

## RT-PCR Amplification and Cloning of Partial DNA Sequence Coding for Oil Palm (*Elaeis oleifera*) Phytoene Synthase Gene

O A Rasid<sup>1\*</sup>, W S Wan Nur Syuhada<sup>2</sup>, A Nor Hanin<sup>1</sup>, S S Masura<sup>1</sup>, M Zulqarnain<sup>2</sup>,  
C L Ho<sup>3</sup>, R Sambanthamurthi<sup>1</sup> and N Suhaimi<sup>3</sup>

<sup>1</sup>Malaysian Palm Oil Board, No 6 Persiaran Institusi, Bandar Baru Bangi 43000, Kajang, Selangor, Malaysia

<sup>2</sup>Department of Genetics and Molecular of Biology, Institute of Biological Sciences,  
University of Malaya, 50603 Kuala Lumpur, Malaysia

<sup>3</sup>Faculty of Biotechnology and Biomolecular Sciences, 43400 UPM Serdang, Selangor, Malaysia.

Received 18 June 2007 / Accepted 10 December 2007

**Abstract.** The potential health benefits of carotenoids as anti cancer and antioxidant agents have recently been demonstrated. Oil palm, *Elaeis oleifera* in particular, is known to be the richest natural source for carotene. However, the species has not been commercially exploited due to its extremely low oil yield. The current work describes the isolation of a cDNA clone coding for phytoene synthase (*psy*) from *E. oleifera* by RT-PCR amplification. A pair of *psy* gene specific primers was successfully used to amplify a 899 bp fragment that codes for a partial length (300 amino acids) of oil palm *psy*. The DNA and amino acid sequences were shown to share a high level of identity to phytoene synthase from other plants at about 83%. Further analysis also showed the presence of conserved aspartate-rich catalytic domains within the clone. Work was also carried out to obtain the expression pattern of oil palm *psy* in developing fruits by real-time PCR analysis. Results indicated that the gene is highly regulated during the course of oil palm fruit development. The pattern of *psy* expression was shown to be well correlated to the accumulation of lutein in the young mesocarp and  $\alpha$ - and  $\beta$ -carotenes in the older tissues. This observation demonstrated that oil palm *psy* was highly regulated for tissue development and accumulation of carotenes for storage.

**Keywords.** Carotenoids, cDNA clone, *Elaeis oleifera*, Expression, Phytoene synthase

### INTRODUCTION

Carotenoids are a large class of isoprenoid-derived pigments that are synthesised *de novo* by many organisms. In plants, carotenoids are important for a number of biological functions such as photosynthesis and photoprotection. In human, carotenoids, in particular  $\beta$ -carotene, are an important source of provitamin A. Deficiency in vitamin A can cause a number of health complications including cataract, xerophthalmia and decreased resistance to diseases (Miller *et al.*, 2002; Canfield, 1995). Another potential health benefit of carotenoids has been reported which is related to their anticancer and antioxidant properties (Handelmann, 2001). It has been suggested that certain carotenoids such as lycopene, are effective anticancer agents for certain types of cancers, such as oral, throat, stomach, and colon cancers (Suda *et al.*, 1986; Mathews-Roth and Krinsky, 1987; Sundram, 1989).

In plants, carotenoids are synthesized and accumulated in plastids. The first committed step in carotenoid synthesis is the formation of the first C40 compound phytoene by con-

densation of two molecules of geranylgeranyl diphosphate (GGDP). This reaction is catalysed by phytoene synthase (*psy*). cDNAs encoding the enzyme have been isolated and characterized from various plants (review in Fraser and Bramley, 2004) including *Arabidopsis*, maize, melon, carrot, *Citrus paradisi* and *unshiu*, daffodil, rice, pepper, marigold and soybean. Moreover, plant phytoene synthase has also been shown to be highly regulated. For example, *psy1*, a tomato fruit specific phytoene synthase, was shown to be significantly increased at the breaker stage (Bartley and Scolnik, 1993; Bartley *et al.*, 1992). The increase in *psy1* expression was followed by an increase in total carotenoid content in the fruits, mainly due to the accumulation of lycopene and  $\beta$ -carotene. It has also been suggested that phytoene synthase could substantially regulate the reaction flux of carotenoid synthesis.

\*Author for Correspondence.

Mailing address: Malaysian Palm Oil Board, No 6 Persiaran Institusi, Bandar Baru Bangi 43000, Kajang, Selangor, Malaysia. Tel: 603-87694585. Fax: 603-89261995. Email: omar@mpob.gov.my

Thus, the gene has been considered to be one of the primary targets for genetic modification of plant carotenoids (Frazer *et al.*, 2002; Shewmaker *et al.*, 1999).

Attempts to genetically modify the carotenoid content in transgenic plants have produced mixed results. Recently, the development of genetically modified Golden Rice 2 has been reported (Paine *et al.*, 2005). The carotene content in the plant has been much improved compared to the first generation of Golden Rice (Ye *et al.*, 2000). The Golden Rice 1 was developed by introducing the daffodil *psy* together with *Erwinia uredovora* carotene desaturase (*crtl*). However, this enzyme was suggested to be the limiting step in  $\beta$ -carotene accumulation. Subsequently, Golden Rice 2 was developed by utilizing *psy* from maize instead of daffodil. The use of maize *psy* has substantially increased carotenoid accumulation in the transgenic plants. The results indicated that the use of different *psy* source with potentially different activity or effectiveness could result in a better accumulation of carotene in the transgenic plants.

