JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Cloning and Sequence Analysis of the Phytoene Synthase Gene from a Unicellular Chlorophyte, *Dunaliella salina*

YUAN YAN,† YUE-HUI ZHU,† JIAN-GUO JIANG,*,† AND DONG-LIN SONG§

College of Food and Biotechnology, South China University of Technology, Guangzhou 510640, People's Republic of China, and Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, People's Republic of China

Phytoene synthase (Psy) catalyzing the dimerization of two molecules of geranylgeranyl pyrophosphate (GGPP) to phytoene may be rate limiting for the synthesis of β-carotene in *Dunaliella salina*. To elucidate the carotenogenic pathway in *D. salina*, the complete Psy gene was first isolated by genome walking technology and suppression PCR. Subsequently, RT-PCR was performed to obtain the Psy cDNA of 1260 bp, which codes 420 amino acids. The Psy gene of total length of 2982 bp was found to consist of five exons and four introns when compared with the cDNA sequence. The *D. salina* Psy amino acid has a 78−89% similarity with many other higher plants, but a phylogenic analysis shows a closer relationship between the microalgal and cyanobacterial Psy.

KEYWORDS: ***Dunaliella salina*; phytoene synthase; sequence analysis; phylogenetic relationship**

1. INTRODUCTION

Carotenoids are a large group of polyene pigments produced by higher plants, certain species of bacteria, fungi, and archaea as well as some algae. Natural carotenoids have various biological functions, such as light harvesting, photoprotection, elimination of free radicals, and as precursors of hormones (*1*). The growing attention paid on these structurally diverse compounds is due to their beneficial effects on human health and their pharmaceutical values (*2*).

The main metabolic pathway of carotenoids is clear and may be common for most of the carotenogenic species (*3*). The early steps of the carotenoid metabolic pathway, which consist of the biosynthesis of phytoene from terpenoid precursors by phytoene synthase as well as the desaturation of phytoene into lycopene followed by two cyclization reactions of lycopene to form α- or β-carotene, have been extensively studied, and the corresponding genes have been isolated from yeast, bacteria, algae, and plants (*4*−*7*).

Since Misawa cloned the carotenoid biosynthesis gene cluster from *Erwinia uredovora* for the first time (*8*) many works have been done on recombinant microorganisms to generate structurally novel or rare high-value carotenoids in nature (*9*−*11*). Besides the assembly of carotenogenic genes into new pathways and molecular incubation of novel metabolic routes, a large pool of available crt genes is also required to produce a mass of high-value carotenoids. Ye obtained "gold rice", which is rich in β-carotene, by engineering the provitamin A biosynthesis pathway into (carotenoid-free) rice endosperm (*12*).

As a green algae *Dualiella salina* could accumulate β-carotene up to 50 mg g^{-1} of dry weight. Under the conditions of high light, high salinity, and nutrient deprivation, it is even up to 10% of the dry weight in *D. salina* (*13*−*15*).

We hypothesized that in *D. salina* there is a specific secondary metabolic pathway to massively synthesize β-carotene, possibly due to a relatively simple reaction sequence or special primary structure of the involved enzymes, such as lycopene cyclases (Lyc) and phytoene synthase (Psy). The crt genes for the putatively efficient carotenoid biosynthetic pathway in *Dunaliella* may significantly contribute to massive production of β-carotene in heterologous host systems such as *Escherichia coli* or in transgenic plants. Additionally, phytoene synthase is often considered to be the first carotenogenic key enzyme in the carotenoid biosynthetic pathway and plays an important control role in the carbon resource flux toward carotenoid synthesis (*16*, *17*).

Therefore, in the present paper the Psy gene was first chosen to illuminate the possibly specific carotenoid metabolic pathway in *D. salina*. Cloning and analysis of the Psy gene would lay a foundation to study the mechanisms for the β-carotene biosynthetic pathway and high accumulation in this green algae.

2. MATERIALS AND METHODS

2.1. Strains and Culture Conditions. *D. salina* cells obtained from the Chinese Academy of Science were grown in defined medium (*18*) containing 2 mol/L NaCl at 26 °C under a 14/10 h dark/light cycle and were collected at the log phase or late log phase. *E. coli* DH5α was used as the host for the multiplication of plasmids.

