.1	29.2	1.4	2.2	95.6	56.9	24.1	24.8	29.9	32.1	27.0	31.4
T	43.1	52.5	52.5	58.4	52.5	49.6	40.9	47.5	49.6	39.4	66.5	29.2	28.5	0	1.5	21.2	43.8	28.5	24.1	38.7	25.6	25.5
L.e.	T/a	T	T	T	T	T	T/A	T/A	T	T/A	T	N	C	A	G	G	T	N	N	T/G	N	N
A.t.									T	T	T	G/a/c	C	A	G	G	T	N	N			
Plant	T	T	T	T	T	T	T	T	T	T	T	G	C	A	G	G						

it shares no sequence homology with *Pds* in the promoter and upstream sequences.

Previously, we have demonstrated by Southern hybridization with cDNA that there is a single *Pds* sequence in the tomato genome [17]. The predicted restriction fragments of the genomic *Pds* are consistent with our previous results, thus confirming that *Pds* is a unique gene in tomato.

Consensus sequences for splice sites in tomato genes

Nucleotide sequences at the intron-exon junctions in the *Pds* gene are generally conserved with the consensus intron borders sequences in eukaryotic genes [7]. We have compiled data on the nucleotide sequences at the exon-intron borders of all known tomato genes that are available in the GenBank database and additional three unpublished genes, CD4A, CD4B and the gene for threonine deaminase. The proposed consensus sequences for splicing junctions in tomato genes, based on 137 introns in 34 genes, is given in Table 1. Most intron sequences start with GT at the 5′ and end with AG at the 3′. The tomato splice junction consensus sequence at the 5′ is t/a N t/a t/a c/a A G || G T A A G T A/T (Table 1A). It agrees generally with the splice junction sequences of plant and animal introns [7]. The tomato 3′ consensus sequence is T/a T T T T T T/A T/A T T/A T N C A G || G T N N G/T N N (Table 1B). It differs from the general consensus of plants at position −4, where any nucleotide can occur, as in the animal sequences [7]. Although plant introns are essentially similar to

```
    1 AAGCTTGACGATTTTCCTTCAGAGTAGAAATTGAAAAGAATCAACTAAAAAGGATAGTCCTTCGATTTGATTTCCGGCTTAAAAATAAACTAATAAGAATGAGAG   105
  106 AGCGAATAATAGAATATTTTGAAATTTTAAAGATATTCAACTATGTTAAATTGCGTTATAAATTTCTTAAATTAGTAGCACCTAATAGTTTAGTTCTCAAAAGTC   210
  211 AAAACTACTACATAATCGTGCTCATTTTTCACATTAAAATGCCTACATGATGTAAAAGTAAAACTCGTAGCATTCTACGTGTTTTACTCAACTCAAACATCCTGT   315
  316 TCATTTTAATAAACGTACGATGAGCTTCTCTCTCCAATTTTCTTTTCTTTTTTTTTTTTAAAAAAATATTTTTTTTTATATCAATCCAAATGGGCTCCAATTTAT   420
  421 CATAAATTAGGTAGAAACTTAGATATTAAAGAAAGAAAAGGGTTTATCTCGCAAGTGTGGCTATGGTGGGACGTGTCAAATTTTGGATTGTAGCCAAACATGAGA   525
  526 TTTGATTTAAAGGGAATTGGCCAAATCACCGAAAGCAGGCATCTTCATCATAAATTAGTTTGTTTATTTATACAGAATTATACGCTTTTACTAGTTATAGCATTC   630
  631 GGTATCTTTTTCTGGGTAACTGCCAAACCACCACAAATTTCAAGTTTCCATTTAACTCTTCAACTTCAACCCAACCAAATTTATTTGCTTAATTGTGCAGAACCA   735