Oil palm is known to be a rich natural source for carotenoids. The commercial materials were shown to contain about 500-700 ppm carotenoids (Choo *et al.*, 1997). The major components are  $\alpha$ -carotene and  $\beta$ -carotene, which make up about 90% of the total content. More interestingly, *E. oleifera*, a related species of *E. guineensis*, has been shown to contain more carotenoids (up to 4000 ppm) compared to *E. guineensis* (Mohd Din *et al.*, 2004; Choo *et al.*, 1997). However, this useful trait could not be introgressed into the commercial planting materials due to its extremely low oil yield. Nevertheless, the species could be potentially exploited for increasing or modifying the carotenoid content in oil palm.

However, details concerning carotenoid synthesis in oil palm are very limited due to the lack of studies to elucidate the pathways (Rasid *et al.*, 2003). To date, partial or full length cDNA clones encoding five of the oil palm carotenoid genes have been reported, namely, phytoene desaturase (Rasid *et al.*, 2005), lycopene  $\beta$ -cyclase (Rasid *et al.*, 2003) 1-deoxy-D-xylulose 5-phosphate synthase (Khemvong and Suvachittanont, GenBank Accession no. AY583783 and AY611205), 1-deoxy-D-xylulose 5-phosphate reductoisomerase (Khemvong and Suvachittanont, GenBank Accession no. AY583783 and AY611205), zeaxanthin epoxidase (Rasid *et al.*, 2005), and full length of lycopene  $\epsilon$ -cyclase and lycopene  $\beta$ -cyclase (Rasid *et al.*, unpublished data).

The current work was aimed at cloning the cDNA clone that codes for the phytoene synthase from *E. oleifera*. Here, we report the isolation and characterization of a cDNA clone encoding the partial length of oil palm phytoene synthase gene. Our results indicate that the gene is highly conserved in the oil palm. Our results also indicate that oil palm *psy* is highly regulated during tissue development and carotenoid accumulation for storage. To our knowledge, this is the first report on the isolation and characterization of oil palm *psy*.

## MATERIALS AND METHODS

**Isolation of Total RNA.** Total RNA was isolated from mesocarp tissues of oil palm fruit and leaves according to Hosen (2001) with modification. cDNAs generated from total RNA from 17 WAA (week after anthesis) of the mesocarp tissues of *E. oleifera* were used in RT-PCR (reverse transcriptase PCR). The cDNA synthesis was carried out using SMART RACE kit (BD Biosciences) according to the manufacturer's protocol.

**Primer Design.** PCR primers for RT-PCR amplification of oil palm *phytoene synthase* were designed by using CConsensus-DEgenerate Hybrid Oligonucleotide Primer (CODEHOP) (Rose *et al.*, 1998). The software is available at <http://blocks.fhrcc.org/codehop.html>. The primers were synthesized based on the conserved regions of the gene. These conserved regions were predetermined using alignment of *psy* sequences from other plants that are available at the GenBank including *Adonis palaestina* (Accession # AAV74394), *Citrus unshiu* (Accession # AAF33237), *Cucumis melo* (Accession # Z37543), *Narcissus pseudonarcissus* (Accession # CAA55391), *Oryza sativa* (Accession # BAD62106), *Prunus armeniaca* (Accession # AAX33349), *Zea mays* (Accession # AAR08445). The software required identification of conserved regions within the gene sequences to design hybrid primers consisting of a short 3'degenerate core region (for selectivity) and a longer 5'consensus clamp region (for specificity). The forward primer PSY5, 5'-GGTGTACAACGTGGT-GCTGAARCARGCNGC-3' and the reverse primer PSY6, 5'-TGAAGTTGTTGTAGTCGTTGGCYTCDATYTC-3' were derived from the amino acid sequences VYNNVVLKQA and FNNYDNAEIE, respectively. The relative position of the primers to the coding region of a known plant *psy* and expected fragment size are shown in Figure 1.

**RT-PCR and Cloning.** RT-PCR amplifications were carried out using Accuprime™ *Taq* DNA Polymerase High Fidelity (Invitrogen) in 35 sequential cycles at 94°C for 30 sec, 53°C for 30 sec and 72°C for 2 min 30 sec on PTC-200 Programmable Thermal Controller (MJ Research, Inc.). The amplifications were performed using combinations of PSY5/PSY6, PSY5/UPM (universal primer mixture provided in the SMART RACE kit) or PSY6/UPM. The PCR products were separated using agarose gel electrophoresis. Amplified fragments were excised from the agarose gel, purified using QIAEXII Gel Extraction Kit (QIAGEN) and finally ligated into PCRII-TOPO vector (Invitrogen). Plasmid DNA from several selected colonies was isolated using GeneTACG Mini s-Preo Plasmid DNA Extraction Kit (GeneTACG Biosciences) according to the manufacturer's protocol. Confirmation of insert was performed with *Eco*R I digestion.

