Songklanakarin J. Sci. Technol.
35 (5), 517-527, Sep. - Oct. 2013

SJST SONGKLANAKARIN JOURNAL OF SCIENCE AND TECHNOLOGY
http://www.sjst.psu.ac.th

Original Article

cDNA cloning and expression analyses of phytoene synthase 1, phytoene desaturase and ζ-carotene desaturase genes from *Solanum lycopersicum* KKU-T34003

Krittaya Supathaweewat[1,2] and Preekamol Klanrit[1,3]*

[1] *Department of Biotechnology, Faculty of Technology,*

[2] *Graduate School,*

[3] *Fermentation Research Center for Value Added Agricultural Products (FerVAAP), Khon Kaen University, Mueang, Khon Kaen, 40002 Thailand.*

Received 18 November 2012; Accepted 1 July 2013

Abstract

We report on the cloning of *Psy1*, *Pds* and *Zds* cDNAs encoding the enzymes responsible for lycopene biosynthesis, namely phytoene synthase 1 (PSY1), phytoene desaturase (PDS) and ζ-carotene desaturase (ZDS), respectively, from high-lycopene tomato cultivar, *Solanum lycopersicum* KKU-T34003. DNA sequence analyses showed that the complete open reading frames of *Psy1*, *Pds* and *Zds* cDNAs were 1,239, 1,752 and 1,767 base pairs in length and encoded proteins of 412, 583 and 588 amino acids, respectively. Phylogenetic and the conserved domain analyses suggest that PSY1, PDS and ZDS from *S. lycopersicum* KKU-T34003 potentially have similar structures and biological functions to the corresponding proteins from other plants. Gene expression studies showed that *Psy1* was expressed only in the petal and the breaker fruit, whereas the expressions of *Pds* and *Zds* were observed in the petal, the breaker fruit and the leaf. The highest expression level for all genes was detected in the breaker-stage fruit, suggesting that carotenoid accumulation was developmentally regulated in the chromoplast-containing tissues.

Keywords: phytoene synthase, phytoene desaturase, ζ-carotene desaturase, lycopene, *Solanum lycopersicum*

1. Introduction

Carotenoids are yellow, orange and red lipid-soluble pigments found in nature. All photosynthetic and many non-photosynthetic organisms can synthesize carotenoids. In plants, the bright colors of carotenoids help attract pollinating insects and other animals, and thus facilitate seed dispersal. Carotenoids are essential structural components of the photosynthetic apparatus, and protect the cells from photooxidative damage as well as serve as precursors for abscisic acid (Krinsky, 1989; Havaux, 1998; Lindgren *et al*., 2003). Apart from their roles in plants, they contribute signi ficantly to human health and nutrition particularly their roles in pro-vitamin A and antioxidant activities (Lee *et al*., 1981). Clinical research evidence shows that many carotenoids, especially β-carotene and lycopene, decrease the risk of certain cancers, age-related macular degeneration and coronary heart disease (Bendich, 1994). Carotenoids are C_{40} tetraterpenes, in which each molecule may contain up to 15 conjugated double bonds (Enfissi *et al*., 2006). Examples of common carotenoids such as β-carotene, lycopene, lutein and zeaxanthin are shown in Figure 1.

Lycopene is a natural red-colored pigment, which has been of significant interest regarding its antioxidant benefits

* Corresponding author.
Email address: kpreek@kku.ac.th

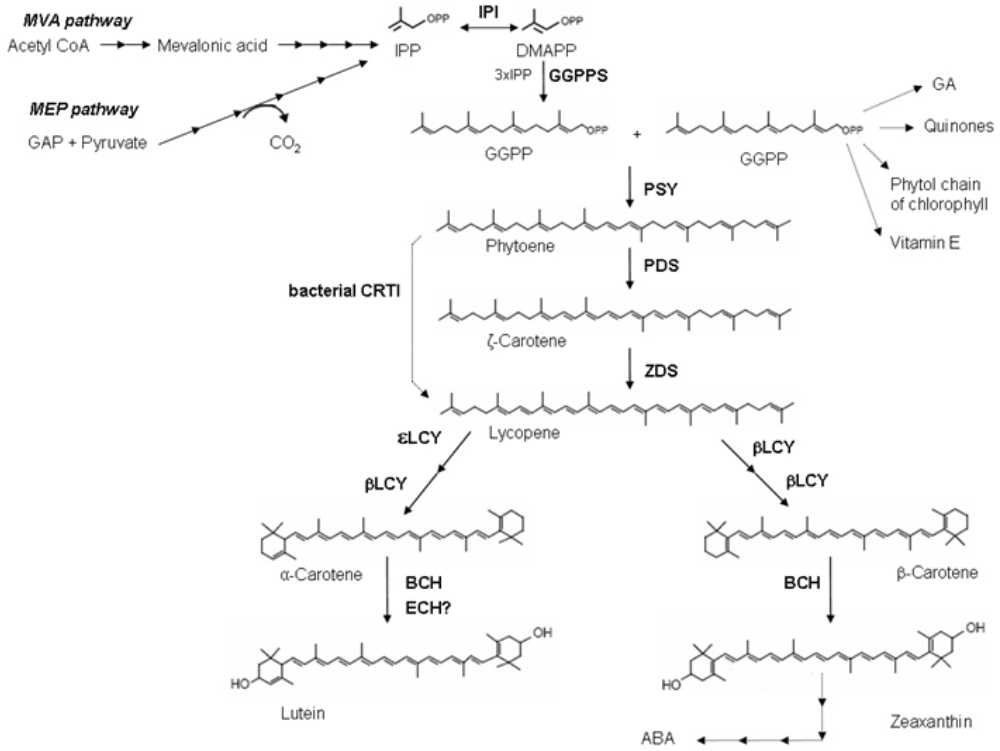

Figure 1. Carotenoid biosynthetic pathway.
Abbreviations: IPP (isopentenyl pyrophosphate); DMAPP (dimethyllallyl pyrophosphate); GGPP (geranylgeranyl pyrophosphate); GA (gibberellic acid); ABA (abscisic acid); GAP (glyceroldehyde-3-phosphate); MVA (mevalonic acid); MEP (methylerythritol pathway); IPI (IPP isomerase); GGPPS (GGPP synthase); PSY (phytoene synthase); PDS (phytoene desaturase); ZDS (æ-carotene desaturase); CRTI (bacterial phytoene desaturase); β- and ε-LCY β- and ε-lycopene cyclase); BCH (β-carotene hydroxylase); ECH (ε-carotene hydroxylase) (modified from Naik *et al*., 2003; Enfissi *et al*., 2006; Ji *et al*., 2009)

in treating and reducing the risk of various diseases such as atherosclerosis, osteoporosis and certain cancers (Rao and Agarwal, 1999). It is commonly present in tomatoes (*Solanum lycopersicum*) as well as other red fruits and vegetables, and also found in certain bacteria, algae and fungi (Young and Britton, 1993). In plants, especially in tomato fruits, lycopene is synthesized and accumulated within a chromoplast. Figure 1 demonstrates the carotenoid biosynthetic pathway with the encoding genes and enzymes. Carotenoids are built from the 5-carbon compound isopentenyl diphosphate (IPP), which is isomerized to dimethylallyl diphosphate (DMAPP) by IPP isomerase (IPI). GGPP synthase (GGPPS) then catalyzes the sequential addition of three IPP molecules to a DMAPP molecule, giving the 20-carbon molecule called geranylgeranyl pyrophosphate (GGPP), which is a precursor for biosynthesis of other compounds such as gibberellins, vitamin E, quinines and phytol chain of chlorophyll (Naik *et al*., 2003). Phytoene synthase (PSY) catalyzes the first committed step in the carotenoid biosynthetic pathway, producing a colorless compound called phytoene (Burkhardt *et al*., 1997). Then, colorless phytoene undergoes four desaturation reactions, which are catalyzed by phytoene desaturase (PDS) and ζ-carotene desaturase (ZDS), resulting in the formation of ζ-carotene (pale-yellow) and lycopene (red), respectively (Figure 1) (Sandmann, 1994; Matthews *et al*., 2003). Then, lycopene is cyclized twice by two individual cyclases, yielding α and β carotenes, which are subsequently converted to other types of carotenoids (Hirschberg, 2001).

