Journal of Bioprocessing & Biotechniques

Araya-Garay et al., J Bioproces Biotechniq 2012, 2:4
DOI: 10.4172/2155-9821.1000125

Research Article Open Access

Cloning and Functional Expression of Z-Carotene Desaturase, A Novel Carotenoid Biosynthesis Gene From *Ficus Carica*

Araya-Garay JM[1], Feijoo-Siota L[1], Veiga-Crespo P[1,2], Sánchez-Pérez A[3] and González Villa T[1,2]*

[1]*Department of Microbiology, University of Santiago de Compostela, Spain*
[2]*School of Biotechnology, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain*
[3]*Discipline of Physiology, Bosch Institute, University of Sydney, NSW 2006, Australia*

Abstract

Carotene desaturation, an essential step in the carotenoid biosynthesis pathway, is catalyzed by two enzymes, phytoene desaturase (PDS) and ζ-carotene desaturase (zeta carotene desaturase, ZDS). Here we describe cloning and *E. coli expression* of zdsfc, a novel *Ficus carica* ζ-carotene desaturase catalyzing dehydrogenation of ζ-carotene into neurosporene and finally lycopene. The ζ-carotene desaturase (ZDS) gene was amplified from the fig tree by rapid amplification of cDNA ends (RACE) and spanned a 1746 bp open reading frame (ORF), encoding a protein of 582 amino acid residues with a predicted molecular weight of 64kD. The N-terminal region of this polypeptide contained a putative transit sequence for plastid targeting. By phylogenetic and sequence analyses, zdsfc showed high homology with previously described ζ-carotene desaturases from higher plant species [1-4]. Additionally, sequence analysis revealed a high degree of conservation among plant ZDSs. The deduced ZDS protein, designated zdsfc, also contains an N-terminus dinucleotide-binding, followed by a conserved region identified in other carotene desaturase sequences. These data, taken together, confirm our cloned zdsfc as an integral part of the ZDS family of proteins.

Keywords: *Ficus carica*; Zeta-carotene desaturase; cDNA; Zeta-carotene; Neurosporene; Lycopene

Introduction

Carotenoids are pigments synthetized by plants, fungi, bacteria, and algae with the main function of protecting them from the action of singlet oxygen and other radicals [5]. In plants, carotenoids are either primary or secondary metabolites. As primary metabolites, carotenoids can function as regulators of plant growth and development, as accessory pigments in photosynthesis, as photoprotectors preventing photo-oxidative damage, or as precursors to the hormone abscisic acid (ABA) [6]. They are also responsible for the color of fruits and flowers, generating distinct yellow, orange, and red colors, thus substantially contributing to plant-animal communication [6,7]. In addition, the colors of many carotenoid-accumulating fruits and flowers also increase their appeal and hence their economic value [8,9]. These pigments also play an important protective role, in human and animal diets, as antioxidants.

The main carotenoid metabolic pathway is well known and may be shared by most of the carotenogenic species [10].The first step in this pathway is the condensation of two molecules of geranylgeranyl pyrophosphate (GGPP) to originate the first true carotenoid C_{40} molecule, phytoene [11]. This two-step reaction is catalyzed by a single soluble enzyme: phytoene synthase (PSY) (Figure 1). Two sequential desaturations of phytoene result in the formation of the first phytofluene, followed by ζ-carotene. Both of these reactions are catalyzed by phytoene desaturase (PDS). Two additional desaturations, catalyzed by ζ-carotene desaturase (ZDS), give rise to first neurosporene and finally lycopene, the symmetrical red carotenoid pigment (Figure 1) [12].

Since 1990, when cloned the carotenoid biosynthesis gene cluster from *Erwinia uredovora*, many carotenoid biosynthetic genes have been identified in plants and other organisms We have previously reported the production of β-carotene by expression of recombinant *Ficus carica* lycopene beta-cyclase in *E. coli* [13,14]. Here, we describe cloning and *Escherichia coli expression* of zdsfc, a novel *Ficus carica* ζ-carotene desaturase catalyzing dehydrogenation of ζ-carotene into neurosporene and finally lycopene. The *E. coli* culture containing zdsfc was able to convert ζ-carotene and neurosporene, two substrates of

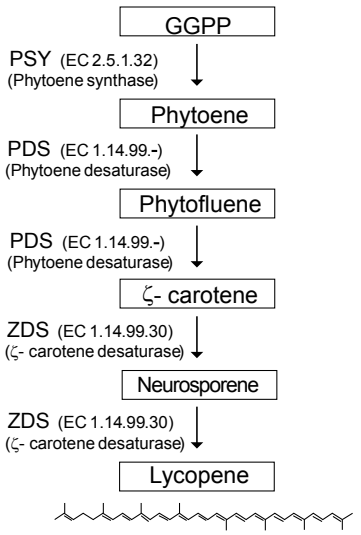

Figure 1: Schematic diagram of the early steps in the carotenoid biosynthetic pathway in plants, from geranyl geranyl pyrophosphate (GGPP) to lycopene, and the enzymes catalyzing those reactions.

***Corresponding author:** González Villa T, Department of Microbiology, University of Santiago de Compostela, Spain, Tel: +34 981592490; Fax: +34 981592490; E-mail: tom.villa@gmail.com

Received April 10, 2012; **Accepted** July 05, 2012; **Published** July 29, 2012

Citation: González Villa T, Araya-Garay JM, Feijoo-Siota L, Veiga-Crespo P, Sánchez-Pérez A (2012) Cloning and Functional Expression of Z-Carotene Desaturase, A Novel Carotenoid Biosynthesis Gene From *Ficus Carica*. J Bioprocess Biotech 2:125 doi: 10.4172/2155-9821.1000125

ζ-carotene desaturase, into neurosporene and lycopene, or lycopene, respectively; thus confirming, *in vivo*, the enzymatic activity of the recombinant ZDS we produced. The *F. carica* enzyme also shared the characteristic properties of other plant ZDSs, and this allowed us to construct a phylogenetic tree illustrating the evolutionary relationship between ZDS and other published carotenoid zeta carotene desaturases (ZDSs) from higher plants.