**DNA Sequence Analysis.** Plasmid DNA for sequencing was prepared using QIAGEN miniprep Kit (QIAGEN). DNA



**Figure 1.** The position of the PSY primers relative to the coding region of *psy* from *Oryza sativa* (REF). The *O. sativa* coding region (1263 bp) is represented by double headed arrow. Small arrows represent the primers. The numbers denote the start and end position of the primers relative to the *O. sativa psy*.

sequencing was performed on an ABI Prism 377 automated DNA sequencer. DNA sequence analysis was carried out using Vector NTI (Invitrogen) and DNA or protein homology search was performed using BLAST 2.0 (Altschul *et al.*, 1997) to the GenBank database.

**Real-time PCR.** The expression study of oil palm *psy* was performed using real-time PCR. Total RNA from mesocarp of 5, 7, 9, 11, 13, 15, 17 and 19 WAA fruits was used in the analysis. Total RNA from kernel, spear and green leaves was also included for comparison. Removal of DNA contamination in total RNA samples was carried out using the QIA-GEN RNase-free DNase Set according to the manufacturer's protocol. The concentration of the RNA samples was determined using the NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies Inc.). The RNA samples were subjected to integrity analysis using Agilent 2100 Bioanalyzer (RNA 6000 Nano Assay Kit) (Agilent Technologies). Intact RNA was converted to cDNA using High Capacity cDNA Archive Kit (Applied Biosystems).

Analysis for conserved regions that were used for primer and probe synthesis was performed using VectorNTI (Invitrogen). Real-time PCR (TaqMan assay) was carried out in mixtures containing 1X TaqMan Universal PCR Master Mix, 1X Assay Mix (containing specific primers and probe) and 45 ng of template cDNA. The internal probe was labeled at the 5' end with dye 6-carboxyfluorescein (FAM) whereas the 3' end was labeled with the non-fluorescent quencher (NFQ). PCR cycling parameters were set at one cycle at 50°C for two min, 95°C for 10 min, and 40 cycles at 95°C for 15 s and 60°C for one min. Real-time detection of fluorescence was performed on the ABI PRISM 7000 Sequence Detection System (Applied Biosystems).

The efficiency of real-time PCR TaqMan reactions was determined by performing real-time PCR reaction using 5 serial 10-fold dilutions of a template comprising of 100, 10, 1, 0.1 and 0.01 ng cDNA. The  $C_T$  values were plotted against the  $\log_{10}$  concentration of template. The slope of the resulting was used to calculate the PCR efficiency using the following equation: PCR efficiency (%) =  $(10^{-1/\text{slope}} - 1) \times 100$  (Toplak *et al.*, 2004). A standard curve slope of -3.32 indicates a PCR reaction with 100% efficiency.

## RESULTS AND DISCUSSION

### RT-PCR and Cloning of Oil Palm Phytoene Synthase.

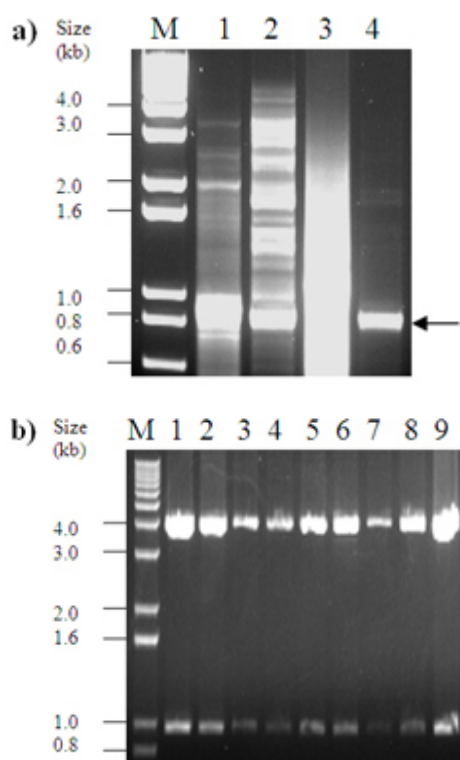
The primers used in the amplification of oil palm *psy* gene were designed based on the conserved regions of the genes. This was primarily due to the lack of information available on oil palm phytoene synthase. The conserved regions were identified by DNA sequence alignment. The primers were then designed using the CODEHOP software. The consensus-degenerate hybrid primers conceived with the CODEHOP programme were designed to overcome problems like non-specific priming and low stringency annealing conditions during PCR amplification (Rose *et al.*, 1998). The primers have been shown to be successful for the amplification of gene orthologs (Wilson *et al.*, 1998, and Baines *et al.*, 2005).

In this work, one of the primer combinations (PSY5/PSY6) was successfully optimised to produce a single amplified fragment of about 0.9kb (Figure 2a). Based on the positions of the primers relative to the known sequences for *psy* from *Oryza sativa* (Figure 1), this fragment is of the expected size. Thus, the amplified fragment was subsequently purified and cloned into the PCR II TOPO vector for further analysis. The presence of the DNA insert in selected colonies was verified by *EcoR* I digestion (Figure 2b). The clone was designated as pEPSY30.

**DNA Sequence Analysis.** The complete DNA sequence was obtained for representative clones originally from RT-PCR products. The sequence analysis result showed that the fragment size was 899 bp. Results from BLAST search indicated that the clone DNA sequence was highly identical to *psy* sequence from other plants at about 81%- 84% identity.