It has been demonstrated that PSY is a rate-limiting enzyme in tomato fruits (Bramley *et al*., 1992), in canola (*Brassica napus*) seeds (Shewmaker *et al*., 1999), in Golden Rice (Paine *et al*., 2005), in *Arabidopsis thaliana* seeds (Lindgren *et al*., 2003) and in marigold flowers (Moehs *et al*., 2001). Phytoene synthase genes and cDNAs have been cloned from many plants and bacteria; for example, *A. thaliana* (Scolnik and Bartley, 1994), *Cucumis melo* (Karvouni *et al*., 1995), *Narcissus pseudonarcissus* (Schledz *et al*., 1996), *Solanum lycopersicum* (Giuliano *et al*., 1993; Giorio *et al*., 2008), tobacco (*Nicotiana tabacum*) (Busch *et al*., 2002), pepper (*Capsicum annuum*) (Romer *et al*., 1993), *Erwinia uredovora* (Misawa *et al*., 1990), *E. herbicola* (Perry *et al*., 1986) and *Rhodobacter capsulatus* (Armstrong *et al*., 1990). Because of this rate-limiting characteristic, PSY has gained much interest to be cloned and targeted for transformation to increase carotenoid levels. The genes encoding PDS and ZDS have also been cloned and characterized from many plants, bacteria, algae and fungi; for example, maize (Matthews *et al*., 2003), tobacco (Busch *et al*., 2002), papaya

(Yan *et al*., 2011), tomato (Mann *et al*., 1994), soybean (Bartley *et al*., 1991), *Chlorella protothecoides* (Li *et al*., 2011), *E. uredovora* (Misawa *et al*., 1990), *E. herbicola* (Perry *et al*., 1986) and *R. capsulatus* (Armstrong *et al*., 1990). They have also been targeted for genetic engineering to produce plants or microorganisms with increased amount of carotenoids.

The tomato fruits are one of the richest sources of carotenoids, especially lycopene. However, there are significantly different levels of lycopene content in various tomato cultivars. The tomato, *S. lycopersicum* KKU-T34003, was used in this study due to its high lycopene content of 17.1 mg/100 g of fresh weight, which was approximately 3-4 times higher than that of the regular tomato (Clinton, 1998). At present, genetic manipulation of the carotenoid pathway is of significant interest in order to generate either plants or microorganisms, which are able to synthesize large amounts of high-value carotenoids including lycopene. Therefore, isolation and analyses of the genes encoding enzymes responsible for lycopene biosynthesis are required as a first step in order to provide an insight into how lycopene is synthesized.

This work is the first report on the cloning of *Psy1*, *Pds* and *Zds* cDNAs encoding PSY1, PDS and ZDS, respectively, from a high-lycopene tomato cultivar, *S. lycopersicum* KKU-T34003. Although *Psy1*, *Pds* and *Zds* cDNAs have already been cloned from other tomato varieties, different sources of the genes or cDNAs potentially affect the expression level and, as a consequence, have an impact on the accumulation of lycopene. The gene expression patterns of *Psy*, *Pds* and *Zds* in different tomato tissues were also demonstrated. Moreover, the phylogenetic tree was constructed to illustrate the evolutionary relationship of PSY, PDS and ZDS families of proteins.

2. Materials and Methods

2.1 Plant material

The tomato seeds (*S. lycopersicum* KKU-T34003) were obtained through the courtesy of Associate Professor Dr. Suchila Techawongstien, the Plant Breeding Research Center for Sustainable Agriculture, Faculty of Agriculture, Khon Kaen University, Thailand.

2.2 Bacterial strain and vector

The bacterial strain and vector used in this study were *Escherichia coli* strain TOP10 (Invitrogen™ Life Technologies Corporation) and pGEM-T® easy vector (Promega Corporation), respectively.

2.3 Media and growth conditions

The tomato seeds were placed in regular soil and the soil pots with a plastic cover were placed in a greenhouse. The plastic cover was removed when the seeds germinated. The plants were watered every day and fertilized twice a week. Luria Bertani (LB) medium was used as a growth medium for *E. coli*, which contained 1.0% tryptone, 0.5% yeast extract and 1.0% NaCl. For the cloning, *E. coli* was cultivated at 37°C in LB medium, which was supplemented with 100 µg/ml of ampicillin. For the screening of recombinants, transformed *E. coli* cells were spread on LB agar plates with 0.5 mM of isopropyl-β-D-thiogalactopyranoside (IPTG) and 80 µg/ ml of 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) and cultivated overnight at 37°C.

2.4 Cloning of *Psy1*, *Pds* and *Zds* cDNAs

Total RNA was extracted from the tomato fruit at the breaker stage using TRIzol® reagent according to manufacturer's instruction (Invitrogen™ Life Technologies Corporation). The extracted RNA was reverse transcribed into cDNA by Moloney Murine Leukemia Virus (M-MuLV) reverse transcriptase (Vivantis Tecnologies), and then the cDNAs were amplified by Polymerase Chain Reaction (PCR) using Easy-A high fidelity PCR cloning enzyme (Stratagene Agilent technologies). The primers (Table 1) used in PCR experiment were designed based on the sequence information available from The National Center for Biotechnology Information, NCBI (www.ncbi.nlm.nih.gov). PCR amplification was carried out under the following conditions: an initial denaturation at 94°C for 2 min; 35 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 40 s and extension at 72°C for 2 min, followed by a final extension at 72°C for 7 min. The PCR products were analyzed on 1% agarose gel stained with ethidium bromide, then purified using GF-1 AmbiClean Kit (Vivantis Tecnologies) and separately ligated into pGEM-T® easy vector. The ligation products were transformed into *E. coli* TOP10 competent cells. The clones were initially selected by blue-white colony selection, followed by *Eco*RI digestion of the extracted plasmids according to the standard protocol of the Molecular cloning: A laboratory manual (Sambrook and Russell, 2001). The recombinant plasmids were purified using Wizard® *Plus* SV minipreps DNA purification system (Promega Corporation), and sequenced by First BASE Laboratories Sdn Bhd, Malaysia.

2.5 Sequence analyses

The DNA sequences were edited, and the number of deduced amino acids and predicted molecular weights were analyzed using bioinformatic tools from BioEdit sequence analysis program and ExPASy (http://expasy.org/tools). The nucleotide and amino acid sequences were compared with the corresponding sequences in NCBI database using the Basic Local Alignment Search Tool, BLAST (http://blast.ncbi.nlm. nih.gov). The amino acid sequence similarities were calculated using Matrix Global Alignment Tool, MatGAT version 2.02 (http://bitincka.com/ledion/ matgat), ClustalW2 (http://www/ebi.ac.uk/Tools/msa/ clustalw2) and Water EM-

Table 1. Primers used in this study for *Psy1*, *Pds* and *Zds* cDNA cloning and gene expression.

