Cloning and characterization of the geranylgeranyl diphosphate synthase (GGPS) responsible for carotenoid biosynthesis in *Pyropia umbilicalis*

Li-En Yang[1,2] · Xing-Qi Huang[1] · Qin-Qin Lu[2] · Jian-Yi Zhu[3] · Shan Lu[1]

Abstract Carotenoid metabolism in red algae is not well understood. Geranylgeranyl diphosphate (GGPP), synthesized by GGPP synthase (GGPS), is a precursor for the biosynthesis of many biologically important metabolites, including carotenoids and chlorophylls. GGPSs have been functionally characterized in many organisms, but not in species of the primitive red algal order Bangiales. Here, we cloned and characterized the gene encoding GGPS (*PuGGPS*) in *Pyropia umbilicalis* (Bangiales). *PuGGPS* encodes a protein of 345 amino acids with an N-terminal transit peptide. The catalytic activity of PuGGPS for the production of GGPP was verified by a color complementation assay in *Escherichia coli* and subsequent high-performance liquid chromatography analysis. Homology modeling of PuGGPS showed that its tertiary structure resembles that of other known GGPSs and that this structure allows for the precise docking of the enzymatic product of PuGGPS, GGPP. When leafy thalli of *P. umbilicalis* were treated with norflurazon, an inhibitor of the key carotenoid metabolism enzyme phytoene desaturase, the expression of *PuGGPS* increased by twofold compared with that of the control in the first 2 h, suggesting a prompt response to metabolic perturbation. Prolonged norflurazon treatment failed to increase *PuGGPS* expression. Sequence analysis showed that PuGGPS shares seven conserved motifs with other previously identified GGPSs from different organisms, including two aspartate-rich GGPS signature motifs. Phylogenetic analysis also indicated that PuGGPS is a member of the type II GGPSs found in eubacteria and plants.

Keywords *Pyropia umbilicalis* · Geranylgeranyl diphosphate · GGPP · GGPP synthase · Carotenoid metabolism

✉ Shan Lu
shanlu@nju.edu.cn

[1] State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, Nanjing 210093, China

[2] Jiangsu Institute of Oceanology and Marine Fisheries, Nantong 226007, China

[3] School of Biology and Food Engineering, Changshu Institute of Technology, Changshu 215500, China

Introduction

Geranylgeranyl diphosphate (GGPP) is an important metabolic intermediate in plants that serves as a precursor for the biosynthesis of tocopherols, gibberellins, carotenoids, chlorophylls, and other diterpenoids (Okada et al. 2000; Lange and Ghassemian 2003). GGPP is usually synthesized via the head-to-tail condensation of three isoprenyl diphosphate (IPP) moieties to the allyl head of dimethylallyl diphosphate (DMAPP), as catalyzed by GGPP synthase (GGPS) (Scolnik and Bartley 1994). Some GGPSs can alternatively use geranyl diphosphate (GPP) or farnesyl diphosphate (FPP) as a substrate to produce GGPP (Hefner et al. 1998). The first *GGPS* genes were identified in the bacteria *Pantoea stewartii*, *Pantoea ananatis* (previously known as *Erwinia herbicola* and *E. uredovora*, respectively), and *Rhodobacter capsulatus*, and were designated as *crtE* genes (Armstrong et al. 1990a, b; Misawa et al. 1990; Mergaert et al. 1993). Subsequently, *GGPS* homologs have been identified in various photosynthetic and non-photosynthetic carotenoid-accumulating organisms, including bacteria, fungi, red yeasts, and higher plants.