2.2. Genomic DNA Isolation and Manipulation. Genomic DNA of *D. salina* was isolated following the method described by Yang (*19*). Amplification by PCR using *D. salina* genomic DNA as template was performed with specific primers that corresponded to the strongly conserved amino acid sequence of Psy from two kinds of algae (*Dunalie bardawil* and *Haematococcus pluvialis*) and six kinds of plants (*Arabidopsis thaliana, Oryza sative, Tagetes erecta, Lycopersicon*

* Author to whom correspondence should be addressed (telephone 86-020-87595610; fax 86-020-87113843; e-mail jgjiang@scut.edu.cn).
† South China University of Technology.
§ Chinese Academy of Sciences.

10.1021/jf048358s CCC: $30.25
Published on Web 01/29/2005

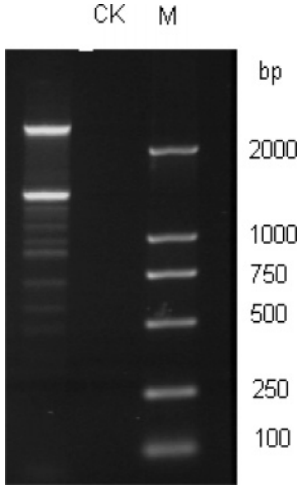

Figure 1. Two fragments amplified from untouched genomic DNA: M, marker; CK, negative control.

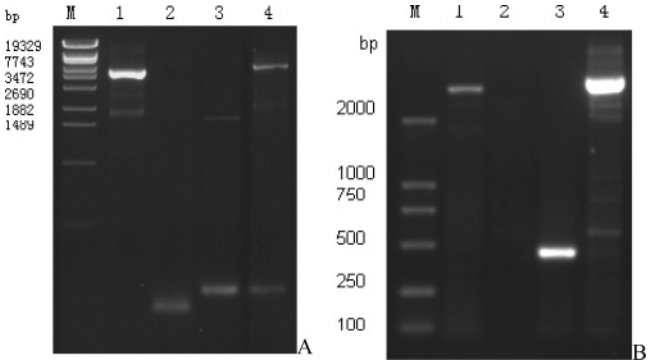

Figure 2. Genome walking PCR products: (**A**, **B**) upstream and downstream walking; M, marker; lanes 1–4, PCR products from *Dra*I-, *Eco*RV-, *Pvu*II-, and *Stu*I-digested genome walker libraries.

esculentum, *Capsicum annuum*, and *Daucus carota*): Psy-F 5′-GAGTACGCCAAGACCTTCTA-3′ (forward) and Psy-R 5′-GTTGTCATAGTCATTCTTCTC-3′ (reverse). The procedure is as follows: 1 cycle of 94 °C, 5 min; 22 cycles of 94 °C, 45 s, 46−35.5 °C, 40 s (decrease 0.5 °C per cycle), and 72 °C, 2 min; 5 cycles of 94 °C, 45 s, 35 °C, 40 s, and 72 °C, 2 min; 1 cycle of 72 °C, 10 min.

The PCR products were separated by electrophoresis in 1.5% agar gels, cloned in the pMD18-T vector (TaKaRa) according to standard methods (*20*), and then sequenced.

For genome walking, *D. salina* genomic DNA was completely digested with E*co*RV, *Dra*I, *Pvu*II, and *Stu*I, respectively, followed by product purification before ligation with Adaptor according to the manual of the Universal GenomeWalker Kit (CLONTECH Laboratories). To obtain the 5′ and 3′ portions of the *D. salina* Psy sequence, two primers, the upstream walking primer UWR 5′-GTGCGTGGCTTGAGCGGCTTAGAAATG-3′ and the downstream walking primer DWF 5′-ATGGTCATCACAGCGGCTTGCTTCACAG-3′, respectively, were designed on the basis of the known Psy gene sequence.

Suppression PCR was carried out using primer pairs AP1/UWR and AP1/DWF as follows: 7 cycles of 94 °C, 2 s, and 72 °C, 3 min; 32 cycles of 94 °C, 2 s, and 67 °C, 3 min; 1 cycle of 67 °C, 4 min.

For other DNA manipulation and technical details, see Sambrook and Russell (*20*).