Materials and Methods

Materials

Plasmids pACCRT-EBI and pACCRT-EBR, encoding *Erwinia uredovora* crtE (GGPP synthase), crtB (phytoene synthase) and crtI (phytoene desaturase), and encoding *Erwinia uredovora* crtE, crtB plus Rodococcus crtR, respectively, were a present from Prof. Misawa (Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Japan). Whereas plasmid pACCRT-EBP, encoding *Erwinia uredovora* crtE, crtB, plus *Synechococcus* crtP was kindly donated by Prof. Gerhard Sandmann (J. W. Goethe Universität, Frankfurt, Germany) [15-17]. Plasmids pCR Blunt II Topo and pUC19, as well as *E. coli* strains TOP10 and BL21 (DE3) and Zero Blunt TOPO PCR Cloning Kit, were from Invitrogen (Carlsbad, CA, USA). The expression plasmid pET21a was purchased from Novagen, (Cambridge, UK). RNeasy Plant Mini Kit was from Qiagen (Valencia, CA, USA). The RT-PCR AMV kit was obtained from Roche Applied Science (Indianapolis, IN, USA). Accuzyme DNA polymerase was from Bioline (Taunton, MA, USA).

Cloning of *Ficus carica* ζ-carotene desaturase cDNA (zdsfc)

Total RNA was isolated, using the RNeasy Plant Mini Kit, from 50 mg of fresh *Ficus carica* leaf, following the instructions recommended by the manufacturer. RT-PCR was carried out using the RT-PCR AMV kit and degenerate primers ZDS DegF and ZDS DegR (Table 1). These oligonucleotide primers were designed using Primer Premier 5.0 (Biosoft International), according to the conserved motifs found on other published zeta-carotene desaturase sequences. Amplification of zdsfc conserved internal DNA fragment was achieved by PCR, using the above-described oligonucleotide primers and 2 µL of first-strand cDNA. The reaction mixture contained polymerase buffer, 0.2 mM of each primer, 1.5 mM $MgCl_2$, 0.2 mM of each deoxynucleotide, and 1 unit of Accuzyme DNA polymerase, following program: 94°C for 2 min, 35 cycles of 94°C for 45 s; 49°C for 1 min and 72°C for 2 min; and a final extension at 72°C for 2 min. The PCR product obtained was subcloned into pCR Blunt II TOPO, using the Zero Blunt TOPO PCR Cloning Kit. All DNA constructs were checked by sequence analysis.

RACE-PCR

The 5′and 3′ ends of zdsfc were obtained by RACE (rapid amplification of cDNA ends), using the 5′/3′ RACE *kit*, 2nd Generation (Roche Applied Science), primers ZDSFc1, ZDSFc2, ZDSFc3, ZDSFc4 and ZDSFc5 (Table 1), and 1 unit of Accuzyme DNA Polymerase. ZDSFc1 was the primer used for first strand cDNA synthesis, whereas the missing 5′ region of the gene was amplified with ZDSFc2 and Oligo dT-Anchor Primer, and incubated as follows: 1 cycle at 94°C for 2 min; 10 cycles of 94°C for 15 s, 55°C for 30 s and 72°C for 40 s; 25 cycles of 94°C for 15 s, 55°C for 30 s and 72°C for 40 s, increasing the elongation time by 20 s for each successive cycle (i.e. elongation time of cycle no.11 is 40 s, of cycle no. 12 is 60 s, of cycle no. 13 is 80 s etc.), and a final extension at 72°C for 7 min.

The PCR-amplified DNA product was then used as a template (dilution 1:20) for a subsequent PCR, with ZDSFc3 and PCR Anchor Primer as primers. This second PCR was incubated as described above, with the exception of the primer annealing step, carried out at 60°C.

Primer name	Primer sequence	Application
ZDS DegF	5′ GGBGAACTTGATTTYCGVTT 3′	RT-PCR
ZDS DegR	5′ ATCACANGCWGCMACATATGCATC 3′	RT-PCR
ZDS-Fc1	5′ CCCCACCTAAGATGAAACCTGCCC 3′	5′ RACE-PCR
ZDS-Fc2	5′ GCGTGCCACCTTTGGAGATGAAC 3′	5′ RACE-PCR
ZDS-Fc3	5′ GCCCCAACTGGAAACCGGAAAT 3′	5′ RACE-PCR
ZDS-Fc4	5′ GCCCAGTTGTAAAGGCTCTCGTTGA 3′	3′ RACE-PCR
ZDS-Fc5	5′ GGTTCATCTCCAAAGGTGGCACGC 3′	3′ RACE-PCR
ZDS-**EcoRI** F	5′ **CGGAATTC**ATGGCTTCTT-GGGTTCTTTTCC 3′	PCR (full ORF)
ZDS-**XhoI** R	5′ **CCGCTCGAG**CTAAACAAGACTA-AGCTTGTC 3′	PCR (full ORF)
	B=C, G, T; Y=C, T; V= A, C, G; N= A, T, C, G; W=A, T; M=C, A	

Table 1: PCR primers used in this study, their oligonucleotide sequence is shown 5' to 3', and the purpose the primers were generated for.
Note: F, forward; R, reverse; ORF, open reading frame.

The 3' end of zdsfc gene was amplified in two sequential PCR reactions, using ZDSFc4 and ZDSFc5, respectively, and the PCR Anchor as primers. The DNA sequence of the PCR-amplified fragments was confirmed by sequence analysis.

PCR primers ZDS-EcoRI F and ZDS-XhoI R were designed to amplify the complete zdsfc coding region (Table 1), using the above described zdsfc DNA fragments as template. These primers contained EcoRI and XhoI restriction sites at their 5' and 3' ends, respectively (in bold in Table 1). The PCR conditions were: 1 cycle at 95°C for 3 min; 35 cycles of 95°C for 10 s, 57°C for 10 s and 72°C for 2 min ; and a final extension at 72°C for 10 min. The resulting PCR product was cloned into a pET21a vector, using the restriction sites present in the primers, thus producing the expression plasmid pET21-zdsfc (Table 2).