Based on comparison to amino acid sequence of PSY from other plants, a possible deduced amino acid sequence for pEPSY30 was obtained. It was shown that the clone codes for 300 amino acid residues. The identity of the deduced amino acid sequence was further verified by BLAST search to the GenBank database. The results showed that the sequence shared a very high level of identity to PSY from other plants including *Daucus carota* (accession number ABB52068), *Adonis palaestina* (accession number AVV74394.1), *Narcissus pseudonarcissus* (accession number CAA55391.1) and *Momordica charantia* with about 83%, 82%, 81% and 81% identities, respectively (Figure 3).

The search for conserved domains that are possibly present in pEPSY30 was performed using CDD software available at <http://www.ncbi.nlm.nih.gov/BLAST> (Marchler-Bauer *et al.*, 2003). The results showed the presence of a highly conserved domain known as Trans\_IPPS\_HH (Trans-Isoprenyl Disphosphate Synthases (Trans\_IPPS), head-to-head (HH) (1'-1) condensation reaction) domain. This catalytic site consists of a large central cavity formed by antiparallel alpha helices with two aspartate-rich regions (DXXXD) (Ohnuma *et al.*, 1996; Ohnuma *et al.*, 1994) located on opposite walls (Figure 3). The X denotes any amino acid. This domain was shown



**Figure 2.** (a) RT-PCR amplification of oil palm phytoene synthase (*psy*) gene using degenerate primers. A combination of PSY5/PSY6 primers was shown to successfully produce a single amplified fragment of the expected size. Lane 1: PSY5/UPM, template: 5'strand; lane 2: PSY5/PSY6, template: 5'strand; lane 3: PSY6/UPM, template: 3'strand; lane 4: PSY5/PSY6, template: 3'strand. Arrow indicates the 899 bp amplified product. (b) Digestion and electrophoresis of representative clones containing RT-PCR product for oil palm *psy*. Digestion was performed using *Eco*R I. All clones were shown to carry the insert (Lane 1-9). M is 1 kb Plus DNA ladder (Invitrogen).

to be highly conserved across plant and animal species. The oil palm domain was about 94%, 91%, 88% 85%, and 26% identical to Trans IPPS domain of phytoene synthase from *Oryza sativa*, *Zea mays*, *Synechocystis sp.*, green alga, and Chain A of human squalene synthase respectively. This conserved domain is present in squalene and phytoene synthases. The former catalyses the 1'-1 condensation reaction of two farnesyl (15C) diphosphates.

**Regulation of Oil Palm Phytoene Synthase.** The attempt to study the expression of these genes was carried out using real-time PCR. The technique has been reported to be sensitive and is able to detect low abundant transcripts. Therefore, RNA samples from a number of oil palm tissues, namely mesocarp 5, 7, 9, 11, 13, 15, 17 and 19 (WAA), kernel 10 WAA, spear leaves and green matured leaves were used in the real-time PCR analysis. The primers and the probes used in the

**Table 1.** List of primers and TaqMan probes specific for oil palm phytoene synthase and Glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

Genes	Primer Sequence
Phytoene synthase	<b>PSY For</b> 5'-GGGCAATATATGTTTGGTGCAGAAG-3'
	<b>PSY Rev</b> 5'-CATAGCCGAAGGCGTAATGTGA-3'
	<b>PSY probe</b> 5'-FAM-AAGCATTTGGGTCCATCTACGA-NFQ
	<b>GAPDH For</b> 5'-ACTGCTACTCAGAAGACTGTTGATG-3'
Glyceraldehyde-3-phosphate dehydrogenase	<b>GAPDH Rev</b> 5'-TGCTGCTAGGAATGATGTTAAAGCT-3'
	<b>GAPDH Probe</b> 5'-FAM-ACCCCTCCAGTCCCTTG-NFQ-3'

real-time PCR were designed from highly conserved regions of the genes. These regions were identified by VectorNTI sequence alignment and BLAST Search at the GenBank. Sequence of primers and probes are shown in Table 1.

The relative quantification of the genes was carried out using the comparative  $\Delta\Delta C_T$  method (Livak and Schmittgen, 2001). The real-time PCR for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene was also included in the analysis as the internal control. Since the analysis is comparative, it is important to ensure that an equal amount of template is used. For this reason, it is also important to accurately estimate the concentration of the RNA samples to be used in the real-time PCR analysis. In this study, the concentration and quality of these samples were determined by using Nano-Drop® ND-1000 Spectrophotometer. The technique could accurately estimate the concentration of the RNA sample as low as 2 ng  $\mu\text{L}^{-1}$ . Finally, the integrity of RNA samples was also determined using Agilent 2100 Bioanalyzer.