2.3. Cloning of *D. salina* Psy cDNA. The total RNA was prepared from 10 mL of *D. salina* cells grown at the late log phase with RNASOLV reagent (Omega) according to the manufacture's instruction. RNA was reverse transcribed in a total volume of 10 μL by using an oligo(dT)$_{18}$ primer, and all volume was immediately added to a PCR master mixture using an RNA PCR Kit (AMV) ver. 3.0 (TaKaRa) according to the manufacturer's protocol. The amplification procedure is as follows: 1 cycle of 94 °C, 2 min; 30 cycles of 94 °C, 30 s, 65−50.5 °C, 30 s (decrease 0.5 °C per cycle), and 72 °C, 1 min; 5 cycles of 94 °C, 30 s, 50 °C, 30 s, and 72 °C, 1 min; 1 cycle of 72 °C, 4 min.

The PCR product was cloned and sequenced as above.

On the basis of the previously obtained Psy gene sequence, the following primers were designed for the PCR amplification: 5′-TATACGCGTTTATGCCCTCCACTTCCG-3′ (forward) and 5′-TTTGAGCTCGACTACTGTGGGCTCTTGGG-3′ (reverse), introducing a *Mlu*I and *Sac*I restriction site at the start and stop coding region of the Psy cDNA molecule, respectively.

2.4. Sequencing and Phylogenic Analysis. Sequence analysis and alignments were done with the DNAStar software package (Lasergene). The amino acid sequence was subjected to TMHMM server (http://genome.cbs.dtu.dk/services/TMHMM-2.0/) for transmembrane analysis, and a hidden Markov model (http://genome.cbs.dtu.dk/services/SignalP-2.0/) for the prediction of protein signal sequence. For construction of a phylogenic tree, Clustal X1.8 (NCBI, Bethesda, MD) and TREECONW 1.3b computer programs were used.

2.5. Nucleotide Sequence Accession Number. The NCBI GenBank accession numbers for the phytoene synthase gene and cDNA reported in this paper are GBAN AF305430 and AY601075, respectively.

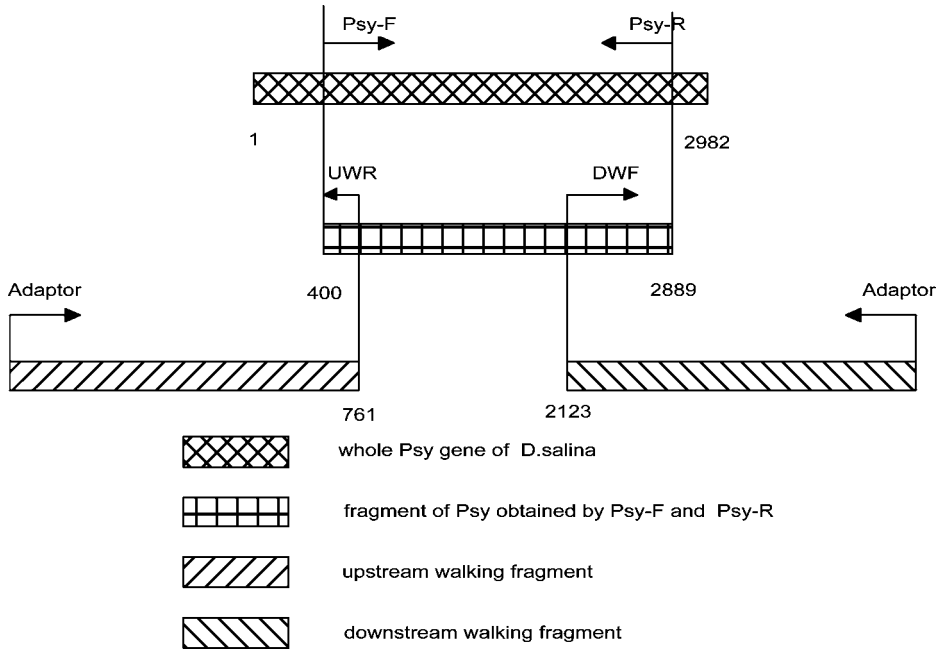

Figure 3. Diagram of overlapping relationship of subclone fragments and positions of corresponding primers.

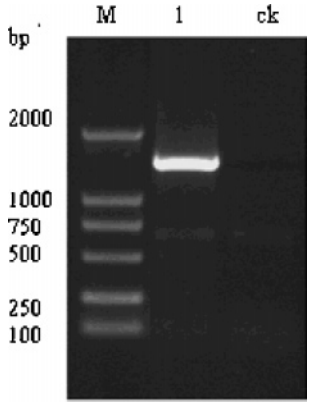

Figure 4. RT-PCR product: M, marker; CK, negative control.