Bacterial transformation and growth

Transformation of plasmid pACCRT-EBI into *E. coli* BL21 (DE3) generated *E. coli* BL-pEBI (Table 2), a bacterial strain expressing the enzymes GGPP synthase, phytoene synthase and phytoene desaturase, from *Erwinia uredovora*, and hence capable of producing lycopene. This recombinant *E. coli* was used as positive control in our zeta-carotene desaturase enzymatic assay (see below). Transformation of plasmid pACCRT-EBP into *E. coli* BL21 (DE3) resulted in *E. coli* BL-EBP (Table 2), a bacterial strain expressing Eu-crtE, Eu-crtB and Sc-crtP. *E. coli* BL-pEBR (Table 2) was generated by transforming *E. coli* BL21 (DE3) with plasmid pACCRT-EBR, encoding Eu-crtE, Eu-crtB and Rc-crtR, and thus capable of producing neurosporene. These recombinant *E. coli* strains were used for heterologous complementation assays (see below). All recombinant *E. coli* strains were grown in LB medium (1% tryptone, 0.5% yeast extract, 1% NaCl) for 48 h at 28°C, with agitation at 180 rpm, in the dark to maximize carotenoid production. Chloramphenicol (34 µg/ml) was used for selection of *E. coli* BL-EBI, BL-EBP and BL-EBR transformants. The culture broth was supplemented with 2% agarose for growth on solid medium. Plasmid pET21a, containing the complete coding area of the zdsfc gene (pET21-zdsfc), was transformed into *E. coli* BL21 (DE3) to produce *E. coli* BL-zdsfc (Table 2). *E. coli* BL-zdsfc was grown in LB medium for 48 h at 28°C, in the presence of ampicillin (50 µg/ml), 1 mM IPTG (Isopropyl-β-D thiogalactopyranoside) was added at the end of the logarithmic growth phase to induce ZDS protein expression.

Recombinant strain	Transformed plasmid	Genes encoded	DNA origin	Enzymatic activity	Enzymatic substrate	Enzymatic product	GenBank accession No.
E. coli BL-EBI	pACCRT-EBI	Eu-crtE	*Erwinia uredovora*	GGPP synthase	FPP/GPP	GGPP	D90087
		Eu-crtB	*Erwinia uredovora*	Phytoene synthase	GGPP	Phytoene	D90087
		Eu-crtI	*Erwinia uredovora*	Phytoene desaturase	Phytoene	Lycopene	D90087
					ζ-carotene	Lycopene	D90087
					Neurosporene	Lycopene	D90087
E. coli BL-EBP	pACCRT-EBP	Eu-crtE	*Erwinia uredovora*	GGPP synthase	FPP/GPP	GGPP	D90087
		Eu-crtB	*Erwinia uredovora*	Phytoene synthase	GGPP	Phytoene	D90087
		Sc-crtP	*Synechoccocus sp.*	Phytoene desaturase	Phytoene	ζ-carotene	X55289
E. coli BL-EBR	pACCRT-EBR	Eu-crtE	*Erwinia uredovora*	GGPP synthase	FPP/GPP	GGPP	D90087
		Eu-crtB	*Erwinia uredovora*	Phytoene synthase	GGPP	Phytoene	D90087
		Rc-crtR	*Rhodobacter capsulatus*	Phytoene desaturase	Phytoene	Neurosporene	CAA77540
E. coli BL-zdsFc	pET21-zdsFc	zds-Fc	*Ficus carica*	Zeta carotene desaturase	ζ-carotene	Lycopene	**JN896309**
					Neurosporene	Lycopene	**JN896309**
E. coli BL-EBPZ	pACCRT-EBP	Eu-crtE	*Erwinia uredovora*	GGPP synthase	FPP/GPP	GGPP	D90087
		Eu-crtB	*Erwinia uredovora*	Phytoene synthase	GGPP	Phytoene	D90087
		Sc-crtP	*Synechoccocus sp.*	Phytoene desaturase	Phytoene	ζ-carotene	X55289
	pET21-zdsFc	zds-Fc	*Ficus carica*	Zeta carotene desaturase	ζ-carotene	Lycopene	**JN896309**
					Neurosporene	Lycopene	**JN896309**
E. coli BL-EBRZ	pACCRT-EBR	Eu-crtE	*Erwinia uredovora*	GGPP synthase	FPP/GPP	GGPP	D90087
		Eu-crtB	*Erwinia uredovora*	Phytoene synthase	GGPP	Phytoene	D90087
		Rc-crtR	*Rhodobacter capsulatus*	Phytoene desaturase	Phytoene	Neurosporene	CAA77540
	pET21-zdsFc	zds-Fc	*Ficus carica*	Zeta carotene desaturase	Neurosporene	Lycopene	**JN896309**

Table 2: Recombinant *E. coli* strains generated in this study, and the DNA plasmids used to transform them. The table also includes the genes encoded by the recombinant plasmids, plant species where the genes originated from, as well as the enzyme names and enzymatic activity of the proteins encoded by those genes.

Protein electrophoresis (SDS-PAGE)

For protein analysis, recombinant E coli BL21 (DE3) cells, containing either pET21a or pET21-zdsfc, were grown at 37°C, with orbital shaking (180 rpm), until they reached the exponential growth phase. The cultures were then induced for recombinant protein expression, by addition of 1mM IPTG, and incubated for a further four hours before sample collection and centrifugation, as recommended by the pET21a system (Novagen). Prior to polyacrylamide gel electrophoresis, the bacterial cells were solubilized, at 85°C for 3 min, in a buffer containing 2.5% (w/v) SDS, 125 mM dithiothreitol, 25% (v/v) glycerol, and 112.5 mM Tris-HCl, pH 6.8. The protein samples were then loaded onto a 12% polyacrylamide gel, on a Mini Protean II cell (Bio-Rad, Hercules, CA), as described previously [18]. The gel was run, at 120 V for 1 h, using Tris-Glycine buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS and pH: 8.3). The molecular weight of the proteins was estimated by comparing their gel migration patterns to those of Precision plus Protein Standard (BioRad), using the Quantity One software (Bio-Rad).