Gene	Primer name	Primer sequence (5′-3′)	Application
Psy1	SolPsyFor1_1	CACCATGTCTGTTGCCTTGTTATG	Cloning
	SolPsyRev1	TCTTTGAAGAGAGGCAGTTTTTG	Cloning
Pds	SolPdsFor1_1	CACCATGCCTCAAATTGGACTT	Cloning
	SolPdsRev1	AACTACGCTTGCTTCCGAC	Cloning
Zds	SolZdsFor1_1	CACCATGGCTACTTCTTCAGCTTATC	Cloning
	SolZdsRev1	GACAAGACTCAACTCATCAG	Cloning
Psy1	SolPsyFor1_2	GGTGGAAAGCAAACTAATAATGG	Gene expression
	SolPsyRev1	TCTTTGAAGAGAGGCAGTTTTTG	Gene expression
Pds	SolPdsFor1_2	GACTGGATGAGAAAGCAAGGTG	Gene expression
	SolPdsRev1	AACTACGCTTGCTTCCGAC	Gene expression
Zds	SolZdsFor1_2	CTGTGATAACATGAGTGCTCG	Gene expression
	SolZdsRev1	GACAAGACTCAACTCATCAG	Gene expression
EF-1α	SolEF1For	CTGAGGCTCTTGACCAGATTAAC	Gene expression
	SolEF1Rev	CTTCCCCTTCTTCTGGGCAG	Gene expression

BOSS sequence alignment (http://www.ebi.ac.uk/Tools/psa). The conserved domains were analyzed using the Conserved Domain Database, CDD (http://www.ncbi.nlm.nih.gov/Structure/cdd/ wrpsb.cgi). The ChloroP 1.1 Prediction Server program (Emanuelsson *et al*., 1999) was employed to identify PSY1, PDS and ZDS signal peptides and predict their cleavage sites. Furthermore, the phylogenetic relationships among the PSY, PDS and ZDS proteins were analyzed using Molecular Evolutionary Genetics Analysis, MEGA software version 4.0.2 (Tamura *et al*., 2007). The neighbor-joining (NJ) method was used to construct the tree. A bootstrap test based on 1,000 replicates was performed to determine the reliability of the phylogenetic tree.

2.6 Gene expression analyses

Total RNA was extracted from leaf, petal and breaker-stage fruit of *S. lycopersicum* KKU-T34003 by TRIzol® reagent as previously mentioned in the cloning section (2.4). The RNA concentrations were quantified using the spectrophotometer DU 730 (Beckman Coulter). Then, the normalized RNA was used as templates for *Psy1*, *Pds* and *Zds* gene expression using AccessQuick™ Reverse Transcription-Polymerase Reaction (RT-PCR) system (Promega Corporation), with the gene-specific primer pairs (Table 1). Elongation factor 1-α (*EF1-α*) was used as an internal control in the experiment (Pokalsky *et al*., 1989). PCR product samples were collected at the PCR cycle numbers 17, 20, 23, 26, 29, 32 and 35. Then, the gene expression patterns were analyzed by agarose gel electrophoresis using 1% agarose gel stained with ethidium bromide. The level of gene expression was determined using the Image Lab™ 2.0 software associated with Molecular Imager® Gel Doc™ XR+ (Bio-Rad Laboratories, Inc.) according to the band intensity, which was translated into relative quantity and compared with the reference band.

3. Results and Discussion

3.1 Cloning of *Psy1*, *Pds* and *Zds* cDNAs and sequence analyses

Figure 2 illustrates the expected PCR product sizes of *Psy1*, *Pds* and *Zds*. DNA sequence analyses showed that the complete open reading frames of *Psy1, Pds* and *Zds* cDNAs were 1,239, 1,752 and 1,767 base pairs in length and encoded proteins of 412, 583 and 588 amino acids (Figures 3, 4 and 5), with the predicted molecular weights of approximately 47, 65 and 65 kDa, respectively. The GenBank accession numbers of *Psy1*, *Pds* and *Zds* cDNAs are KC767847 (*Psy1*), KC767848 (*Pds*) and KC767849 (*Zds*). Analysis of the homology of the deduced amino acid sequences of PSY1, PDS and ZDS revealed various degrees of similarity to the corresponding proteins in other plants, with a very high degree of similarity, ranging from 85.9-100.0% (PSY1), 95.9-100.0% (PDS) and 97.1-99.3% (ZDS), to plant proteins in the family Solanaceae. PSY1, PDS and ZDS from *S. lycopersicum* KKU-T34003 showed the highest sequence similarity with the corresponding proteins from tomato (*S. lycopersicum*) [Accession numbers: NP_001234812 (PSY1); AAA08868 (PDS); ABR 57231 (ZDS)], followed by red pepper (*Capsicum annuum*)

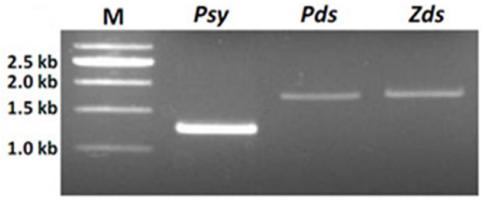

Figure 2. Agarose gel electrophoresis showing RT-PCR amplified products of *Psy1*, *Pds* and *Zds* cDNAs.