  736 CTCCCTATATCTTCTAGGTGCTTTCATTCGTTCCGAGgtaagaaaagattttgtttctttgaatgctttatgccactcgtttaacttctgaggtttgtggatct   840
  841 tttaggcgactttttttttttttgtatgtaaaatttgtttcataaatgcttctcaacataaatcttgacaaagagaaggaatttaccaagtatttaggttcaga   945
  946 aatggataatttcttactgtgaaatatccttatggcaggttttactgttatttttcagTAAAATGCCTCAAATTGGACTTGTTTCTGCTGTTAACTTGAGAGTC  1050
    1                                                                M  P  Q  I  G  L  V  S  A  V  N  L  R  V      14
 1051 CAAGGTAGTTCAGCTTATCTTTGGAGCTCGAGGTCGTCTTCTTTGGGAACTGAAAGTCGAGATGGTTGCTTGCAAAGGAATTCGTTATGTTTTGCTGGTAGCGAA  1155
   15 Q  G  S  S  A  Y  L  W  S  S  R  S  S  S  L  G  T  E  S  R  D  G  C  L  Q  R  N  S  L  C  F  A  G  S  E      49
 1156 TCAATGGGTCATAAGTTAAAGATTCGTACTCCCCATGCCACGACCAGAAGATTGGTTAAGGACTTGGGGCCTTTAAAGgtttggtttgaattctatagactgatg  1260
   50 S  M  G  H  K  L  K  I  R  T  P  H  A  T  T  R  R  L  V  K  D  L  G  P  L  K                                 75
 1261 acttgaattattatctcaaggcatatatttctctaaataaaggaactatcttgtcattcagGTCGTATGCATTGATTATCCAAGACCAGAGCTGGACAATACAG  1365
   76                                                               V  V  C  I  D  Y  P  R  P  E  L  D  N  T  V    90
 1366 TTAACTATTTGGAGGCTGCATTTTTATCATCAACGTTCCGTGCTTCTCCGCGCCCAACTAAACCATTGGAGATTGTTATTGCTGGTGCAGgtgatattccgggtc  1470
   91   N  Y  L  E  A  A  F  L  S  S  T  F  R  A  S  P  R  P  T  K  P  L  E  I  V  I  A  G  A  G                  120
 1471 ATCTATATTTGTAGCATTCGTTTTGCTTATTtcgaaggaagcatttttttgctgtattatcatttgaatatttattgtacctgtacatgcctgattaactgtttg  1575
 1576 actacaaaatttgaggttattaagcctcttgtcatttgcgactaaatagaaagtcactgctgattggactaggagcagttacacaagttgtaaattttggacta  1680
 1681 ttcagccatcactctctttatttctgctgttataatgtcttatctctctatcatttgagccatgataagtcactgaacatgctattgtctgtgacatgtccaaat  1785
 1786 ggtgactctgaagtattagaggttgtatatatcaattctccagtgcaagattgatcccatcccactatacctatatattgcagtatgtggagttccaaaatttga  1890
 1891 acactaaattaagttacacaataaggggttgcaacctttaatgcatgcaatctgtgagaacgtgaagcagatatatgatagatcttcatatttcatctttactat  1995
 1996 gctggagcagtgacctctaagctgccttgaacttgttcattgacttggttgcaaaatactgtaacattcattttcaactagtactataatcaactagtagaaaat  2100
 2101 tatattactacctaaagaagtgttatttgcatgatttgttccattgtacttcaaagcatataataaagaaggcaacgcaaggtactagatagctatttcagataa  2205
 2206 atggagcagtaacatgattctatgataattcaacttttggaatatcagGTTTGGGTGGTTTGTCTACAGCAAAATATTTGGCAGATGCTGGTCACAAACCGATA  2310
  121                                                    L  G  G  L  S  T  A  K  Y  L  A  D  A  G  H  K  P  I     138
 2311 CTGCTGGAGGCAAGGGATGTTCTAGGTGGAAAGgttaagaagatcaaaaatttccttttttttaaaaaatcctttttctttttgtgtctttccctattagtagtc  2415
  139 L  L  E  A  R  D  V  L  G  G  K                                                                             149