The *GGPS* gene family is the largest and most diverse group of isoprenyl synthase genes (van Schie et al. 2013). In higher plants, it usually comprises genes involved in the

biosynthesis of GGPP and other isoprenoids. There are 12 *GGPS* homologs in the genome of *Arabidopsis thaliana*, but only five (*GGPS1*, *GGPS3*, *GGPS4*, *GGPS7*, and *GGPS11*) have been functionally characterized as producers of GGPP (Okada et al. 2000). Although carotenoid biosynthesis exclusively occurs in the plastid in higher plants, the subcellular localization of GGPSs has revealed that they are also present in other compartments, such as mitochondria (Okada et al. 2000; Beck et al. 2013). It has been proposed that plants use different GGPSs to supply GGPP for different metabolic branches beyond this intermediate (Okada et al. 2000; van Schie et al. 2013).

Although the carotenoid metabolic pathway has been well studied in higher plants (Ruiz-sola and Rodríguez-Concepción 2012), it remains poorly understood in red algae. To the best of our knowledge, only two red algal enzymes involved in carotenoid metabolism, a lycopene cyclase of *Cyanidioschyzon merolae* (Cunningham et al. 2007) and a β-carotene hydroxylase (PuCHY1) of *Pyropia umbilicalis* (previously as *Porphyra umbilicalis*) (Yang et al. 2014), have been functionally characterized. Here, we cloned the gene encoding GGPS in *P. umbilicalis* (*PuGGPS*) and determined that it functions in the biosynthesis of GGPP for carotenoid biosynthesis.

Materials and methods

Leafy thalli of *Pyropia umbilicalis* strain P.umb.1 (deposited as UTEX #LB-2591 CCAP 1379/4), which is the same strain previously used for transcriptome sequencing (Chan et al. 2012), were obtained from Dr. Susan Brawley (University of Maine). *P. umbilicalis* thalli were cultivated in Provasoli's enriched seawater (Starr and Zeikus 1993) under previously described conditions (Yang et al. 2013). The materials were rinsed briefly with distilled water, surface dried with paper towels, ground into a fine powder with a mortar and pestle in liquid nitrogen and either stored at −80 °C or used promptly. Leafy thalli used for the quantification of gene expression were treated with 5 μM norflurazon (Sigma, USA) for 2 and 4 h, and untreated samples were used as controls. All materials were cultivated under the same conditions.

Transcriptome analysis and homolog identification

A. thaliana is one of the eukaryotic organisms in which the carotenoid metabolic pathway has been well studied (Ruiz-sola and Rodríguez-Concepción 2012). The open reading frames (ORFs) of 12 genes encoding GGPSs of *A. thaliana* were used as queries to blast against the transcriptome database (http://dbdata.rutgers.edu/nori) of *P. umbilicalis* (Chan et al. 2012). Two *GGPS* genes identified from the

cyanobacterium *Synechococcus* sp. PCC 6301 and the unicellular red alga *Cyanidioschyzon merolae* were also incorporated as queries. We used the tblastx algorithm with an *e*-value cutoff of 1×10^{-10} to screen for putative homologs, as previously reported (Yang et al. 2014). The queries are listed in Table 1.

Cloning of *PuGGPS*

To obtain the coding sequence of a transcript (*P_umbilicalis_esContig5139*) that resembles the known *GGPS* genes, total RNA from *P. umbilicalis* was isolated and reverse transcribed as previously described (Yang et al. 2013). Based on the sequence of the putative homologous transcript, a pair of primers, GGPS-HF and GGPS-ER, was designed to amplify the coding region from our complementary DNA (cDNA) pool. Genomic DNA was also extracted from *P. umbilicalis* as previously described (Yang et al. 2013). All primers used in this study are listed in Table 2. All amplicons were subcloned into a pMD19-T vector (Takara, Japan) and sequenced. The full-length cDNA sequence was deposited in GenBank under the accession number KP863500, and the gene was named *PuGGPS*. High-fidelity PrimeStar DNA polymerase (Takara) was used throughout this study unless otherwise indicated, and dimethylsulfoxide (DMSO) was added to the polymerase chain reaction (PCR) system at a final concentration of 5 % (*v/v*) to overcome the high GC contents of the templates (Sun et al. 2010).