Bioinformatic analyses

To construct the phylogenetic tree (Figure 2), the inferred fig tree (*Ficus carica*) ZDS amino acid sequence was compared to 22 homologous amino acid sequences from three citrus trees (*Citrus sinensis, C. unshiu and C. maxima*), physic nut tree (*Jatropha curcas*), pawpaw (*Carica papaya*), strawberry (*Fragaria x ananassa*), two apple trees (*Malus x domestica*), peppers (*Capsicum annuum*), tomato (*Solanum lycopersicum*), sweet potato (*Ipomea sp.*), gentian (*Gentiana lutea*), chrysanthemum (*Chrysanthemum x morifolium*), marigold (*Tagetes erecta*), two wild carrots (*Daucus carota*), daffodil (*Narcissus tazetta var. Chinensis*), orchid (*Oncidium Gower Ramsey*), cress (*Arabidopsis thaliana*), turnip (*Brassica rapa*), sorghum (*Sorghum bicolour*), wheat (*Triticum aestivum*), and green algae (*Dunaliella salina*). Vector NTI Advance 11 software (Invitrogen) and BioEdit Sequence Alignment Editor version 7.0.5.3, were used to analyze the nucleotide and deduced amino acid sequences, and for sequence alignment, respectively. The NCBI database was searched for plant ZDS sequences using the BLAST software [19]. The ChloroP 1.1 Prediction Server program (Emanuelsson et al. 1999 [20]) and TargetP 1.1 Server were used to identify the ZDS signal/sorting peptide and for predicting its cleavage site [21,22]. Phylogenetic and molecular evolutionary analyses were conducted using MEGA, version 5.0, package program [23]. Data were analyzed by the neighbor-joining method [24] . The reliability of the neighbor-joining tree was estimated by calculating bootstrap confidence limits (BCL) based on 1000 replicates [24]. The evolutionary distances were computed using the Jones-Taylor-Thornton method and are in the units of the number of base substitutions per site. The rate variation among sites was modeled with a Gamma distribution (shape parameter=1) [25]. The GenBank accession numbers of the amino acid sequences used in the analysis are shown in Figure 2. Protein sorting was predicted by PSORT, a web server for analyzing and predicting protein-sorting signals, from the Institute for Molecular and Cellular Biology (Osaka, Japan) [26]. Finally, PSIpred v3.0 was used for hydrophobicity and protein secondary structure predictions [27].

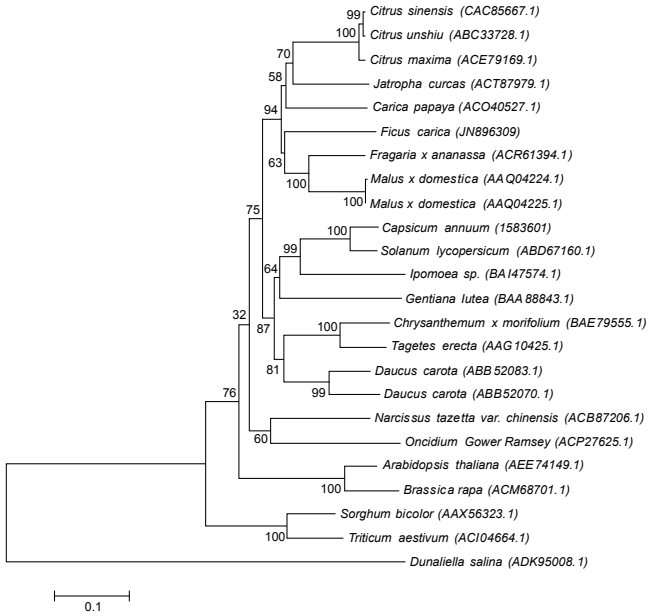

Figure 2: A phylogenetic tree illustrating the inferred evolutionary relationship of the ZDS family of proteins. The tree was generated based on the alignment of the ZDS-Fc amino acid sequence with published plant ZDS protein sequences, and was inferred using the Neighbor-Joining method. The GeneBank accession numbers for the amino acid sequences are in parenthesis, the percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [23]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

HPLC analysis

For HPLC analysis, *E. coli* BL-EBPZ and BL-EBRZ (Table 2) were grown overnight in LB broth, supplemented with chloramphenicol (34 µg/ml) and ampicillin (50 µg/ml), at 37°C with agitation (200 rpm). Recombinant protein expression was then induced, by IPTG addition (1 mM), and the cultures incubated, in the dark, at 28°C for 48h with orbital shaking (200 rpm) After incubation, the cells were centrifuged (4000xg for 5 min), and the bacterial pellet washed twice with deionized water, re-suspended in acetone (Sigma) and homogenized by vortexing (10 min at 4°C). After a new centrifugation (13000xg, 2 min, 4°C), the cell supernatant was collected, dried down under N2 flow, and stored at -80°C for high performance liquid chromatography (HPLC) analysis. All sample manipulations were carried out on ice, under dim light, to prevent photodegradation, isomerization, or other structural changes in the recombinant terpenoids. For HPLC analysis, the dry residues were suspended in 2 ml of chlorophorm:metanol:acetone (3:2:1, v/v) and filtered through a 0.22 µm polycarbonate filter. The HPLC device was equipped with a photodiode array detector, set to scan from 250 to 540 nm throughout the elution procedure, and controlled by the Empower2 software program. Twenty µl of the terpenoid samples were loaded onto a C_{30} carotenoid column (250 mm x 4-6 mm, 5 µm; YMC Europa) and the flow rate was maintained at 1ml/min. The mobile phase was: A, methyl *tert*-butil ether; B, water; and C, methanol. The linear ternary gradient elution program was performed as follows: Initially A-B-C (5:5:90); followed by 0-12 min, A-B-C (5:0:95); 12-20 min, A-B-C (14:0:86); 20-30 min, A-B-C (25:0:75); 30-50 min, A-B-C (50:0:50); 50-70 min, A-B-C (75:0:25); and finally back to A-B-C (5:5:90) for column re-equilibration at 23°C. A Maxplot chromatogram, which plots each carotenoid peak and its corresponding maximum absorbance wavelength, was obtained for each HPLC sample; and the recombinant proteins were identified by comparing their HPLC retention profiles to those of standards run in the same HPLC conditions, or to published data. The lycopene standard was obtained from Sigma-Aldrich (Madrid, Spain), and used as previously described, whereas ζ-carotene and neurosporene were produced by the recombinant *E. coli* cultures BL-EBP and BL-EBR, respectively [28,29]. Quantitative analyses were carried out by comparing the Maxplot chromatogram peak areas of the carotenoid samples to a calibration curve constructed with the lycopene standard.