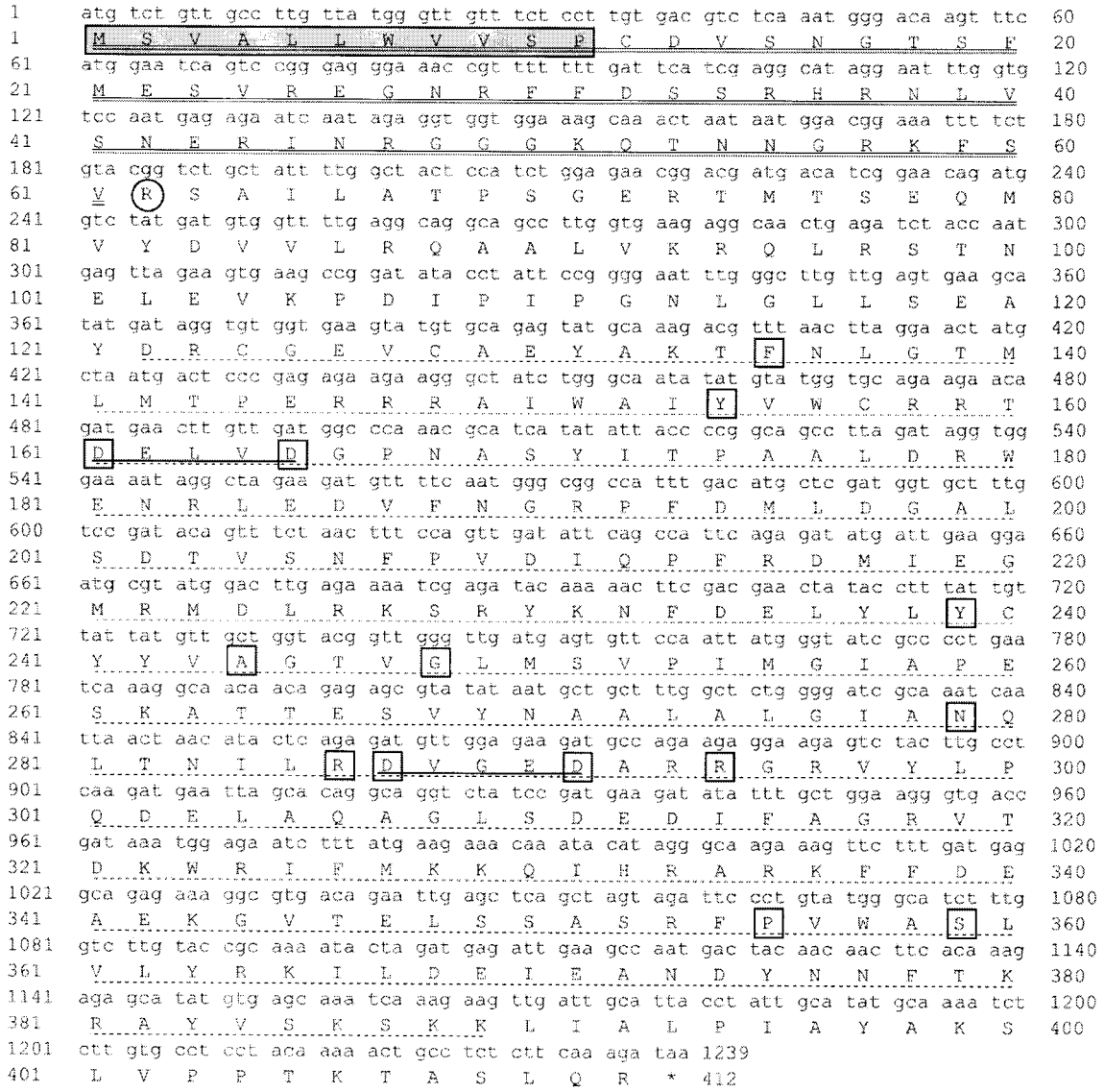

Figure 3. Nucleotide and deduced amino acid sequences of *Psy1* cDNA (accession number: KC767847). A double underline, a dotted underline and a star indicate the putative plastid transit peptide (TP) which directs the gene product to plastids, the conserved region of PSY called trans-isoprenyl diphosphate synthase and the stop codon, respectively. The underlined amino acids indicate two aspartate-rich regions (DXXXD). Square boxes show the important catalytic residues, whereas the circle is the location of the potential TP cleavage site. The highlighted rectangular box indicates the motif present in dicotyledons.

[Accession numbers: ACE78189 (PSY1); CAA48195 (PDS); CAA61985 (ZDS)] and tobacco (*Nicotiana tabacum*) [Accession numbers: ADK25054 (PSY1); ABY25272 (PDS); AEG7 3891 (ZDS)].

The conserved domain of PSY, namely trans-isoprenyl diphosphate synthase, which was predicted to be involved in the catalytic activity and substrate recognition, was found in the region from aspartate$_{122}$ to lysine$_{389}$ (Hirschberg, 2001; Cunningham, 2002). As predicted by ChloroP 1.1 program (Emanuelsson *et al*., 1999), another functional region designated the plastid transit peptide (TP), starting from the first amino acid to arginine$_{62}$ was also observed (Giorio *et al*., 2008), with the transit peptide cleavage site between residues 62R and 63S (Figure 3). Moreover, the catalytic sites were predicted to consist of a large central cavity formed by mostly antiparallel alpha helices with two aspartate-rich regions (DXXXD, where X encoded any amino acid), and another 14 catalytic amino acid residues found within the sequence were also important for catalyzing the reaction. The amino acid residues at the N terminus, MSVALLWVVSP, showed the signature sequence of PSY proteins present in dicotyledonous plants (Giorio *et al*., 2008).

PDS and ZDS are also crucial for lycopene biosynthesis in plants. Both PDS and ZDS belong to a dinucleotide-

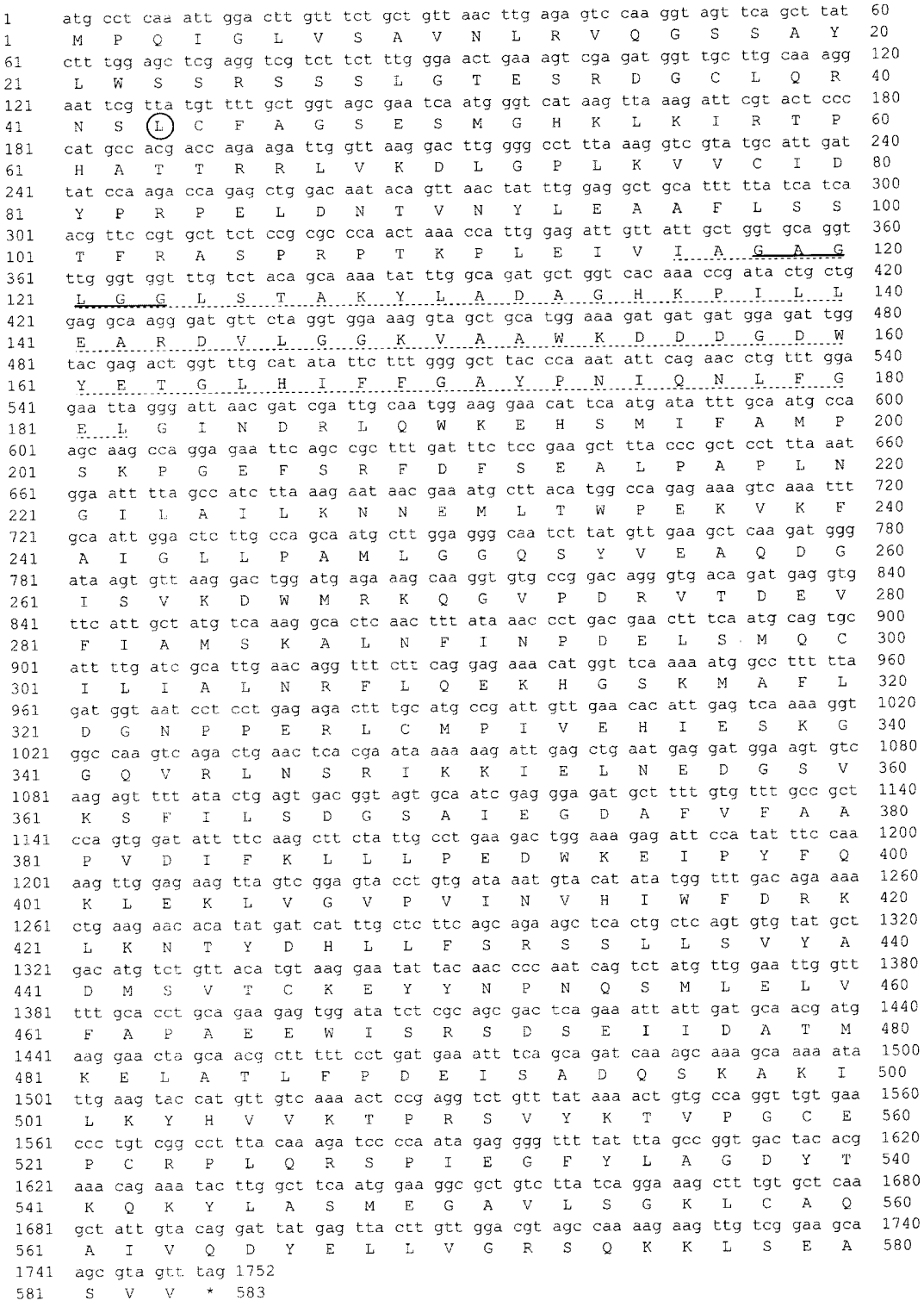

Figure 4. Nucleotide and deduced amino acid sequences of *Pds* cDNA (accession number: KC767848). A dotted underline and a star indicate the conserved region of PDS (dinucleotide binding domain, GXGXXG) and the stop codon, respectively. The circle is the location of the potential TP cleavage site.