Table 1 Queries of the geranylgeranyl diphosphate synthase (GGPS) genes from *Arabidopsis thaliana* and other organisms used to blast against the *Pyropia umbilicalis* transcriptome database

GenBank accession	Source
BT012158.1 (GGPS1)	*A. thaliana*
NM_127418.1 (GGPS2)	*A. thaliana*
DQ446521.1 (GGPS3)	*A. thaliana*
BT005328.1 (GGPS4)	*A. thaliana*
NM_112311.1 (GGPS5)	*A. thaliana*
BT010463.1 (GGPS6)	*A. thaliana*
AK117933.1 (GGPS7)	*A. thaliana*
NM_112907.1 (GGPS8)	*A. thaliana*
NM_113869.1 (GGPS9)	*A. thaliana*
NM_114027.2 (GGPS10)	*A. thaliana*
NM_119845.3 (GGPS11)	*A. thaliana*
AF324662.2 (GGPS12)	*A. thaliana*
NC_010137.1	*Cyanidioschyzon merolae*
AB016093.1	*Synechococcus* sp. PCC 6301

Table 2 Primers used in the present study

Primer[a]	Sequence[b]
GGPS-HF	5′-ATGCTCCGCACCGAGCCC-3′
GGPS-ER	5′-TCAGTTCCTTCGCTCAAATGATGAAG-3′
GGPS-SacI-HF	5′-GTAC<u>GAGCTC</u>CATGCTCCGCACCGAGC-3′
GGPS-HindIII-ER	5′-CCC<u>AAGCTT</u>GGGTCAGTTCCTTCGCTCA-3′
ACT-QF	5′-ATGAGCAAGCCCGCTGTTGTCATC-3′
ACT-QR	5′-TCCAGCGGATACTTGATGACCAGCA-3′
GGPS-QF	5′-GCGCTGATGAAGTTCCTAGCCAC-3′
GGPS-QR	5′-CCATAAACAACGTGGTTGGTCGGC-3′
PDS-QF	5′-TCCCGCCGCCAACCAAGAAA-3′
PDS-QR	5′-ATGTTGGGGTAGGCGCCGAAAA-3′

[a] The -QF or -QR suffix after the name of the gene indicates the forward or reverse orientation, respectively

[b] Restriction sites in primers are underlined

Heterologous expression and functional characterization of PuGGPS

The ORF of *PuGGPS* was amplified using the primers GGPS-SacI-HF and GGPS-HindIII-ER to incorporate restriction sites. The amplicon was purified, digested with *Sac* I-*Hin*d III, cloned into pET-32b (Merck Millipore, Germany), and sequenced for confirmation. The construct was designated as pET-PuGGPS. To test the enzymatic activity of PuGGPS in vivo, pET-PuGGPS was co-transformed into *E. coli* BL21(DE3) (Merck Millipore) with pAC-94N, which harbors all the genes needed for the biosynthesis of β-carotene in bacteria (*crtB*, *crtI* and *crtY* of *Pantoea stewartii*) except for *GGPS* (*crtE*) (Cunningham and Gantt 2007). The vector pET-NtGGPS1, which was constructed in our lab and harbors an active *GGPS* gene of *Nicotiana tabacum* (GGPS1, KC484701.1) (Orlova et al. 2009), was co-transformed with pAC-94N into *E. coli* BL21(DE3) as a positive control. An empty pET-32b vector was co-transformed with pAC-94N as a negative control. Transformant colonies accumulating β-carotene were selected from Luria broth (LB) plates and then inoculated in LB medium as previously reported (Yang et al. 2014). Isopropyl β-D-1-thiogalactopyranoside (IPTG) (0.5 mM) was used to induce the expression of *PuGGPS* for 3 h at 250 rpm and 28 °C. A 200-μL aliquot of the culture was added to 20 mL of fresh medium containing 0.5 mM IPTG and inoculated under the same conditions. After inoculation, cells in the 20 mL culture were collected by centrifugation at 5000×*g* for 5 min. The pelleted cells were sequentially mixed with 400 μL of 80 % acetone, 250 μL of ethyl acetate, and 250 μL of distilled water for 15 s, with high-speed vortexing after each step. After centrifugation at 10,000×*g* for 5 min at 4 °C, the organic phase was transferred into a new tube and dried under a stream of nitrogen gas. The dried extract was re-dissolved in 100 μL of ethyl acetate and either analyzed