 2416 ctttttggggtaggggttctgtttgtttcatataaatgtgctgtctttctggtagattctacgtaacctagctaataccttagtctagatcactatgttaagta  2520
 2521 gtttcaaaacaaaaaactctctcttagtttgttttatgttttgatgcatgttttctctaaagattagatgtcctttttgtctacttgcttttgttacatttgcat  2625
 2626 ttaggtgcaagttggtatttgcatacttaataatgcatgttttaattttacttttctgcctattctatgtcctttaatcaagaatcgatttgcacgctatttctt  2730
 2731 cctcgtattttatgcatctatatgctttcaagttgaaatggctagaatggattatcttgcttatggagtcaccaagaattctagttgtataggcaaacaattgat  2835
 2836 ttagtattagtctttatgtgttctactgttgagaatttatatctccctttagGTAGCTGCATGGAAAGATGATGATGGAGATTGGTACGAGACTGGTTTGCATAT  2940
  150                                                     V  A  A  W  K  D  D  D  G  D  W  Y  E  T  G  L  H  I    167
 2941 ATTCTgtaagtttgacctctcattgttatatgtttacgttaatcttctatatactgtcattgtatttttttttttgatctctagccaattagacatctcctatcc  3045
  168  F  F                                                                                                       169
 3046 tcgtttgtcgtttatcgtttatcttttacaaaaatagcctattattgtcagtaaatctgtattttgtctagcttctcctttctcatcttattattcatatagtga  3150
 3151 ctcatacaaattggtgcttgatctcttttaagTTGGGGCTTACCCAAATATTCAGAACCTGTTTGGAGAATTAGGGATTAACGATCGATTGCAATGGAAGGAACA  3255
  170                                   G  A  Y  P  N  I  Q  N  L  F  G  E  L  G  I  N  D  R  L  Q  W  K  E  H    193
 3256 TTCAATGATATTTGCAATGCCAAGCAAGCCAGGAGAATTCAGCCGCTTTGATTTCTCCGAAGCTTTACCCGCTCCTTTAAATGGtgagctaatcacgagtaaatt  3360
  194  S  M  I  F  A  M  P  S  K  P  G  E  F  S  R  F  D  F  S  E  A  L  P  A  P  L  N  G                         221
 3361 tctccctcttgtagttatttgttaaacttccctaataagctgtaaagttgattagaattctaaaaaaaaaatctgtaaaattgataagtcaatcacacctatgg  3465
 3466 gactttactaaccttaaaagagcataaaagttcattacttcttcattggaccttttgtgtgcagctaaaatattaaattctttgatataatttgcagGAATTTA  3570
  222                                                                                                    I  L     223
 3571 GCCATCTTAAAGAATAACGAAATGCTTACATGGCCAGAGAAAGTCAAATTTGCAATTGGACTCTTGCCAGCAATGCTTGGAGGGCAATCTTATGTTGAAGCTCAA  3675
  224 A  I  L  K  N  N  E  M  L  T  W  P  E  K  V  K  F  A  I  G  L  L  P  A  M  L  G  G  Q  S  Y  V  E  A  Q     258
 3676 GATGGGATAAGTGTTAAGGACTGGATGAGAAAGCAAGtgcgtaatcaattatgttgctttttaagtgcataaaccctggctatatagttaccagttcacacacaa  3780
  259 D  G  I  S  V  K  D  W  M  R  K  Q  G                                                                       271
 3781 attttcatttgaggtgttagattgtctactgcctctcttttgctaagagcatacatctagcaaaatctcagcagcagcaagttgctttgcaggGTGTGCCGGACA  3885
  272                                                                                                V  P  D  R   275
 3886 GGGTGACAGATGAGGTGTTCATTGCTATGTCAAAGGCACTCAACTTTATAAACCCTGACGAACTTTCAATGCAGTGCATTTTGATCGCATTGAACAGGTTTCTTC  3990