The amplification efficiency of the TaqMan assay is shown in Figure 4. The quantitation was linear over a range of 5 log units of *psy* gene indicating a wide dynamic range and high reliability. It was shown that the PCR reaction for *psy* was 100% efficient. The threshold cycle ( $C_T$ ) values obtained from the amplification plot were used to estimate the relative abundance of the genes. The value marks the first cycle at which the signal is significantly above the background. Then, the average  $C_T$  values for the target gene were normalized to the average  $C_T$  values of the internal control (GAPDH). The difference in  $C_T$  values ( $\Delta C_T$ ) is equal to the difference in expression after being normalized to the endogenous control. Finally, the comparative  $C_T$  values ( $\Delta\Delta C_T$ ) were obtained by subtracting the  $\Delta C_T$  of calibrator sample (K10) from  $\Delta C_T$  of test sample. The calibrator is normally an untreated sample. However for expression studies, the calibrator can be any one of the samples used but normally the sample with the lowest

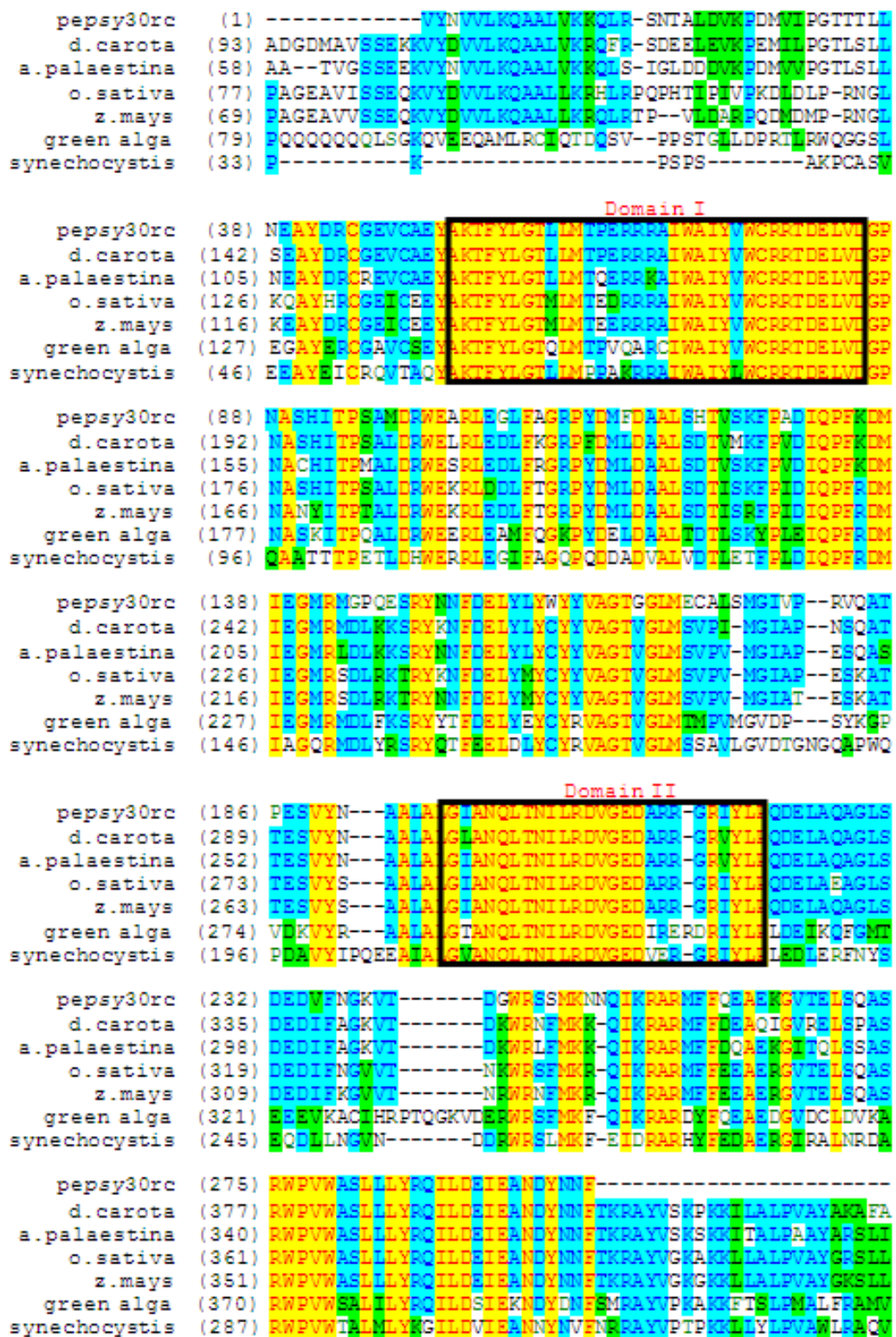
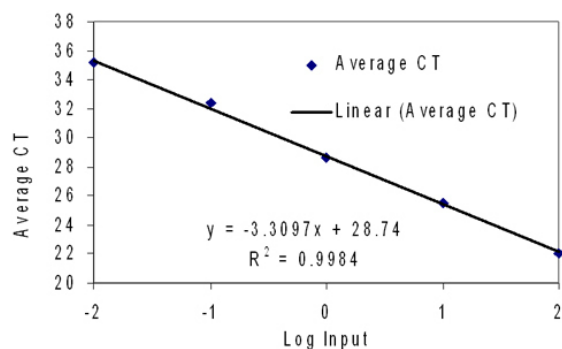
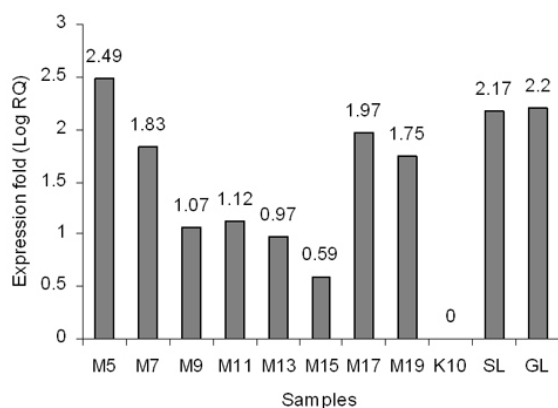