immediately or stored at −80 °C. A Waters 2695 separation module and 2998 photodiode array detector (PDA) (Waters, USA) were used for high-performance liquid chromatography (HPLC) analysis of carotenoids on a Spherisorb ODS2 column (5 μm, 4.6×250 mm) (Waters) using a 45-min gradient of ethyl acetate (0–100 %) in acetonitrile-water-triethylamine (9:1:0.01, *v/v*) at a flow rate of 1 mL min^{-1} and monitored at 440 nm (Norris et al. 1995). At least five replicates were performed for each sample.

Homology modeling

The protein sequence of PuGGPS was imported into Accelrys Discovery Studio Client (Accelrys, USA) with its transit peptide sequence removed. The truncated sequence was aligned with proteins with available crystal structures in the Protein Data Bank (PDB) using the blast search algorithm. The crystal structures of five proteins with high sequence identities, including chain A of GGPS of *Sinapis alba* (PDB ID: 2J1P), GGPS of *Lactobacillus brevis* (3M9U), GPS of *Mentha piperita* (3KRA), and FPSs of *Marinomonas* sp. MED 121 (4F62) and *Staphylococcus aureus* (1RTR), were used as templates for homology modeling (Levit et al. 2012).

Gene expression profile

To discern whether PuGGPS is involved in carotenoid biosynthesis in *P. umbilicalis*, the expression levels of transcripts encoding phytoene desaturase (PuPDS) (GenBank accession number KP863501) and PuGGPS were relatively quantified after leafy thalli were treated with the carotenoid biosynthesis inhibitor norflurazon for 2 and 4 h. RNA was isolated from norflurazon-treated and control materials using RNAiso Plus and reverse transcribed using a PrimeScript II First Strand cDNA Synthesis Kit (Takara) as described previously (Yang et al. 2013). Quantitative PCR was performed with a StepOne Plus Real-Time PCR System (Applied Biosystems, USA) using a SYBR Premix Ex Taq (Tli RNaseH Plus) Kit (Takara). A *P. umbilicalis* homolog (*PuACT2*) of the *Pyropia yezoensis* Actin2 gene was used as a reference (Kitade et al. 2008). The reaction system and thermal profile were described previously (Yang et al. 2014). The primer pairs used for the amplification of fragments of *PuPDS*, *PuGGPS*, and *PuACT2* are listed in Table 1 (each with a -QF or -QR suffix after the name of the gene to indicate the forward or reverse orientation, respectively). For each treatment, PCR was performed in triplicate for each triplicate sample. The expression levels were calculated according to the $2^{-\Delta CT}$ method (Sun et al. 2010).

Phylogenetic analysis

We searched GenBank for GGPS homologs of different types according to previous reports (Wang and Ohnuma 2000; Vandermotern et al. 2009; van Schie et al. 2013). Sequences were aligned using the ClustalX program (Chenna et al. 2003). A phylogenetic tree was constructed using the neighbor-joining (NJ) method in MEGA 5 with 1000 bootstrap replicates (Tamura et al. 2011).

Results

We used the ORFs of the 12 *A. thaliana GGPS* genes as queries to blast against the transcriptome of *P. umbilicalis*, and all queries returned only a single transcript species: *P_umbilicalis_esContig5139*. When we used the *GGPS* genes from *Cyanidioschyzon merolae* and *Synechococcus* sp. PCC 6301 as queries, the same transcript species was identified. We also used these 14 queries to blast against the transcriptome of *Porphyra purpurea*, another Bangiales red algal species, and similarly found only one transcript species (data not shown).