  276   V  T  D  E  V  F  I  A  M  S  K  A  L  N  F  I  N  P  D  E  L  S  M  Q  C  I  L  I  A  L  N  R  F  L  Q   310
 3991 AGgttagaatcctgatcccacattcaaaacaaagagagagagagaagtacttttgtcccttccagactctcacataacatgaaatctctaactaaggagcttaaa  4095
 4096 tgtgattgggttgtccataagttcatcttgtttggtacgcaacactagattgctgtaggaaacattcataatggcagttttgaagatgaaacttttgaagttacc  4200
 4201 atccgtcaaataatatactttggttttgcttgttacttctggagggacaggtttctctaggttgagggcatattgtttaactgctggagaagtttttgtttaatc  4305
 4306 gacatactctcagaagcttctactagcagaaatggaactgatttatgctaacttggacacttatattaaccatcctctcctcctttttcttttcccattactttt  4410
 4411 ctgaatcatcttcaggaacttctattaaacttctccaaattctaccatgcaaatgggcctcttgtaaattataccctgaatctcatggatgaaggactactttac  4515
 4516 tgattcacttttaaacagctcagttgtcatctagtatgcattttattgtcttaaagaatttaaggtgaactacttataacttggaaacaaacaaaccccgaaaga  4620
 4621 ataagaaatgatggtgaatttctccattcagtttagcttttctgatatattgattataattaacagGAGAAACATGGTTCAAAAATGGCCTTTTTAGATGGTAAT  4725
  311                                                                   E  K  H  G  S  K  M  A  F  L  D  G  N     323
 4726 CCTCCTGAGAGACTTTGCATGCCGATTGTTGAACACATTGAGTCAAAAGGTGGCCAAGTCAGACTGAACTCACGAATAAAAAAGATTGAGCTGAATGAGGATGGA  4830
  324 P  P  E  R  L  C  M  P  I  V  E  H  I  E  S  K  G  G  Q  V  R  L  N  S  R  I  K  K  I  E  L  N  E  D  G     358
 4831 AGTGTCAAGAGTTTTATACTGAGTGACGGTAGTGCAATCGAGGGAGATGCTTTTGTGTTTGCCGCTCCAGgtataatatccattatactagttgattaatccagt  4935
  359 S  V  K  S  F  I  L  S  D  G  S  A  I  E  G  D  A  F  V  F  A  A  P  V                                      382
 4936 tttcagtttcttaatatgagttatgattttgctgattttgatgaaccaattagTGGATATTTTCAAGCTTCTATTGCCTGAAGACTGGAAAGAGATTCCATAT  5040
  383                                                          D  I  F  K  L  L  L  P  E  D  W  K  E  I  P  Y     398
 5041 TTCCAAAAGTTGGAGAAGTTAGTCGGAGTACCTGTGATAAATGTACATATATGGttagtgttgaatatatttacttgttagtgtttgagagatgcatacctatg  5145
  399 F  Q  K  L  E  K  L  V  G  V  P  V  I  N  V  H  I  W                                                        416
 5146 tatgtgcaatgaaatgtctatacatgtatgtttataaatattgacttgcatatcctcagtcttgtgtgaatttcccttgaatgaggaattatggatgttgggttt  5250
 5251 tccgaaacttagatagtaaaagtgcaaatttctgtacatattttaatgagaagaactaccatcgttctggccattagtggctaatattctgctgaaaagaaaaaa  5355
```

```
5356  attctcatcatgaattcatatgtatccttaatatactgaagcgactacagttatcggtaccaaaccaataacgattgtataaattaaactgcaGTTTGACAGAAA  5460
 417                                                                                                F  D  R  K   420
5461  ACTGAAGAACACATATGATCATTTGCTCTTCAGCAGgttaatttcattcactttattcttccaaagcagtttgtgtgtcgatgacttcattttcatttagttccc  5565
 421   L  K  N  T  Y  D  H  L  L  F  S  R                                                                        432