Figure 3. The conserved core motifs amino acid domain. This alignment shows that protein sequences are extremely polymorphic. The protein sequences were aligned using Align X as implemented in the VNTI software. Strictly conserved sequences are in yellow. Identification of the aspartate-rich domains I and II of different species (shown in boxes). Amino acid sequences identical to those pEPSY30 are shown in yellow regions. *Synechocystis sp* (accession number: P37294); phytoene synthase, chloroplast precursor *Zea mays* (accession number: P49085); *Oryza sativa* (japonica cultivar-group) (accession number: AAK07735); green alga, *Dunaliella bardanil* (accession number: T10702); Chain A, Crystal structure of Human squelene (accession number: 1EZFA).



**Figure 4.** Real-time PCR amplification efficiency curve generated from 10-fold serial dilutions of cDNA. The average  $C_T$  value is plotted against the log of input template amount. The  $C_T$  value decreased log-linearly with increasing amount of template.



**Figure 5.** The expression level of oil palm *psy* gene in developing fruits (M5-M19) obtained by real-time PCR analysis. High level of *psy* expression in the early and late developmental stages of the fruit. Similar level of expression was also observed in spear (SL) and green (GL) leaves. K10 was referred as the calibrator.

$C_T$  value. These  $\Delta\Delta C_T$  values represent the difference in the expression after being normalized to internal control and relative to a calibrator. The relative expression (RQ) of the target genes was calculated using the  $\Delta\Delta C_T$  values according to the equation  $2^{-\Delta\Delta C_T}$ . Fold-differences were expressed as log values of RQ values.

The results from the real-time PCR relative quantitation analysis are summarized in Table 2. As indicated in the table, the standard deviation (SD) for average  $C_T$  value for all the target genes and the internal control was very small. This indicated the high degree of accuracy between replicates. The summarized data also indicates that the GAPDH transcript is abundant in all the samples used, as signified by the  $C_T$  values. In contrast, the transcript level of the oil palm *psy* seemed to be very much less compared to the transcript level of GAPDH. The  $C_T$  value for *psy* was 24 to 35, as compared to 20.8 to 23.1 for GAPDH. This observation clearly suggests that *psy* is expressed at a low levels in oil palm mesocarp tissues.

Overall, *psy* expression was shown to be relatively high in the young mesocarp tissue (5 WAA). It was about 2.5 fold higher than in kernel (calibrator). However, the expression progressively decreased to about 1.83 fold at 7<sup>th</sup> week and 1.07 fold at 9<sup>th</sup> week. Then, the expression remained constant at about one fold from the 9<sup>th</sup> to the 13<sup>th</sup> week. The expression further decreased to the lowest level in the 15<sup>th</sup> week. The expression then was increased to about two fold in 17 and 19 WAA tissues (Figure 5).

The expression of oil palm *psy* gene seemed to be well correlated to the accumulation of carotenoids in fruit tissues. Tay and Gwendoline (2006) have shown the presence of a relatively high amount of lutein in young mesocarp (one WAA) tissues. The content of this carotenoid remains high during mesocarp development until 8 WAA. After 8 WAA, the content of this carotenoid starts to decline. Thus, it

**Table 2.** The summarized results of Real-Time RT-PCR analysis for oil palm *psy* in mesocarp tissues of different developmental stages (M5-M19 WAA), kernel and spear leaves. The relative quantitation was obtained using the comparative ( $\Delta\Delta C_T$ ) method. The RT-PCR analysis for GAPDH gene was used as the internal control. The standard error for the average  $C_T$  for target genes and internal control is given in parenthesis. K10 was used as the calibrator.  $C_T$ , threshold cycle and RQ, relative quantitation.

Sample	<i>psy</i> Average $C_T$	GAPDH Average $C_T$	$\Delta C_T$	$\Delta\Delta C_T$	Relative Quantification (RQ)	Expression Fold (log RQ)
M5	23.333 (0.115)	23.84 (0.027)	-0.507 (0.118)	-8.286	312.129	2.49
M7	23.296 (0.043)	21.612 (0.021)	1.684 (0.048)	-6.095	68.356	1.83
M9	28.065 (0.027)	23.843 (0.074)	4.222 (0.079)	-3.557	11.77	1.07
M11	29.959 (0.055)	25.904 (0.024)	4.055 (0.06)	-3.724	13.214	1.12
M13	31.927 (0.108)	27.371 (0.061)	4.556 (0.124)	-3.223	9.337	0.97
M15	33.046 (0.385)	27.242 (0.022)	5.804 (0.386)	-1.975	3.93	0.59
M17	23.706 (0.217)	22.472 (0.021)	1.234 (0.218)	-6.545	93.38	1.97
M19	28.102 (0.012)	26.12 (0.064)	1.982 (0.065)	-5.797	55.60	1.75
K10	35.46(0.194)	27.681(0.038)	7.779(0.198)	0	1	0
SL	23.329 (0.05)	22.759 (0.049)	0.570 (0.07)	-7.208	147.85	2.17
GL	27.033 (0.118)	26.57 (0.093)	0.463 (0.15)	-7.316	159.344	2.20