3. RESULTS

3.1. Isolation of the Psy Gene from *D. salina*. The genomic DNA was prepared from cells grown at late logarithmic phase. PCR was done with the untouched genomic DNA by using primer Psy-F/R, and two major bands, 2490 and 1500 bp in length, in 1.5% agar gel were recovered and sequenced (**Figure 1**). A complete BLAST homologous search in the GenBank database showed that the larger fragment has a certain similarity with the Psy gene from other species.

Four genome walker libraries were constructed by ligation of the adaptor with *Dra*I-, *Eco*RV-, *Pvu*II-, and *Stu*I-digested genomic DNAs respectively. They were used as templates to isolate the 5′ upstream and 3′ downstream regions of the *D. salina* Psy gene.

Upstream genome walking was performed with AP1 and primer UWR and a notable 3500 bp band observed only from the *Dra*I-digested genome DNA library, as shown in **Figure 2A**. It was cloned and sequenced, and a BLAST retrieve found its Psy gene homologous sequence and overlap in its C terminus with the previously obtained fragment. A similar downstream genome walking using primer DWF and AP1 found that two fragments, 3000 and 400 bp from *Stu*I and *Pvu*II, respectively, showed similarity with other homologous Psy genes after their sequences were determine (**Figure 2B**). The latter band was consistent with the prediction that there is a *Pvu*II restriction site 400 bp away from the DWF region.

The *D. salina* Psy gene was assembled according to overlapping sequences from the three fragments above and is 2982 bp in length (**Figure 3**).

3.2. Cloning Psy cDNA by Reverse Transcription PCR. Total RNAs extracted from freshly harvested cells were used as templates to amplify *D. salina* Psy cDNA. A 1260 bp fragment (**Figure 4**) was successfully obtained and predicted to be Psy cDNA by using BLAST programs (NCBI) to perform a complete homologous search in the GenBank database.

The differences between the *D. salina* Psy gene and the cDNA sequence were compared. Four conjectural introns separating five exons were 720, 292, 251, and 459 bp in length, respectively (**Figure 5**). A 47.3 kDa peptide coded by Psy cDNA consists of 420 amino acids, p*I* 8.67. Analysis with the TMHMM server showed that there is no transmembrane region contained in this peptide. It has a signal sequence in the C terminus involved in protein transposition according to the analysis with a hidden Markov model.

The BlastP search results demonstrated that the cloned *D. salina* Psy showed at the protein level 81% identity and 89% similarity with *D. bardawil* , 70 and 80% with the *H. pluvialis* counterpart, 67 and 80% with the *A. thaliana* counterpart, and 65 and 78% with *D. carota* Psy; these data demonstrated that the cloned *D. salina* Psy belongs to the Psy family.

Amino acid sequence alignment (data not shown) of *D. salina* Psy with the known alga and plant Psy showed that the conserved regions do not lie in the N terminus; the functional region of this carotenoid synthase may begin from its 119th amino acid, and the first 20 amino acids may function as a transmembrane region.

3.3. Phylogenetic Analysis. Neighbor-joining included in the TREECONW 1.3b software package was used to build a phylogenic tree; analysis of the complete homologous Psy sequences recovered with the tree are shown in **Figure 6**. Psy sequences and GenBank accession numbers are as follows: nine plant species, *C. annuum* (CAA48115), *T. erecta* (AAM45379),

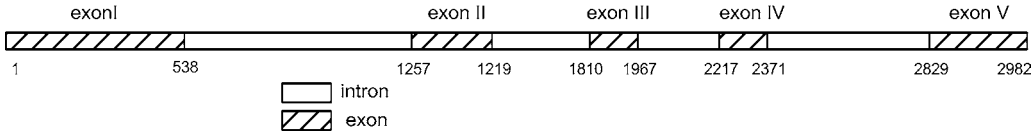

Figure 5. Diagram of exons and introns of Psy gene in *D. salina*.