5566  ttcctcctcttccagtgtcattattatttgatacaacattgagtagacggattatcaatttcctgcaaatttcctagtgacttaaatcattagattcatatcatt  5670
5671  gtaaatggctatgtttaatagaaagttaatgaagtatttattggtctaaataagcatttgttgttgcccctccccttcccggtgaaaatccttctctaagcatt  5775
5776  agctggtgagattcagttgaggatagacatgacatgtcacggcaaccacttctccgtggcatgaaaaatggactagagatccaagtcgtcacatttcatcagcta  5880
5881  tatgttaattatgttccctttagtttttaaatgacaccggctggtggagacagataaatagagctccttttttcaattcttaaaaatattgtgttgtttatgatac  5985
5986  ttctattggttccttttcttaatttctctttcatattcatgcagAAGCTCACTGCTCAGTGTGTATGCTGACATGTCTGTTACATGTAAGgtattcatatgtaac  6090
 433                                                 S  S  L  L  S  V  Y  A  D  M  S  V  T  C  K                   447
6091  catttatattgctctcaaattataatctgtggacttgtgtataatttgctgtgaattcgataatgctgttggtaagtgtatgaattcccgctgttgtaatcttca  6195
6196  tttggatatacaatgatattggttttgtatatatggaatagccagggctatgatcctgtctgtgcacgggcccggagccaccacatgaggcaagttacacaggac  6300
6301  aagtaaagaaattttgacgaaacttagccggagatttttttttccatttagctaatttctaagaatttaatgggataaaattcttttattagctatctaggt  6405
6406  tcttgctgccttgggggtggaggggtaagaagagaacttaattcagtgttacctgcagttttggaaagagtctacagatctacttctcagttttattggttgtgg  6510
6511  aaatctgtacaaatataaagcattgccaaaaagttggagaagttaaacttgaaaaaagaggcttatgatagaacactcagacaattgtctatcgtaaaaatattt  6615
6616  tcctttgctgtgtatagtcattatcccattctcttcatgtgtgttttacgaccctctgacatagctggcctctatgcagGAATATTACAACCCCAATCAGTCTAT  6720
 448                                                                                   E  Y  Y  N  P  N  Q  S  M   456
6721  GTTGGAATTGGTTTTTGCACCTGCAGAAGAGTGGATATCTCGCAGCGACTCAGAAATTATTGATGCAACGATGAAGGAACTAGCAACGCTTTTTCCTGATGAAAT  6825
 457   L  E  L  V  F  A  P  A  E  E  W  I  S  R  S  D  S  E  I  I  D  A  T  M  K  E  L  A  T  L  F  P  D  E  I   491
6826  TTCAGCAGATCAAAGCAAAGCAAAAATATTGAAGTACCATGTTGTCAAAACTCCGAGgtaatatagcatttgtccttcatagttgctcatcatgatatgtttttt  6930
 492   S  A  D  Q  S  K  A  K  I  L  K  Y  H  V  V  K  T  P  R                                                   510
6931  cactcttcatacaaatatttagatcatatttgagagtcctgcatttactagaatcgagtccgaatggagtctatacatgaatccttgttgatctgctttttttct  7035
7036  ctcagGTCTGTTTATAAAACTGTGCCAGGTTGTGAACCCTGTCGGCCTTTACAAAGATCCCCAATAGAGGGGTTTTATTTAGCCGGTGACTACACGAAACAGAAA  7140
 511        S  V  Y  K  T  V  P  G  C  E  P  C  R  P  L  Q  R  S  P  I  E  G  F  Y  L  A  G  D  Y  T  K  Q  K    543
7141  TACTTGGCTTCAATGGAAGGCGCTGTCTTATCAGGAAAGCTTTGTGCTCAAGCTATTGTACAGgtaagttctcactggttgtcagcagttgtttctttcattcct  7245
 544  Y  L  A  S  M  E  G  A  V  L  S  G  K  L  C  A  Q  A  I  V  Q                                              564
7246  ggtatgtttgattatattctgtcttctttgttatggtagagaagtgtagaaccatttcttaggcattacatttggtagcagtctctttctgtaagcctgtttctc  7350