binding domain family, which is predicted to be involved in the catalytic activities of dehydrogenases (Bartley *et al.*, 1990). Our study showed that this domain was found near the N terminus, within the region from isoleucine$_{116}$ to leucine$_{182}$ of PDS (Figure 4) and isoleucine$_{84}$ to valine$_{150}$ of ZDS (Figure 5). The consensus Gly-rich sequence with the pattern,

```
1     atg gct act tct tca gct tat ctt tct tgt cct gca act tct gct act gga aag aaa cat  60
1      M   A   T   S   S   A   Y   L   S   C   P   A   T   S   A   T   G   K   K   H  20
61    gtt ttc cca aat ggg tca cct gga ttc ttg gtt ttt ggt ggt acc cgt ttg tcc aac cgg  120
21     V   F   P   N   G   S   P   G   F   L   V   F   G   G   T   R   L   S   N   R  40
121   tta gtg acc cga aag tcg gtt att cgg gct gat ttg gat tct atg gtt tct gat atg agt  180
41     L   V   T   R   K   S   V   I   R   A   D   L   D   S   M   V   S   D   M   S  60
181   acc aac gct cca aaa ggg cta ttt cca ccc gag cct gaa cat tat cgg ggg cca aag ctg  240
61     T   N   A   P   K   G   L   F   P   P   E   P   E   H   Y   R   G   P   K   L  80
241   aaa gta gct att att gga gct ggg ctt gca ggc atg tcg act gct gtg gag ctc ttg gat  300
81     K   V   A   I   I   G   A   G   L   A   G   M   S   T   A   V   E   L   L   D  100
301   caa gga cat gag gtg gat ata tac gaa tca agg act ttt att ggt ggg aaa gtg ggt tct  360
101    Q   G   H   E   V   D   I   Y   E   S   R   T   F   I   G   G   K   V   G   S  120
361   ttt gtt gat aga cgt ggg aac cac att gaa atg gga ctg cac gtg ttc ttt ggt tgt tat  420
121    F   V   D   R   R   G   N   H   I   E   M   G   L   H   V   F   F   G   C   Y  140
421   aat aat ctg ttc cgt ctg ttg aaa aag gtg ggt gct gaa aaa aat ctg cta gtg aag gag  480
141    N   N   L   F   R   L   L   K   K   V   G   A   E   K   N   L   L   V   K   E  160
481   cat act cac aca ttt gta aat aaa ggg ggt gaa ata ggg gaa ctt gat ttc cgc ttt cca  540
161    H   T   H   T   F   V   N   K   G   G   E   I   G   E   L   D   F   R   F   P  180
541   gtt gga gca ccc tta cat gga att aat gca ttt ctg tct act aat cag tta aag att tat  600
181    V   G   A   P   L   H   G   I   N   A   F   L   S   T   N   Q   L   K   I   Y  200
601   gat aaa gct aga aat gct gta gct ctt gcc ctt agt cca gtg gtg cgg gct tta gtt gat  660
200    D   K   A   R   N   A   V   A   L   A   L   S   P   V   V   R   A   L   V   D  220
661   ccg gat ggt gca ttg cag cag ata cgc gat cta gat aat gta agc ttt tct gag tgg ttt  720
221    P   D   G   A   L   Q   Q   I   R   D   L   D   N   V   S   F   S   E   W   F  240
721   ctg tct aaa ggt ggg acg cgt gct agc aac cag agg gtg tgg gat cct gtt gca tat gct  780
241    L   S   K   G   G   T   R   A   S   N   Q   R   V   W   D   P   V   A   Y   A  260
781   ctt gga ttc att gac tgt gat aac atg agt gct cgg tgt atg ctc act ata ttt gca tta  840
261    L   G   F   I   D   C   D   N   M   S   A   R   C   M   L   T   I   F   A   L  280
841   ttt gcc aca aaa aca gag gct tcc cta tta cgc atg ctt aaa ggt tct cct gac gtt tat  900
281    F   A   T   K   T   E   A   S   L   L   R   M   L   K   G   S   P   D   V   Y  300
901   ttg agt ggt cca att aag aag tac atc atg gac aaa ggg ggc agg ttc cat ctg agg tgg  960
301    L   S   G   P   I   K   K   Y   I   M   D   K   G   G   R   F   H   L   R   W  320
961   gga tgc aga gag gta ctc tat gag acg tcc tct gat gga agc atg tat gtt agt ggg ctt 1020
321    G   C   R   E   V   L   Y   E   T   S   S   D   G   S   M   Y   V   S   G   L  340
1021  gcc atg tca aag gcc act cag aag aaa att gta aag gct gat gca tat gtg gct gca tgt 1080
341    A   M   S   K   A   T   Q   K   K   I   V   K   A   D   A   Y   V   A   A   C  360
1081  gat gtc cct gga att aaa aga ttg gtt cct cag aag tgg agg gaa ttg gaa ttc ttt gac 1140
361    D   V   P   G   I   K   R   L   V   P   Q   K   W   R   E   L   E   F   F   D  380
1141  aac att tac aaa ttg gtc gga gtg cct gtt gtt acc gta caa cta cgc tac aat ggc tgg 1200
381    N   I   Y   K   L   V   G   V   P   V   V   T   V   Q   L   R   Y   N   G   W  400
1201  gtt aca gag ttg cag gac ttg gag cgt tcg agg caa ttg aag cgc gct gca gga ttg gac 1260
401    V   T   E   L   Q   D   L   E   R   S   R   Q   L   K   R   A   A   G   L   D  420
1261  aat ctc ctc tat acg cca gat gca gat ttc tct tgc ttt gca gat ctt gca ttg gca tct 1320
421    N   L   L   Y   T   P   D   A   D   F   S   C   F   A   D   L   A   L   A   S  440
1321  cca gat gat tac tac att gag gga caa ggc tca ttg ctt caa tgt gtc ctt aca cct ggt 1380
441    P   D   D   Y   Y   I   E   G   Q   G   S   L   L   Q   C   V   L   T   P   G  460
1381  gac cct tac atg cct cta tca aat gat gaa atc att aaa aga gtt aca aag cag gtt ttg 1440
461    D   P   Y   M   P   L   S   N   D   E   I   I   K   R   V   T   K   Q   V   L  480
1441  gca tta ttt cct tcg tcc caa ggt ctt gag gtt acc tgg tca tca gtt gtg aag ata gga 1500
481    A   L   F   P   S   S   Q   G   L   E   V   T   W   S   S   V   V   K   I   G  500
1501  caa tct tta tat cgt gaa gga cct ggt aaa gac cca ttc aga cct gat cag aag acg cca 1560
501    Q   S   L   Y   R   E   G   P   G   K   D   P   F   R   P   D   Q   K   T   P  520
1561  gtg gaa aat ttc ttt ctt gct ggc tca tat aca aaa cag gac tac atc gat agc atg gaa 1620
521    V   E   N   F   F   L   A   G   S   Y   T   K   Q   D   Y   I   D   S   M   E  540
1621  gga gca act ctt tca ggt agg caa gct tct gca tac ata tgt aat gtt gga gag cag ctg 1680
541    G   A   T   L   S   G   R   Q   A   S   A   Y   I   C   N   V   G   E   Q   L  560
1681  atg gcg ttg cgt aaa aag atc act gct gct gag ttg aat gac atc tct aaa ggt gtg tcc 1740
561    M   A   L   R   K   K   I   T   A   A   E   L   N   D   I   S   K   G   V   S  580
1741  cta tct gat gag ttg agt ctt gtc tga 1767
581    L   S   D   E   L   S   L   V   *  588
```

Figure 5. Nucleotide and deduced amino acid sequences of *Zds* cDNA (accession number: KC767849). A dotted underline and a star indicate the conserved region of ZDS (dinucleotide binding domain, GXGXXG) and the stop codon, respectively. The circle is the location of the potential TP cleavage site.