Zeta-carotene desaturase activity determination

Zeta-carotene desaturase enzymatic activity was determined by functional complementation assays on *E. coli* strains BL-pEBP and BL-pEBR. These strains showed a light yellow and dark yellow coloration, due to ζ-carotene and neurosporene accumulation, respectively. Both *E. coli* BL-pEBP and BL-pEBR were -transformed with the plasmid pET21-zdsfc, to generate *E. coli* BL-EBPZ and *E. coli* BL-EBRZ, respectively. Transformed cells were plated onto LB agar plates containing chloramphenicol (34µg/ml) and ampicillin (50µg/ml). These plates also contained 10µl of a 100 mM aqueous solution of IPTG, added one hour before the cells were plated. The cells were incubated at 28ßC, for 48 h in the darkness, to maximize terpenoid production. After incubation, the desaturase enzymatic activity on the zeta-carotene and neurosporene substrates produced by the *E. coli* strains, was identified by the pigments present on the recombinant colonies, and confirmed by HPLC analysis. The negative control for our assay was an *E. coli* BL21 (DE3) containing the plasmid pET21a alone, whereas recombinant *E. coli* BL-pEBI was used as the positive control.

Results

Cloning the full-length zdsfc cDNA

The strategy used to construct the full-length coding region of *Ficus carica* zds (zdsfc) is outlined in Figure 3, and the primers used for this purpose are listed in Table 1. As a first step, we analyzed the published zds sequences, obtained from different plant species, and identified conserved regions in the gene. Based on these regions, we designed degenerate primers ZDS DegF and ZDS DegR. We used these primers to RT-PCR amplify the RNA we isolated from *Ficus carica* leaves, and this generated a DNA fragment of 567 bp (Figure 3A). By sequence analyses, the amplified DNA was shown to span to the central area of zdsfc, and displayed a high similarity (80% to 83%) to other plant zeta-carotene desaturase genes, such as those from *Citrus unshi* (AB072343.1), *Citrus maxima* (EU798286.1), *Malus x domestica* (AF429983.1), *Citrus sinensis* (AJ319762.1), *Jatropha curcas* (GQ337075.1), *Citrus x paradisi* (AF372617.1), *Fragaria x ananassa* (FJ795343.1), *Carica papaya* (FJ812088.1) and *Daucus carota* (DQ192189.1).

The missing 5'and 3' regions of zdsfc were then obtained by RACE-PCR, using the *Ficus carica* cDNA, generated above by RT, as template, and using primers designed from either the 5' or 3' area, respectively, of the above described central area of zdsfc. This resulted in a 727 bp encompassing the 5'-coding region of the gene (Figure 3B), and a 1318 bp spanning the missing 3'-coding region of zdsfc (Figure 3C). The DNA fragments also share a common area, at either the 5' or 3'-end, respectively, with the central part of the zdsfc gene, used to design the PCR primers (Figure 3E). Hence, the three DNA fragments together span the full-length coding region of zdsfc, and were used to

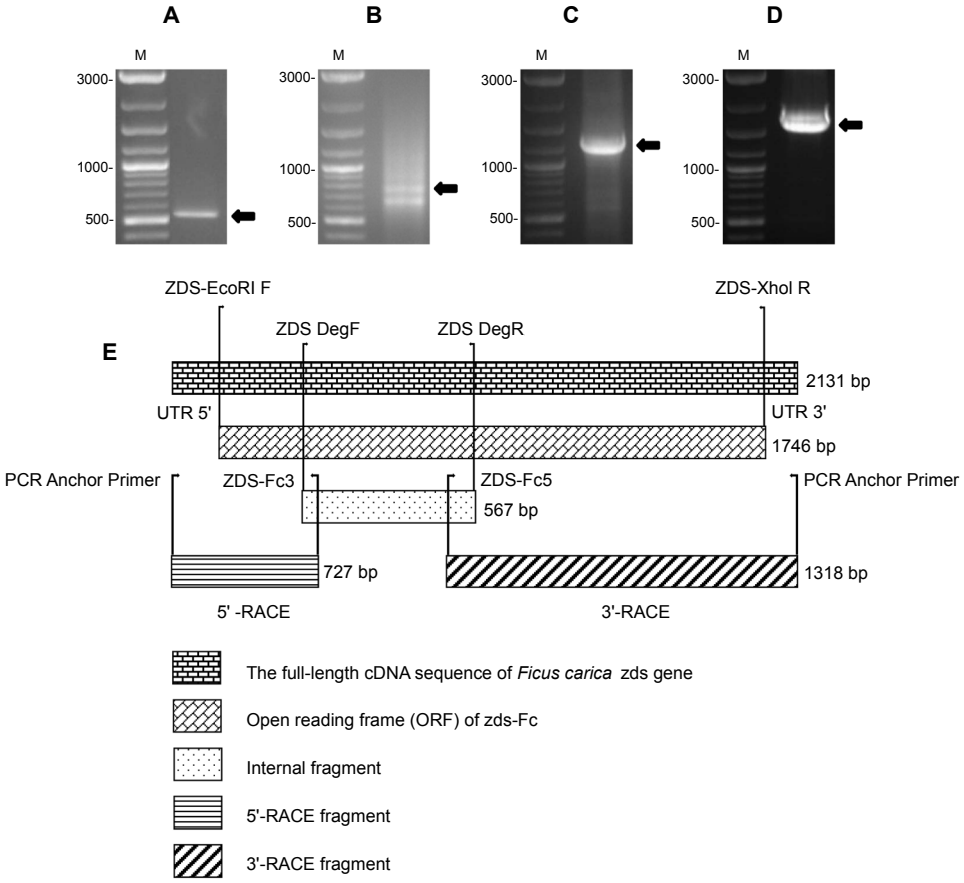

Figure 3: Isolation of the full-length zeta-carotene desaturase cDNA from *Ficus carica*. Agarose gels (1%) stained with ethidium bromide show RT-PCR amplified products: A) Internal zds-Fc cDNA fragment of 567 bp, amplified with primers ZDS DegF and ZDS DegR; B) DNA fragment of 727 bp, corresponding to the 5´-end of the zds-Fc cDNA, obtained by 5´-RACE; C) The 1318 bp cDNA fragment amplified by 3´-RACE, corresponding to the 3´-coding region of zds-Fc; D) A DNA fragment of 2131 bp, spanning the complete open reading frame of the zds-Fc gene from *Ficus carica*. E) Diagram of the strategy used to construct the full-length coding region of *Ficus carica* zds-Fc (1746 bp), indicating the length of the PCR-amplified fragments and the PCR primers used.