7351  attaatttggcgggagagaggggttgtcaagtagtacaacttcacttctgacggagcttctttctctcaacttgtaccttcaacagaaatggatgtcgaacatgc  7455
7456  ctatactaattctgcctagcgtaatagtttctccaaccaaatttaatagtttgtgcacacaccagttagtcaaggagtactgtagtctcccattcttaattatgg  7560
7561  tcttggtttgatcccttggaaaaggagaaaatgttggagggagcaatttccccttaaacggcacgcaaacagataaggttggacgagtgagtttgaatactgga  7665
7666  tggattaacaacaaaaaattggacggctacagatgcagtgttgtgtgtcttagaacctcagtagtagcgtgatatatgcagagagttgtgttattcatctttata  7770
7771  cacctatatttgcagGATTATGAGTTACTTGTTGGACGTAGCCAAAAGAAGTTGTCGGAAGCAAGCGTAGTTTAGCTTTGTGGTTATTATTTAGCTTCTGTACAC  7875
 565                 D  Y  E  L  L  V  G  R  S  Q  K  K  L  S  E  A  S  V  V  *                                  583
7876  TAAATTTATGATGCAAGAAGCGTTGTACACAACATATAGAAGAAGAGTGCGAGGTGAAGCAAGTAGGAGAAATGTTAGGAAAGCTCCTATACAAAAGGATGGCAT  7980
7981  GTTGAAGATTAGCATCTTTTTAATCCCAAGTTTAAATATAAAGCATATTTTATGTACCACTTTCTTTATCTGGGGTTTGTAATCCCTTTATATCTTTATGCAATC  8085
8086  TTTACGTTAGTTAATATCTATCTATCGATATTCTAGTATCTTATACTATAGATCCAACTGAACCAAGAATTATGAACCGTGTCTTCCAGAAATTCTAATAATGAT  8190
8191  GGGAGCAATATAAATATAAGGATGTCTTTGACAATAAAAGGGCGGTGGAAGAGTTATAGTGAAAGCTT                                       8258
```

Fig. 2. Nucleotide sequence of the *Pds* gene of tomato. The deduced amino acid sequence of PDS is written under the first nucleotide of the coding sequence. The introns are shown in lower-case letters. Predicted TATA and CAAT consensus sequences are underlined.

eukaryotic introns in terms of 5′ and 3′ of the splice sites, they generally have less pronounced polypyrimidine stretch at 5 to 15 nucleotides upstream to the 3′ splice site and, instead, they contain A and T [7, 20]. The tomato sequence is similar to other plants in this respect. This A-T-rich sequence is more pronounced in dicots than monocots and is required for efficient splicing [20].

Acknowledgments

Part of this work was supported by a grant to J.H. from the Authority for Research and Development of the Hebrew University of Jerusalem.

References

1. Armstrong GA, Alberti M, Hearst JE: Conserved enzymes mediate the early reactions of carotenoid biosynthesis in nonphotosynthetic and photosynthetic prokaryotes. Proc Natl Acad Sci USA 87: 9975–9979 (1990).
2. Armstrong GA, Alberti M, Leach F, Hearst JE: Nucleotide sequence, organization, and nature of the protein products of the carotenoid biosynthesis gene cluster of *Rhodobacter capsulatus*. Mol Gen Genet 216: 254–268 (1989).
3. Bartley GE, Schmidhauser TJ, Yanofsky C, Scolnik PA: Carotenoid desaturases from *Rhodobacter capsulatus* and *Neurospora crassa* are structurally and functionally conserved and contain domains homologous to flavoprotein disulfide oxidoreductases. J Biol Chem 265: 16020–16024 (1990).