GXGXXG, which has been proposed to play an important role in binding to the dinucleotides [FAD/NAD(P)], was found within the sequence (Dailey and Dailey, 1998). Furthermore, the plastid transit peptides were also observed at the N terminus of PDS and ZDS, with predicted transit peptide cleavage sites between residues 43L/44C and 49R/50A,

respectively (Figures 4 and 5). Similar sequence characteristics were reported in pepper (Hugueney *et al*., 1992), maize (Matthews, *et al*., 2003), papaya (Yan *et al*., 2011), tomato (Giorio *et al*., 2008) and tobacco (Busch *et al*., 2002). These findings provided a strong support that PSY1, PDS and ZDS from *S. lycopersicum* KKU-T34003 might have similar structures and functions to the corresponding proteins from other plants.

The phylogenetic relationships among the deduced amino acid sequences of PSY1, PDS or ZDS from *S. lycopersicum* KKU-T34003 and the corresponding proteins of several plants and bacteria illustrate three distinct clusters of proteins (Figure 6). In tomato, there are two active forms of PSY, namely PSY1 and PSY2, which function mainly in the fruit and leaf tissue, respectively (Bramley *et al*., 1992; Fraser *et al*., 1999). Considering the results from the sequence alignment of PSYs and the phylogenetic tree, the finding of 2 active forms of PSY could also be explained that the presence of tomato KKU-T340003 *Psy1* and *Psy2* in the genome was possibly a result of a gene duplication event (Zhang, 2003), as also observed in other plants in the families Solanaceae and Apiaceae such as tobacco (Busch *et al*., 2002) and carrot (Just *et al*., 2007). PDS was more closely related to ZDS, as supported by the previous finding that these two enzymes were functionally and structurally similar (Matthews *et al*., 2003). In particular, PSY, PDS and ZDS are more closely related to the corresponding enzymes in plants than those in bacteria (CRTB and CRTI). The high bootstrap values demonstrated on the phylogenetic tree confirm that they are distinct groups of organisms.

3.2 Gene expression analyses

The expression study showed that *Psy1* gene was expressed only in the petal and the breaker fruit, with the highest expression level in the breaker fruit (Figure 7A);

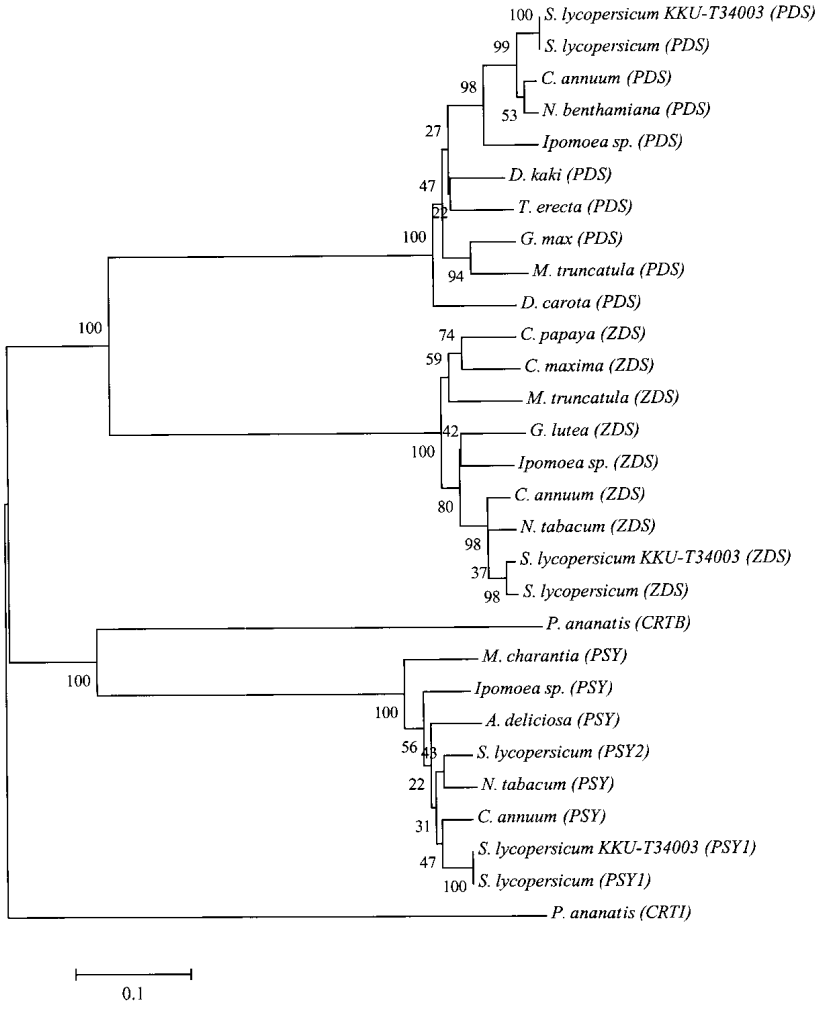

Figure 6. Phylogenetic analysis of the deduced amino acid sequences of PSY1, PDS and ZDS from *S. lycopersicum* KKU-T34003 and those of other plants including bacteria from GenBank database. Clustering was performed by the neighbor-joining with the software package, MEGA version 4.0. Bootstrap values are indicated near the base of each branch (1,000 replicates). Horizontal branch lengths represent relative evolutionary distances with the scale bar corresponding to 0.1 amino acid substitution per site.

however, the expression was not detected in the leaf tissue (data not shown). The differential expression in different tissues could be explained in part by the presence of two active forms of PSY, namely, PSY1 and PSY2, which were found in fruit and leaf of tomato, respectively (Bramley *et al*., 1992; Bartley and Scolnik, 1993; Fraser *et al*., 1999). Our results were also consistent with those of Giorio *et al*. (2008) and Namitra *et al*. (2011) who reported that *Psy1* gene expression in the fruit was increased between the mature green and pink stages, and expressed at a higher level than that of the petal. Cloning, sequence analysis and gene expression study of *Psy2* from the leaf of *S. lycopersicum* KKU-T34003 are being carried out in our laboratory as additional data in order to gain more insights about the function of *Psy* gene family.

Similar patterns of *Pds* and *Zds* gene expression were observed. The highest expression level was found in the breaker fruit, followed by the petal, whereas the lowest expression was observed in the leaf (Figure 7A). An indirect approach to estimate the protein abundance is to examine the gene expression level. Therefore, it is plausible to explain that high levels of gene expression observed in the petal and fruit leads to the higher activity of these enzymes in converting phytoene (coloress) to ζ-carotene (pale yellow) and lycopene (red) in the tomato fruit. This observation was consistent with the high amount of carotenoid, especially lycopene accumulated in the petal and fruit tissues. Comparing the *Psy1*, *Pds* and *Zds* gene expression in the fruit, it was found that *Psy1* was expressed at the highest level, followed by *Zds* and *Pds*, correspondingly. This result was also in agreement with that of Giuliano *et al*. (1993), who reported that the expression of *Psy1* gene was much higher than that of *Pds* and *Zds* in tomato fruit at the breaker stage. Moreover, comparison of gene expression among the three in the petal showed that *Pds* was expressed at the highest level, whereas *Psy1* and *Zds* were expressed at approximately the same lower level. A similar result was also demonstrated in papaya by Yan *et al*. (2011), who reported that both *Pds* and *Zds* genes were expressed in fruit, leaf and flower, with the highest expression in mature fruit. The *Psy*1, *Pds* and *Zds* gene expression could vary in different plant tissues, which potentially affected the subsequent accumulation of carotenoids. However, endogenous and environmental factors, for example the type and the stage of the tissue development, the mutations and the ultraviolet light, could also be the causes of increased or reduced gene and protein expressions (Giuliano *et al*., 1993; Namitra *et al*., 2011; Gady *et al*., 2012; Lazzeri *et al*., 2012). The above observations indicate that the gene expression is a key of the regulatory mechanism that controls carotenoid biosynthesis *in vivo*.