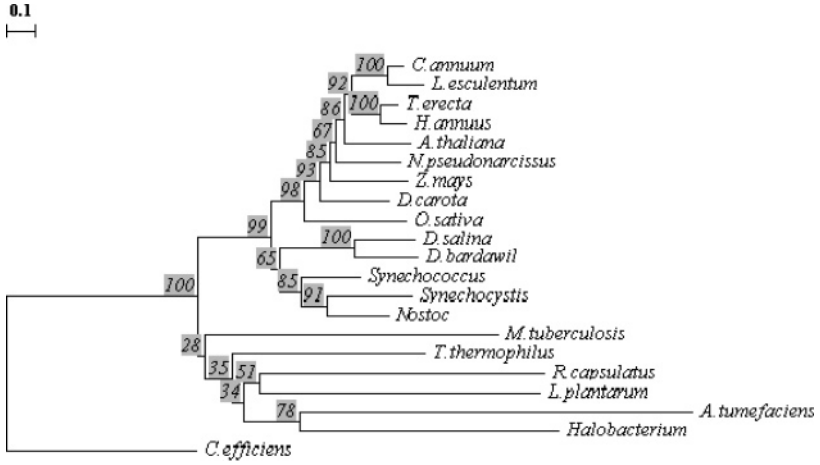

Figure 6. Phylogenic tree of Psy sequences from various species (numbers associated with the branches are bootstrap values; branch lengths are proportional to the number of amino acid substitutions and are measured by the distance scale bar).

A. thaliana (AAM62787), *D. carota* (BAA84763), *O. sative* (AAK07735), *Helianthus annuus* (CAC19567), *L.esculentum* (CAA42969), *Narcissus pseudonarcissus* (S54135), *Zea mays* (S68307); green algae *D. bardawil* (T10702); three cyanobacteria, *Synechocystis* PCC6803 (NP_441168), *Synechococcus* PCC7942 (CAA45350), *Nostoc* sp. PCC7120 (NP_485873); and seven bacteria, *Mycobacterium tuberculosis* H37Rv (NP_217974), *Thermus thermophilus* HB27 (YP_006040), *Rodobacter capsulatus* (P17056), *Lactobacillus plantarum* WCFS1 (NP_786525), *Agrobacterium tumefaciens* (AAL42564), *Halobacterium* sp. NRC-1 (NP_280449), and *Corynebacterium efficiens* YS-314 (NP_737251).

As shown from the tree, Psy catalyzing an early carotenoid biosynthetic pathway step is well conserved among bacteria, cyanobacteria, and eukaryotic organisms. Two defined clusters of Psys can be found. One comprises the enzyme from bacteria. In a separate group, chlorophyte, cyanobacteria, and higher plants align together. The cyanobacterial Psys have a closer relationship with *Dunaliella* Psys than with the plant Psys.

4. DISCUSSION

D. salina is one of the commonly recognized natural sources to produce β-carotene, but the previous work mainly focused on the improvement of culture conditions and extraction methods; however, there is a lack of study on enzymes directly involved in the carotenoid metabolic pathway by using molecular biology technology. Psy is the first carotenoid synthase in plants and plays an important role in production levels of carotenoids. The work of obtaining the Psy cDNA from *D. salina* would help in our understanding of the regulation mechanism of Psy and even the mechanisms for high accumulation of β-carotene in *D. salina*, which aids in the production of masses of natural β-carotene. Additionally, works on transforming the Psy gene of *D. salina* or the whole carotenoid synthesis pathway in crops will help to improve the β-carotene content, which will contribute to solving the problem of vitamin A deficiency in developing countries.

From the known Psy gene of eukaryotic organisms, the Psy gene in *D. salina* has a smaller number of exons than the common five to seven exons in plants, and the exons of Psy in *D. salina* and other higher plants have similar locations. Alignments at the protein level show the sequence differences between plant and bacterial Psys are mainly found at the N terminus due to the presence of a signal peptide responsible for the localization of these enzymes in chloroplasts and chromoplasts (*21*). This is also demonstrated by analysis using a hidden Markov model. The relatively conserved C terminus may be involved in the catalytic activity or substrate recognization/binding (*22*).

Studies on the N-terminal sequence differences and upstream regions of *D. salina* Psy would be helpful to illuminate the possible specific regulation mechanisms for Psy.

LITERATURE CITED

(1) Vershinin, A. Biological functions of carotenoids—diversity and evolution. *Biofactors* **1999**, *10*, 99−104.

(2) Van den Berg, H.; Faulks, R.; Granado, H. F.; Hirschberg, J.; Olmedilla, B.; Sandmann, G.; Southon, S.; Stahl, W. The potential for the improvement of carotenoid levels in foods and the likely systemic effects. *J. Sci. Food Agric.* **2000**, *80*, 880−912.