4. Bartley GE, Scolnik PA: Carotenoid biosynthesis in photosynthetic bacteria: Genetic characterization of the *Rhodobacter capsulatus crtI* protein. J Biol Chem 264: 13109–13113 (1989).

5. Bartley GE, Viitanen PV, Pecker I, Chamovitz D, Hirschberg J, Scolnik PA: Molecular cloning and expression in photosynthetic bacteria of a soybean cDNA coding for phytoene desaturase, an enzyme of the carotenoid biosynthesis pathway. Proc Natl Acad Sci USA 88: 6532–6536 (1991).
6. Britton G: Biosynthesis of carotenoids. In Goodwin TW (ed) Plant Pigments, pp. 133–180. Academic Press, New York (1988).
7. Brown JWS: A catalogue of splice junction and putative branch point sequences from plant introns. Nucleic Acids Res 14:9549–9559 (1986).
8. Chamovitz D, Pecker I, Hirschberg J: The molecular basis of resistance to the herbicide norflurazon. Plant Mol Biol 16: 967–974 (1991)
9. Chamovitz D, Pecker I, Sandmann G, Böger P, Hirschberg J: Cloning a gene for norflurazon resistance in cyanobacteria. Z Naturforsch 45c:482–486 (1990).
10. Chamovitz D, Sandmann G, Hirschberg J: Molecular and biochemical characterization of herbicide resistant mutants of cyanobacteria reveals that phytoene desaturase is a rate limiting step in carotenoid biosynthesis. J Biol Chem, in press (1993).
11. Giuliano G, Bartley GE, Scolnik P: Regulation of carotenoid biosynthesis during tomato development. Plant Cell 5: 379–387 (1993).
12. Hugueney P, Römer S, Kuntz M, Camara B: Characterization and molecular cloning of a flavoprotein catalyzing the synthesis of phytofluene and ζ-carotene in *Capsicum* chromoplasts. Eur J Biochem 209: 399–407 (1992).
13. Linden H, Misawa N, Chamovitz D, Pecker I, Hirschberg J, Sandmann G: Functional complementation in *Escherichia coli* of different phytoene desaturase genes and analysis of accumulated carotenes. Z Naturforsch 46c: 1045–1051 (1991).
14. Martinez-Ferez IM, Vioque A: Nucleotide sequence of the phytoene desaturase gene from *Synechocystis* sp. PCC 6803 and characterization of a new mutation which confers resistance to the herbicide norflurazon. Plant Mol Biol 18: 981–983 (1992).
15. Misawa N, Nakagawa M, Kobayashi K, Yamano S, Izawa I, Nakamura K, Harashima K: Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products in *Escherichia coli.* J Bact 172: 6704–6712 (1990).
16. Pecker I, Chamovitz D, Linden H, Sandmann G, Hirschberg J: A single polypeptide catalyzing the conversion of phytoene to z-carotene is transcriptionally regulated during tomato fruit ripening. Proc Natl Acad Sci USA 89: 4962–4966 (1992).
17. Pecker I, Chamovitz D, Mann V, Sandmann G, Böger P, Hirschberg J: Molecular characterization of carotenoid biosynthesis in plants: The phytoene desaturase gene in tomato. In: Murata N (ed) Research in Photosynthesis, vol. 3, pp. 11–18. Kluwer Academic Publishers, Dordrecht (1993).
18. Ray J, Moureau P, Bird C, Bird A, Grierson D, Maunders M, Truesdale M, Bramley P, Schuch W: Cloning and characterization of a gene involved in phytoene synthesis from tomato. Plant Mol Biol 19: 401–404 (1992).
19. Sanger F, Nicklen S, Coulsen AR: DNA sequencing with chain termination inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
20. Simpson CG, Brown JWS: Efficient splicing of an AU-rich antisense intron sequence. Plant Mol Biol 21: 205–211 (1993).