PCR amplify a 2131 bp DNA fragment. Apart from the zdsfc coding region, the PCR product also contained 197 nucleotides of 5´-UTR and 188-nucleotides of 3´-UTR (data not shown). Finally, we sequenced the zdsfc cDNA from *Ficus carica* and deposited it in as the GenBank data base (accession number of JN896309). The complete open reading frame (ORF) of the gene has a length of 1746 bp (Figure 3D).

Ficus carica ZDS protein characterization and recombinant expression

The zdsfc gene coded for a polypeptide of 582 amino acids, with an estimated molecular weight of 64 kD. The ZDSFc polypeptide sequence was compared to published zeta-carotene desaturase proteins from other plants. ChloroP 1.1 and TargetP 1.1 predicted a putative chloroplast transit peptide targeting sequence, encompassing the N-terminal amino acids 1 to 43 of ZDSFc; this correlates well with the transit peptides found in maize, *Arabidopsis, Narcissus,* and *Triticum*, which are 30, 34, 42 and 32 residues long, respectively [2,30].The PSORT program predicted a putative polypeptide cleavage site, located: between amino acids 33 and 34, whereas PSIpred found, in our *F. carica* protein (ZDSFc), a highly conserved motif characteristic of higher plants ZDS enzymes. Another conserved motif present in ZDSFc is a typical pyridine dinucleotide binding domain (FAD-binding domain), with a secondary structure of β-sheet-α-helix loop-β-sheet, that is also characteristic of other plant ZDSs [3,31].

The degree of identity (and homology) between the ZDSFc polypeptide and other plant ZDSs was examined by BlastP search. This analysis revealed that our enzyme has an 84% amino acid identity (91% similarity) with *Jatropha curcas*, 85% (91%) with *Citrus maxima*, 85% (91%) with *Citrus unshiu*, 84% (91%) with *Citrus sinensis*, 84% (90%) with *Malus x domestica*, 83% (91%) with *Fragaria x ananassa*, and 83% (90%) with *Carica papaya*. These data, taken together, confirm the cloned zdsfc as an integral part of the ZDS family of proteins.

Expression of *Ficus carica* zds was confirmed by SDS/PAGE analyses of *E. coli* lysates transformed with the plasmid pET21a, either alone, or containing the complete coding area of the zdsfc gene (*E. coli* BL-zdsfc, Table 2). A protein with a molecular weight of 64 kD was observed only in the bacterial cells expressing zdsfc, and only after protein expression was induced by addition of IPTG (data not shown).

Phylogenetic analysis of plant ZDS proteins

We conducted a phylogenetic study of ZDSFc on the basis of its predicted polypeptide sequence and that of a variety of clone plant ZDS sequences, available from GenBank. The phylogenetic tree shown in Figure 2 was inferred using the Neighbor-Joining method [24]. This

test showed that our expressed protein has a high homology to other plant zeta-carotene desaturases. The zeta-carotene desaturase from marine microalgae *Dunaliella salina* was included as an out-group and, hence, appears in an isolated branch. The phylogenetic analyses clearly demonstrated that *Ficus carica* ZDS (JN896309) clusters better with *Fragaria x ananassa* (FJ795343.1), *Malus x domestica* (AF429983.1), *Carica papaya* (FJ812088.1), *Jatropha curcas* (GQ337075.1), *Citrus maxima* (EU798286.1), *Citrus sinensis* (AJ319762.1), and *Citrus unshiu* (DQ309869.1), than with the other ZDSs included in Figure 2. Contrary to expectations, the monocot ZDS proteins from *Narcissus tazetta var. chinensis* and *Oncidium Gower Ramsey* were not more closely related to the grass carotenoid desaturase than the ZDSs from dicot plants, such as *Arabidopsis thaliana or Brassica rapa*. The Bootstrap values on the nodes (Figure 2) indicate the number of times that each group occurred with 1000 replicates. Models with the lowest BIC scores (Bayesian Information Criterion) are considered to describe the substitution pattern as best. So, the evolutionary distances were computed using the Jones-Taylor-Thornton method (Nei and Kumar 2000). The rate variation among sites was modeled with a Gamma distribution (shape parameter=1) with 5 rate categories, and by assuming that a certain fraction of the sites are evolutionarily invariable.

Carotenoid production and ZDSFc enzymatic activity determination

We constructed a recombinant *E. coli* strain (*E. coli* BL-pEBI, Table 2), encoding GGPP and phytoene synthases, as well as phytoene desaturase, hence producing lycopene. We used the cell extracts from this bacterial culture, expressing lycopene, as a positive control for our HPLC analyses. We also constructed two recombinant *E. coli* strains (*E. coli* BL-pEBP and *E. coli* BL-pEBR, Table 2) to use as negative controls. Induction of recombinant protein expression in *E. coli* BL-pEBP should result in the accumulation of ζ-carotene, which can be used as a substrate by zeta carotene desaturase. On the other hand, the recombinant *E. coli* BL-pEBR should produce neurosporene. HPLC analyses confirmed the carotenoids present in the bacterial extracts from our engineered controls as zeta-carotene (*E. coli* BL-pEBP), neurosporene (*E. coli* BL-pEBR) and lycopene (*E. coli* BL-pEBI).