(3) Sandmann, G. Molecular evolution of carotenoid biosynthesis from bacteria to plants. *Physiol. Plant.* **2002**, *116*, 431−440.

(4) Misawa, N.; Shimada, H. Metabolic engineering for the production of carotenoids in non-carotenogenic bacteria and yeasts. *J. Biotechnol.* **1998**, *59*, 169−181.

(5) Sieiro, C.; Poza, M.; de Miguel, T.; Villa, T. G. Genetic basis of microbial carotenogenesis. *Int. Microbiol.* **2003**, *6*, 11−16.

(6) Linden, H. Carotenoid hydroxylase from *Haematococcus pluvialis*: cDNA sequence, regulation and functional complementation. *Biochim. Biophys. Acta* **1999**, *1446*, 203−212.

(7) Hirschberg, J. Carotenoid biosynthesis in flowering plants. *Curr. Opin. Plant Biol.* **2001**, *4*, 210−218.

(8) Misawa, N.; Nakagawa, M.; Kobayashi, K.; Yamano, S.; Izawa, Y.; Nakamura, K.; Harashima, K. Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. *J. Bacteriol.* **1990**, *172*, 6704−6712.

(9) Schmidt-Dannert, C. Engineering novel carotenoids in microorganisms. *Curr. Opin. Biotechnol.* **2000**, *11*, 255−261.

(10) Lee, P. C.; Momen, A. Z. R.; Mijts, B. N.; Schmidt-Dannert, C. Biosynthesis of structurally novel carotenoids in *Escherichia coli*. *Chem. Biol.* **2003**, *10*, 453−462.

(11) Umeno, D.; Tobias, A. V.; Arnold, F. H. Evolution of the C30 carotenoid synthase CrtM for function in a C40 pathway. *J. Bacteriol.* **2002**, 6690−6699.

(12) Ye, X.; Al Babili, S.; Kloti, A., et al. Engineering the provitamin A (β-carotene) biosynthesis pathway into (carotenoid-free) rice endosperm. *Science* **2000**, *287*, 303−305.

(13) Borowitzka, M. *Microalgae Biotechnology*; Elsevier: Amsterdam, The Netherlands, 1989.

(14) Ben Amotz, A.; Avron, M. On the factors which determine massive β-carotene accumulation in the halotolerant alga *Dunaliella bardawil*. *Plant Physiol.* **1983**, *72*, 593−597.

(15) Boussiba, S.; Vonshak, A. Astaxanthin accumulation in the green alga *Haematococcus pluvialis*. *Plant Cell Physiol.* **1991**, *32*, 1077−1082.

(16) Shewmarker, C. K.; Sheehy, J. A.; Daley, M. Seed specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. *Plant J.* **1999**, *20*, 401−412.

(17) Bramley, P.; Teulieres, C.; Blain. Biochemical characterization of transgenic tomato plants in which carotenoids synthesis has been inhibited through the expression of antisense RNA to pTOM5. *Plant J.* **1992**, *2*, 343−349.

(18) Sheffer, M.; Avron, M. Isolation of the plasma-membrane of the halotolerant alga *Dunaliella salina* using sulforhodamine B as a probe. *Biochim. Biophys. Acta* **1986**, *857*, 155−164.

(19) Yang, Z. Y.; Zhang, Q. Q.; Jiao, X. Z. Construction of a genomic DNA library of *Dunaliella salina* [J]. *Acta Phytophysiol. Sinica* **2000**, *26* (1), 25−28.

(20) Sambrook, J.; Russell, D. W. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001

(21) Cunningham, F. X., Jr.; Gantt, E. Genes and enzymes of carotenoid biosynthesis in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998**, *49*, 557−583.

(22) Misawa, N.; Truesdale, M. R.; Sandmann, G.; Fraser, P. D.; Bird, C.; Schuch, W.; Bramley, P. M. Expression of a tomato cDNA coding for phytoene synthase in *Escherichia coli*, phytoene formation in vivo and in vitro, and functional analysis of various truncated gene products. *J. Biochem.* **1994**, *116*, 980−985.

Received for review October 2, 2004. Revised manuscript received December 10, 2004. Accepted December 13, 2004. The project was supported by the National Natural Science Foundation of China (20076020).

JF048358S