4. Conclusions

In this study, we have successfully cloned *Psy1, Pds* and *Zds* cDNAs encoding phytoene synthase, phytoene desaturase and ζ-carotene desaturase from *S. lycopersicum* KKU-T34003. The trans-isoprenyl diphosphate synthase of PSY1 was found within the sequence whereas PDS and ZDS belonged to family of a dinucleotide-binding domain protein. Results from conserved domain together with the phylogenetic analyses suggested that PSY1, PDS and ZDS from *S. lycopersicum* KKU-T34003 might have similar structures and functions to the corresponding proteins from other plants and were possibly evolutionarily related. The gene expression analyses showed the highest expression of *Psy1*, *Pds* and *Zds* genes in the breaker fruit of tomato, confirming that these carotenogenic genes were expressed at high levels in the chromoplast-containing tissues. The cDNAs obtained from our study represent a high potential for being exploited in combination with other genes or cDNAs to generate the recombinant constructs, which are essentially able to produce bacterial strains or plant varieties with large amounts of lycopene in the future.

Acknowledgements

This research was financially supported by The Coordinating Center for Thai Government Science and Technology Scholarship, National Science and Technology Development Agency (NSTDA) (PDS & ZDS), Fermentation Research Center for Value Added Agricultural Products (FerVAAP), Khon Kaen University and The Khon Kaen

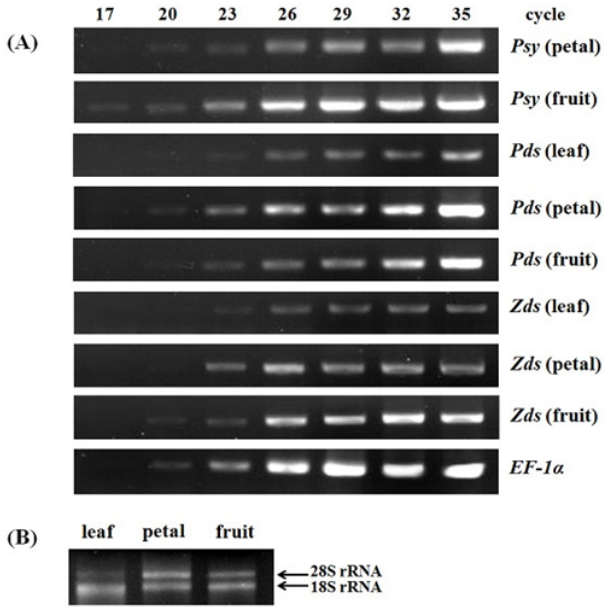

Figure 7. (A) Gene expression analysis of tomato *Psy1*, *Pds* and *Zds* genes from leaf, petal and fruit tissues carried out using semi-quantitative RT-PCR with gene-specific primers. *EF-1α* (accession number X14449) was used as an internal control (Pokalsky *et al*., 1989). (B) The extracted RNA samples from leaf, petal and fruit tissues were equalized and adjusted to be approximately 5 ng for employing as templates in the semi-quantitative RT-PCR reactions.

University's Graduate Research Fund Academic Year 2011 (PSY). We are grateful to Associate Professor Dr. Suchila Techawongstien (The Plant Breeding Research Center for Sustainable Agriculture, Faculty of Agriculture, Khon Kaen University) for providing seeds and tissues of the tomato (*S. lycopersicum* KKU-T34003), and also grateful to the Department of Biotechnology, Faculty of Technology, Khon Kaen University for supporting the equipment and facilities for this research.

References

Armstrong, G.A., Schmidt, A., Sandmann, G. and Hearst, G.E. 1990. Genetic and bio-chemical characterization of carotenoid biosynthesis mutants of *Rhodobacter capsulatus*. Journal of Biological. Chemistry. 265, 8329-8338.

Bartley, G.E., Schmidhauser, T.J., Yanofsky, C. and Scolnik, P.A. 1990. Carotenoid desaturases from *Rhodobacter capsulatus* and *Neurospora crassa* are structurally and functionally conserved and contain domains homologous to flavoprotein disulfide oxidoreductases. Journal of Biological Chemistry. 265, 16020-16024.

Bartley, G.E. and Scolnik, P.A. 1993. cDNA cloning, expression during development, and genome mapping of PSY2, a second tomato gene encoding phytoene synthase. Journal of Biological Chemistry. 268 (34), 25718-25721.

Bartley, G.E., Viitanen, P.V., Pecker, I., Chamovitz, D., Hirschberg, J. and Scolnik, P.A. 1991. Molecular cloning and expression in photosynthetic bacteria of a soybean cDNA coding for phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway. Proceedings of the National Academy of Sciences of the United States of America. 88, 6532-6536.

Bendich, A. 1994. Recent advances in clinical research involving carotenoids. Pure and Applied Chemistry. 66 (5), 1017-1024.

Bramley, P., Teulieres, C., Blain, I., Bird, C. and Schuch, W. 1992. Biochemical characterization of transgenic tomato in which carotenoid biosynthesis has been inhibited through the expression of antisense RNA to pTOM5. Plant Journal. 2, 343-349.

Burkhardt, P.K., Beyer, P., Wunn, J., Kloti, A., Armstrong, G.A., Schledz, M., von Lintig, J. and Potrykus, I. 1997. Transgenic rice (*Oryza sativa*) endosperm expressing daffodil (*Narcissus pseudonarcissus*) phytoene synthase accumulates phytoene, a key intermediate of provitamin A biosynthesis. Plant Journal. 11, 1071-1078.

Busch, M., Seuter, A. and Hain, R. 2002. Functional analysis of the early steps of carotenoid biosynthesis in tobacco. Plant Physiology. 128, 439-453.

Clinton, S.K. 1998. Lycopene: Chemistry, Biology, and Implications for Human Health and Disease. Nutrition Reviews. 56 (2), 35-51.

Cunningham, F.X. Jr. 2002. Regulation of carotenoid synthesis and accumulation in plants. Pure and Applied Chemistry. 74 (8), 1409-1417.

Dailey, T.A. and Dailey, H.A. 1998. Identification of an FAD superfamily containing protoporphyrinogen oxidases, monoamine oxidases and phytoene desaturase: expression and characterization of phytoene desaturase of *Myxococcus xanthus*. Journal of Biological Chemistry. 273, 13658-13662.

Emanuelsson, O., Nielsen, H., Von Heijne, G. 1999. ChloroP, a neural networkbased method for predicting chloroplast transit peptides and their cleavage sites. Protein Science. 8, 978-984.

Enfissi, M.A., Fraser, P.D. and Bramley, P.M. 2006. Genetic engineering of carotenoid formation in tomato. Phytochemistry Reviews. 5, 59-65.

Fraser, P.D., Kiano, J.W., Truesdale, M.R., Schuch, W. and Bramley, P.M. 1999. Phytoene synthase-2 enzyme activity in tomato does not contribute to carotenoid biosynthesis in ripening fruit. Plant Molecular Biology. 40, 687-698.