In order to test the enzymatic activity of our recombinant zdsfc, we constructed two more recombinant *E. coli* strains (*E. coli* BL-EBPZ and *E. coli* BL-EBRZ, Table 2). The only difference between the negative control *E. coli* BL-pEBP and *E. coli* BL-EBPZ is the presence in the latter of an additional plasmid (pET21-zdsFc), encoding *Ficus carica* ZDS. Production of a functional ZDSFc enzyme by *E. coli* BL-EBPZ would result in the conversion of the zeta-carotene, produced by *E. coli* BL-pEBP, into neurosporene and lycopene (Figure 1). For its part, *E. coli* BL-EBRZ is only different from the negative control (*E. coli* BL-pEBR) by the presence in the former of an additional plasmid (pET21-zdsFc), encoding *Ficus carica* ZDS. As above, production of a functional ZDSFc enzyme by *E. coli* BL-EBRZ would result in the conversion of the carotenoid accumulated by *E. coli* BL-pEBR (neurosporene) into the product of ZDS activity, in this case lycopene (Figure 1). As expected, ZDS expression by *E. coli* BL-EBPZ resulted in the enzymatic conversion of part of the ζ-carotene, accumulated by the *E. coli* BL-EBP culture, into neurosporene and lycopene. This was confirmed by the HPLC detection of two new compounds, not present in the *E. coli* BL-pEBP culture. The HPLC elution profile for these compounds recorded maximum absorbance peaks at 416, 441 and 471 nm (identified as neurosporene) and 445, 472, and 504 nm (identified as lycopene). Accordingly, ZDS expression by *E. coli* BL-EBRZ resulted in the enzymatic conversion of part of the neurosporene accumulated by the *E. coli* BL-EBR culture into lycopene, a compound not present in the *E. coli* BL-pEBR culture. The new pigment was confirmed as lycopene by HPLC analysis, and exhibited the expected maximum absorbance peaks at 445, 472 and 504 nm. These results, taken together, confirm the enzymatic ability of ZDS to act on both zeta-carotene and neurosporene, as expected from a plant zeta-carotene desaturase.

The neurosporene content of the recombinant *E. coli* BL-EBPZ culture was estimated at 34 μg per g, of cells (dry weight) and the lycopene content at 11 μg per gram of cells (Table 3). Whereas, the lycopene content of the recombinant *E. coli* BL-EBRZ culture was estimated at 13 μg per gram of cells (dry weight).

Carotenoid production was also apparent by the accumulation of pigment in the recombinant *E. coli* cultures and was found to be dependent on the temperature at which the bacteria were grown. Both *E. coli* BL-EBPZ and *E. coli* BL-EBRZ cultures accumulated most pigment when grown at 28°C in the dark.

Discussion

Carotenoids are widely distributed natural pigments responsible for the yellow, orange and red colors of fruits, roots and flowers, as they invariably occur in the chloroplasts of higher plants. But the importance of carotenoids in food goes far beyond their role as natural pigments. Carotenoids are not only the starting material for the synthesis of vitamin A, an essential vitamin, but also carry out a variety of biological functions independent of the provitamin A activity, attributed to the antioxidant property of carotenoids resulting in deactivation of free radicals [8,30]. They have been reported to enhance the immune system, as well as decrease the risk of degenerative diseases such as cancer, cardiovascular disease, age-related macular degeneration, and cataract formation [8,31].

With consumers these days looking for natural products that would offer them more health benefits, carotenoids emerge as a logical target for the food industry. Plant carotenoid production is insufficient to satisfy this market; hence it is essential to generate a sustainable alternative capable of producing large amounts of natural, high quality carotenoids. Carotenoid production can be rapidly increased by the use of recombinant DNA technology, but before this can be done, the carotenoid biosynthetic genes from many different plants and other organisms need to be cloned and characterized. Moreover, heterologous expression of biosynthetic pathways in *E. coli* continues to be a powerful approach for developing metabolic engineering applications in plants. The utility of the bacterial system lies in its inherent similarity to the biochemistry of the plant plastid, thus generating carotenoids that structurally identical to those produced by plants [32].

Here, we describe the cloning and characterization of a novel ζ-carotene desaturase from the fig tree *Ficus carica*. Our strategy was to analyze the published zds sequences, from other plant species, identify conserved regions in the gene and design degenerate PCR primer that were used to clone the central region of the zdsfc gene. The rest of the

E. coli strain	ζ-carotene	Neurosporene	Lycopene
E. coli BL-pEBP	201	n.d.	n.d.
E. coli BL-EBPZ	122	34	11
E. coli BL-pEBR	n.d.	215	n.d.
E. coli BL-EBRZ	n.d.	186	13

Table 3: Carotenoid production by four of the recombinant *E. coli* strains we generated to assess the enzymatic activity of recombinant zds-Fc. The figures indicate the μg of carotenoid per g of cells (dry weight) produced by the bacterial cultures, whereas n.d. stands for not determined.

gene was then obtained by RACE-PCR. The sequence amplified for the zeta-carotene desaturase gene from *Ficus carica* is 2131 bp long and spans a 1746 bp open reading frame, encoding a protein of 582 amino acid residues, with an estimated molecular weight of 64 kD.

The ZDSFc polypeptide sequence was compared to published zeta-carotene desaturase proteins from other plants and was found to share many characteristics common to other plant ZDS s so, as it is the case for other enzymes in the carotenoid biosynthesis pathway, ZDS appears to be highly conserved in higher plants [2]. Another common characteristic was the presence in ZDS of an N-terminal putative chloroplast transit peptide targeting sequence, which was predicated to be 43 residues long, this is also typical of other enzymes involved in carotenoid biosynthesis in plants [8,10,33]. Finally, the sequences also contained a typical pyridine dinucleotide binding domain (FAD-binding domain), with a secondary structure of β-sheet-α-helix loop-β-sheet, that is characteristic of other plant ZDSs [3,29] . These data, taken together, confirms the cloned zdsfc as an integral part of the ZDS family of proteins.

Based on the shared sequences between the *F. carica* enzyme and other plant ZDSs, we constructed a phylogenetic tree illustrating the evolutionary relationship between the fig tree ZDS and other published carotenoid zeta carotene desaturases from higher plants. *Ficus carica* ζ-carotene desaturase was further characterized as part of the ZDS family of plant enzymes, and it clustered better with the dicot plant trees ZDSs such as *Citrus unshiu, Jatropha curcas, Carica papaya* and *Malus x domestica*.

Recombinant expression of *Ficus carica* zds in *E. coli* produced a protein with a molecular weight of 64 kD, consistent with that estimated from the amino acid sequence of the polypeptide. In order to test the *in vivo* enzymatic activity of our recombinant zdsfc, we engineered four recombinant *E. coli* strains (*E. coli* BL-EBP, BL-EBR, BL-EBPZ, and BL-EBRZ, Table 2). Whereas the first two bacterial cultures only accumulated ζ-carotene and neurosporene, respectively, the last two recombinant strains (*E. coli* BL-EBPZ, and BL-EBRZ), expressing our recombinant zdsFc, were able to convert the ζ-carotene and neurosporene into neurosporene and lycopene, or lycopene, respectively. These results demonstrated, *in vivo*, that ZDS, a single desaturase plant enzyme, efficiently catalyzed the last two desaturation steps leading to the production of Lycopene (Table 3). Hence, as described for other plant enzymes the *Ficus carica* zds gene we cloned encodes a zeta-carotene desaturase which is capable of catalyzing the conversion of ζ-carotene into first neurosporene and then lycopene [1-3].