should be expected that the genes required for the formation of this carotenoid must be present at a relatively high amounts at these stages. In this work, it has been shown that *psy*, the first enzyme for lutein formation, was expressed at a relatively high level in five through seven WAA mesocarp tissues. The high level of *psy* expression during these stages corresponds to the high level of lutein observed. The high content of lutein is essential for plant development as this carotenoid is required for photosynthetic apparatus. The expression of oil palm *psy* was also correlated with the accumulation of  $\alpha$ - and  $\beta$ -carotenes in oil palm fruits. The increase of *psy* expression at the 17<sup>th</sup> and 19<sup>th</sup> weeks of mesocarp development is in good agreement with the findings that  $\alpha$ - and  $\beta$ -carotenes were shown to significantly increase at 12 WAA and reach maximum levels at 18 WAA fruit (Tay and Gwendoline, 2006; Khemvong and Suvachittanont, 2005). This finding clearly suggested that oil palm *psy* is highly regulated in the developing fruits. The expression is tightly correlated to the requirement for development in the young mesocarp tissues and for the carotenoid accumulation for storage in the older mesocarp tissues.

The hypothesis was further evidenced by the high level of *psy* expression in young and green leaves. The expression of *psy* in these tissues was only slightly lower than in the 5 WAA mesocarp tissues. The high expression of this gene in these tissues could reflect the large requirement for carotenoids including lutein and other xanthophylls in photosynthesis, while, in young leaf tissues, carotenoids are also needed for tissue development.

## CONCLUSION

This paper reports the isolation of a gene that codes for one of the genes involved in the carotenoid biosynthetic pathway of oil palm. We have isolated and completely sequenced partial cDNA clones coding for oil palm phytoene synthase. Based on the sequence comparison either at the nucleic acid or deduced amino acid level, we concluded that the clones that we obtained indeed code for oil palm phytoene synthase as evidenced by their high identity to other plant phytoene synthase genes. Conserved aspartate-rich domains that are essential for catalytic activity were also observed within the sequence. The expression study showed that the pattern of oil palm *psy* expression was well correlated to the accumulation of lutein in the young mesocarp tissues and  $\alpha$ - and  $\beta$ -carotenes in the older tissues. This finding clearly demonstrates the regulation of oil palm *psy* for tissue development and accumulation of carotenes for storage. Work is ongoing to isolate the full length of this gene and other genes that are involved in the pathway.

## ACKNOWLEDGEMENTS

The authors would like to thank the Director-General of MPOB for permission to publish this paper. We also would like to acknowledge Dr. Rajinder Singh and Dr. Parveez for their comments, the Genomics group for the DNA sequencing and Pn. Suhaila Abd. Wahab for her assistance.

## REFERENCES

- Altchul, S.F., Madden, T.L., Shaffer, A.A., Zhang, J., Zhang, G., Miller, W. and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.
- Baines, J.E., McGovern, R.M., Persing, D. and Gostout, B.S. 2005. Consensus degenerate hybrid oligonucleotide primers (CODEHOP) for the detection of novel papillomaviruses and their application to esophageal and tonsillar carcinomas. *Journal of Virological Methods* 123: 81-87.
- 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. *The Journal of Biological Chemistry* 268: 25718-25721.
- Bartley, G.E., Viitanen, P.V., Bacot, K.O. and Scolnik, P.A. 1992. A tomato gene expressed during fruit ripening encodes an enzyme of the carotenoid biosynthesis pathway. *The Journal of Biological Chemistry* 267: 5036-5039.
- Canfield, L.M. 1995.  $\beta$ -carotenoid metabolites: Potential importance to human health. *Malaysian Oil Science & Technology* 4: 43-46.
- Choo, Y.M., Ma, A.N. and Yap, S.C. 1997. Carotenes, vitamin E and sterols in oils from *Elaeis guineensis*, *Elaeis oleifera* and their hybrids. *Palm Oil Developments* 27: 1-9.
- Fraser, P.D. and Bramley, P.M. 2004. The biosynthesis and nutritional uses of carotenoids. *Progress in Lipid Research* 43: 228-265.
- Fraser, P.D., Romer, S., Shipton, C.A., Mills, P.B., Kiano, J.W., Misawa, N., Drake, R.G., Schuch, W. and Bramley, P.M. 2002. Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Plant Biology* 99: 1092-1097.
- Handelmann, G.J. 2001. The evolving role of carotenoid in human biochemistry. *Nutrition* 17: 818-822.
- Hosein, F. 2001. Isolation of high quality RNA from seeds