Gady, A.L.F., Vriezen, W.H., Van de Wa, M.H.B.J., Huang, P., Bovy, A.G., Visser, R.G.F. *et al.* 2012. Induced point mutations in the phytoene synthase 1 gene cause differences in carotenoid content during tomato fruit ripening. Molecular Breeding. 29, 801-812.

Giorio, G., Stigliani, A.L. and D2 Ambrosio, C. 2008. Phytoene synthase genes in tomato (*Solanum lycopersicum* L.)-new data on the structures, the deduced amino acid sequences and the expression patterns. Federation of European Biochemical Societies Journal. 275 (3), 527-535.

Giuliano, G., Bartley, G.E. and Scolnik, P.A. 1993. Regulation of carotenoid biosynthesis during tomato development. Plant Cell. 5 (4), 379-387.

Havaux, M. 1998. Carotenoids as membrane stabilizers in chloroplasts. Trends in Plant Science. 3, 147-151.

Hirschberg, J. 2001. Carotenoid biosynthesis in flowering plants. Current Opinion in Plant Biology. 4, 210-218.

Hugueney, P., Römer, S., Kuntz, M. and Camara, B. 1992. Characterization and molecular cloning of a flavoprotein catalyzing the synthesis of phytofluene and zeta-carotene in Capsicum chromoplasts. European Journal of Biochemistry. 209(1), 399-407.

Ji, J., Wang, G., Wang, J. and Wang, P. 2009. Functional analysis of multiple carotenogenic genes from *Lycium barbarum* and *Gentiana lutea* L. for their effects on beta-carotene production in transgenic tobacco. Biotechnology Letters. 31 (2), 305-312.

Just, B.J., Santos, C.A., Fonseca, M.E., Boiteux, L.S., Oloizia, B.B. and Simon, P.W. 2007. Carotenoid biosynthesis structural genes in carrot (*Daucus carota*): isolation, sequence-characterization, single nucleotide polymorphism (SNP) markers and genome mapping. Theoretical and Applied Genetics. 114, 693–704.

Karvouni, Z., John, I., Taylor, J.E., Watson, C.F., Turner, A.J. and Grierson, D. 1995. Isolation and characterization of a melon cDNA clone encoding phytoene synthase. Plant Molecular Biology. 27 (6), 1153-1162.

Krinsky, N.I. 1989. Antioxidant function of carotenoids. Free Radical Biology and Medicine. 7, 617-635.

Lazzeri, V., Calvenzani, V., Petroni, K., Tonelli, C., Castagna, A. and Ranieri, A. 2012. Carotenoid profiling and biosynthetic gene expression in flesh and peel of wild-type and *hp-1* tomato fruit under UV-B depletion. Journal of Agricultural and Food Chemistry. 60, 4960-4969.

Lee, C., McCoon, P. and LeBowitz, J. 1981. Vitamin A value of sweet corn. Journal of Agricultural and Food Chemistry. 29, 1294-1295.

Li, M., Gan Z., Cui, Y., Shi, C and Shi, X. 2011. Cloning and Characterization of the ζ-Carotene Desaturase Gene from *Chlorella protothecoides* CS-41. Journal of Biomedicine and Biotechnology., 7 pages, doi:10.1155/2011/731542.

Lindgren, L.O., Stallberg, K.G. and Höglund, A.S. 2003. Seed-specific overexpression of an endogenous Arabidopsis phytoene synthase gene results in delayed germination and increased levels of carotenoids, chlorophyll and abscisic acid. Plant Physiology. 132, 779-785.

Mann, V., Pecker, I. and Hirschberg, J. 1994. Cloning and characterization of the gene for phytoene desaturase (Pds) from tomato (*Lycopersicon esculentum*). Plant Molecular Biology. 24(3), 429-34.

Matthews, P.D., Luo, R. and Wurtzel, E.T. 2003. Maize phytoene desaturase and ζ-carotene desaturase catalyse a poly-Z desaturation pathway: implications for genetic engineering of carotenoid content among cereal crops. Journal of Experimental Botany. 54 (391), 2215-2230.

Misawa, N., Nakagawa, M, Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K and Harashima, K. 1990. Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. Journal of Bacteriology. 172 (12), 6704-6712.

Moehs, C.P., Tian, L., Osteryoung, K.W. and DellaPenna, D. 2001. Analysis of carotenoids biosynthetic gene expression during marigold petal development. Plant Molecular Biology. 45, 281-293.

Naik, P.S., Chanemougasoundharam, A., Khurana, S.M. and Kalloo, G. 2003. Genetic manipulation of carotenoid pathway in higher plants. Current Science. 85 (10), 1423-1430.

Namitra, K.K., Archana, S.N. and Negi, P.S. 2011. Expression of carotenoid biosynthetic pathway genes and changes in carotenoids during ripening in tomato (*Lycopersicon esculentum*). Food and Function. 2 (3-4), 168-173.

Paine, J.A., Shipton, C.A., Chaggar, S., Howells, R.M., Kennedy, M.J., Vernon, G., Wright, S.Y., Hinchliffe, E. and Drake, R. 2005. Improving the nutrional value of golden rice through increased pro-vitamin A content. Nature Biotechnology. 23, 482-487.

Perry, K.L., Simonitch, T.A., Harrison-Lavole, K. J. and Liu, S. 1986. Cloning and regulation of *Erwinia herbicola* pigment genes. Journal of Bacteriology. 168, 607-612.

Pokalsky, A.R., Hiatt, W.R., Ridge, N., Rasmassen, R., Houck, C.M. and Shewmaker, C.K. 1989. Structure and expression of elongation factor 1α in tomato. Nucleic Acid Research. 17 (12), 4661-4673.

Rao, A.V. and Agarwal, S. 1999. Role of lycopene as antioxidant carotenoid in the prevention of chronic diseases: a review. Nutrition Research. 19, 305-323.

Romer, S., Hugueney, P., Bouvier, F., Camara, B. and Kuntz, M. 1993. Expression of the genes encoding the early carotenoid biosynthetic enzymes in *Capsicum annuum*. Biochemical Biophysical Research Communications. 196 (3), 1414-1421.

Sambrook, J. and Russell, D.W. 2001. Molecular Cloning: A Laboratory Manual. Third Edition. New York: Cold Spring Harbor Laboratory Press.

Sandmann, G. 1994. Phytoene desaturase: genes, enzymes and phylogenetic aspects. J. Plant Physiology. 143, 444-447.

Schledz, M., al-Babili, S., von Lintig, J., Haubruck, H., Rabbani, S., Kleinig, H. *et al.* 1996. Phytoene synthase from *Narcissus pseudonarcissus*: functional expression, galactolipid requirement, topological distribution in chromoplasts and induction during flowering. Plant Journal. 10 (5), 781-792.

Scolnik, P.A. and Bartley, G.E. 1994. Nucleotide sequence of an *Arabidopsis* cDNA for phytoene synthase. Plant Physiology. 104 (4), 1471-1472.

Shewmaker, C.K., Sheehy, J.A., Daley, M., Colburn, S. and Ke, D.Y. 1999. Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. Plant Journal. 20 (4), 401-412.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0, Molecular Biology and Evolution. 24 (8), 1596-1599.

Yan, P., Gao, X.Z., Shen, W.T. and Zhou, P. 2011. Cloning and expression analysis of phytoene desaturase and ζ-carotene desaturase genes in *Carica papaya*. Molecular Biology Reports. 38, 785-791.

Young, A. and Britton, G. 1993. Carotenoids in photosynthesis. London: Chapman & Hall.

Zhang, J. 2003. Evolution of gene duplication: an update. Trends in Ecology and Evolution. 18, 292-298.