In summary, this study, represents the first time that enzymes the Ficus carica zds was, not only cloned and characterized, but was also shown, by recombinant expression in *E. coli*, to catalyze the desaturation of ζ-carotene to produce first neurosporene and finally lycopene.

Acknowledgements

J.M. A-G is the recipient of an AECID scholarship from the Spanish Foreign Affairs Ministry.

The authors wish to thank Dr. Norihiko Misawa (Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Japan) for the gift of plasmids pACCRT-EBR and pACCRT-EBI; and Dr. Gerhard Sandmann (J. W. Goethe-Universität, Frankfurt, Germany) for the gift of plasmid pACCRT-EBP. Thanks are also due to both the Faculty of Pharmacy and the School of Biotechnology for their support throughout this project.

References

1. Al-Babili S, Oelschlegel J, Beyer P (1998) A cDNA encoding for beta carotene desaturase (Accession No. AJ224683) from *Narcissus pseudonarcissus* L. Plant Physiol 117: 719.
2. Cong L, Wang C, Li Z, Chen L, Yang G, et al. (2010) cDNA cloning and expression analysis of wheat (*Triticum aestivum* L.) phytoene and zeta-carotene desaturase genes. Mol Biol Rep 37: 3351-3361.
3. Matthews PD, Luo R, Wurtzel ET (2003) Maize phytoene desaturase and zeta-carotene desaturase catalyse a poly-Z desaturation pathway: implications for genetic engineering of carotenoid content among cereal crops. J Exp Bot 54: 2215-2230.
4. Yan P, Gao XZ, Shen WT, Zhou P (2011) Cloning and expression analysis of phytoene desaturase and ζ-carotene desaturase genes in *Carica papaya*. Mol Biol Rep 38: 785-791.
5. Krinsky NI (1989) Antioxidant functions of carotenoids. Free Radic Biol Med 7: 617-635.
6. Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. Curr Opin Plant Biol 4: 210-218.
7. DellaPenna D, Pogson BJ (2006) Vitamin synthesis in plants: Tocopherols and carotenoids. Annu Rev Plant Biol 57: 711-738.
8. Fraser PD, Bramley PM (2004) The biosynthesis and nutritional uses of carotenoids. Prog Lipid Res 43: 228-265.
9. Rodríguez Concepción M (2006) Early steps in isoprenoid biosynthesis: Multilevel regulation of the supply of common precursors in plant cells. Phytochem Rev 5: 1-15.
10. Sandmann G, Romer S, Fraser PD (2006) Understanding carotenoid metabolism as a necessity for genetic engineering of crop plants. Metab Eng 8: 291-302.
11. Park H, Kreunen SS, Cuttriss AJ, Dellapenna D, Pogson BJ (2002) Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation, and photomorphogenesis. Plant Cell 14: 321-332.
12. Sandmann G (1994) Phytoene desaturase - Genes, enzymes and phylogenetic aspects. J Plant Physiol 143: 444-447.
13. Misawa N, Nakagawa M, Kobayashi K, Yamano S, Izawa Y, et al., (1990) Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. J Bacteriol 172: 6704-6712.
14. Araya-Garay JM, Feijoo-Siota L, Veiga-Crespo P, Villa TG (2011) cDNA cloning of a novel gene codifying for the enzyme lycopene β-cyclase from *Ficus carica* and its expression in the *Escherichia coli*. Appl Microbiol Biotechnol 92: 769-777.
15. Misawa N, Shimada H (1997) Metabolic engineering for the production of carotenoids in non-carotenogenic bacteria and yeasts. J Biotechnol 59: 169-181.
16. Takaichi S, Sandmann G, Schnurr G, Satomi Y, Suzuki A, et al. (1996) The carotenoid 7,8-dihydro-ψ end group can be cyclized by the lycopene cyclases from the bacterium *Erwinia uredovora* and the higher plant *Capsicum annuum*. Eur J Biochem 241:291-296.
17. Linden H, Vioque A, Sandmann G (1993) Isolation of a carotenoid biosynthesis gene coding for zeta-carotene desaturase from *Anabaena* PCC 7120 by heterologous complementation. FEMS Microbiol Lett 106:99-104.
18. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.
19. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410.
20. Emanuelsson O, Nielsen H, Von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci 8: 978-984.
21. Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300: 1005-1016.
22. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of

prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10: 1-6.

23. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol.

24. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406-425.

25. Nei M, Kumar S (2000) Molecular Evolution and Phylogenetics. Oxford University Press, New York.

26. Nakai K, Horton P (1999) PSORT: A program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci 24: 34-36.

27. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J.Mol.Biol 292: 195-202.

28. Melendez-Martinez AJ, Vicario IM, Heredia FJ (2003) A routine high-perfomance liquid chromatography method for carotenoid determination in ultrafrozen orange juices. J Agric Food Chem 51: 4219-4224.

29. Wirenga RK, Terpstra P, Hol WGJ (1986) Prediction of the occurrence of the ADP-binding βαβ-fold in proteins, using an amino acid sequence fingerprint. GBB-University of Groningen.

30. Richmond A (2000) Microalgal biotechnology at the turn of the millennium: A personal view. J Appl Psychol 12: 441-451.

31. van den Berg H, Faulks R, Fernando Granado H, Hirschberg J, Olmedilla B, et al. (2000) The potential for the improvement of carotenoid levels in food and the likely systemic effects. J Sci Food Agric 80: 880-912.

32. Gallagher CE, Cervantes-Cervantes M, Wurtzel ET (2003) Surrogate biochemistry: use of *Escherichia coli* to identify plant cDNAs that impact metabolic engineering of carotenoid accumulation. Appl Microbiol Biotechnol 60: 713-719.

33. Bouvier F, Rahier A, Camara B (2005) Biogenesis, molecular regulation and function of plant isoprenoids. Prog.Lipid Res 44: 357-429.