- and tubers of the Mexican yam bean (*Pachyrhizus erosus*). *Plant Molecular Biology Reporter* 19: 65a-65e.
- Khemvong, S. and Suvachittanont, W. 2005. Molecular cloning and expression of a cDNA encoding 1-deoxy-D-xylulose 5-phosphate synthase from oil palm *Elaeis guineensis* Jacq. *Plant Science* 169: 571-578.
- Livak, K.J. and Schmittgen, T.D. 2001. Analysis of relative gene expression data using Real-Time quantitative PCR and the 2- $\Delta\Delta C_T$  method. *Methods* 25: 402-408.
- Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C., Liebert, C.A., Liu, C., Lu, F., Marchler, G.H., Mullokandov, M., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, P.A., Yamashita, R.A., Yin, J.J., Zhang, D. and Bryant, S.H. 2005. CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Research* 33: 10.1093/nar/gki069
- Mathews-Roth, M.M. and Krinsky, N.I. 1987. Carotenoids affect development of UV-B induced skin cancer. *Photochemistry and Photobiology* 46: 507-509.
- Miller, M., Humphrey, J., Johnson, E., Marinda, E., Brookmeyer, R. and Katz, J. 2002. Why do children become vitamin A deficient?. *Proceeding of the XX international vitamin A consultative group meeting*. 2867s-2880s.
- Mohd Din, A., Kushairi, A.A., Noh, A., Isa, Z.A., Junaidah, J. and Rajanaidu, N. 2004. Innovations in oil palm breeding and genetics. Book of Abstracts of *AGRICONGRESS 2004*, Selangor, Malaysia, pp. 448-449.
- Ohnuma, S-i., Nakazawa, T., Hemmi, H., Hallberg, A-M., Koyama, T., Ogura, K. and Nishino, T. 1996. Conversion from farnesyl diphosphate synthase to geranylgeranyl diphosphate synthase by random chemical mutagenesis. *The Journal of Biological Chemistry* 271: 10087-10095.
- Ohnuma, S-i., Suzuki, M. and Nishino, T. 1994. Archaeobacterial ether-linked lipid biosynthetic gene. *The Journal of Biological Chemistry* 269: 14792-14797.
- Paine, J.A., Shipton, A., Chaggar, S., Howells, R.M., Kennedy, M.J., Vermon, G., Wright, S.Y., Hinchliffe, E., Adams, J.L., Silverstones, A.L. and Drake, R. 2005. Improving the nutritional value of golden rice through increased pro-vitamin A content. *Nature Biotechnology* 23: 482-487.
- Rasid, O.A., Singh, R., Harikrisna, K., Ho, C.L. and Cheah, S.C. 2003. Isolation of partial DNA sequence coding for lycopene  $\beta$ -cyclase from oil palm (*E. guineensis*). *Proceedings of the PIPOC A-P13*: 781-788.
- Rasid, O.A., Singh, R., Sambanthamurthi, R., and Cheah, S.C. 2005. RT-PCR amplification and cloning partial DNA sequence coding for oil palm (*E. guineensis*) phytoene desaturase. *Proceedings of the PIPOC A-P43*: 1119-1128.
- Rasid, O.A., Singh, R., Sambanthamurthi, R. and Cheah, S.C. 2005. RTPCR amplification and cloning of partial DNA sequence coding for oil palm (*E. guineensis*) zeaxanthin epoxidase. *Proceedings of the 2005 Conference on Biotechnology of Plantation Commodities*, Selangor, Malaysia, pp. 553.
- Rose, T.M., Schultz, E.R., Henikoff, J.G., Pietrokovski, S., Mccallum, C.M., Henikoff, S. 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Research* 26: 1628-1635.
- Shewmaker, C.K., Sheesy, J.A., Daley, M., Colburn, S. and Ke, D.Y. 1999. Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. *The Plant Journal* 20: 401-412.
- Suda, T., Takahashi, N., Shinki, T., Yamaguchi, A. and Tanioka, Y. 1986. The common marmoset as an animal model for vitamin D-dependent rickets, type II. *Advances in Experimental Medicine and Biology* 196: 423-435.
- Sundram, K., Khor, H.T., Ong, A.S.H. and Pathmarathan, R. 1998. Effect of dietary palm oils on mammary carcinogenesis in female rats induced by 7,12-dimethylbenz (a) anthracene. *Cancer Research* 49: 1447-1451.
- Tay, Y.P.B. and Gwendoline, E.C.L. 2006. Identification of lutein in crude palm oil and evaluation of carotenoid at various ripening stage of the oil palm fruit. *Journal of Oil Palm Research* 18: 189-197.
- Toplak, N.; Okršlar, V.; Stanič-racman, D.; Gruiden, K. and Žel, J. 2004. A high-throughput method for quantifying transgene expression in transformed plants with real-time PCR analysis. *Plant Molecular Biology Reporter* 22: 237-250.
- Wilson, C.A., Wong, S., Muller, J., Davidson, C.E., Rose, T.M. and Burd, P. 1998. Type C retrovirus released from porcine primary peripheral blood mononuclear cells infects human cells. *Journal of Virology* 72: 3082-3087.
- Ye, X., Salim, A., Klöti, A., Zhang, J., Lucca, P., Beyer, P. and Potrykus, I. 2000. Engineering the Provitamin A ( $\beta$ -Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science* 287: 303-305.