Sequence analysis of the *GGPS* homolog transcript revealed a translation initiation codon (ATG) at the 259th nucleotide. Using a primer pair designed according to the transcript sequence, both cDNA and genomic DNA sequences corresponding to this putative gene were amplified, cloned, and sequenced. ORF of this putative gene was found to be 1038 bp, encoding a protein of 345 amino acid residues (data not shown), of which the relative molecular weight was approximately 35 kDa. No introns were found when comparing the cDNA and genomic DNA sequences of *PuGGPS*. The sequences sharing the highest sequence identities with PuGGPS were GGPSs from *Thalassiosiora pseudonana* (65 %),
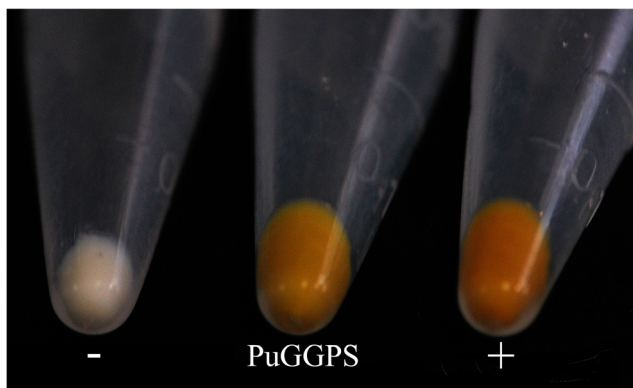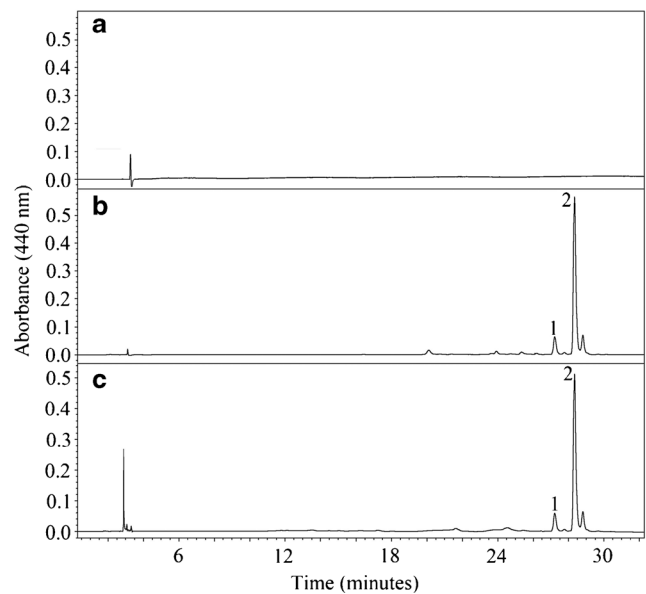

Fig. 2 HPLC analysis of pigment accumulation in *E. coli* cells harboring different combinations of plasmids after IPTG induction. **a** pET-32b + pAC-94N as the negative control, **b** pET-PuGGPS + pAC-94N, and **c** pET-NtGGPS1 + pAC-94N as the positive control. The peaks are lycopene (*1*) and β-carotene (*2*)

Synechococcus sp. JA-3-3Ab (58 %), and *Cyanidioschyzon merolae* (57 %). A transit peptide that targets this protein to the plastid was predicted at the N-terminus (M^1-D^{65}) by the online server TargetP (Emanuelsson et al. 2000).

Because previous studies have shown that GGPS homologs might not produce GGPP alone and that they are involved in different metabolic pathways, we assessed the function of this PuGGPS by constructing a pET-PuGGPS expression vector, transforming it into *E. coli* BL21(DE3) and inducing the expression of a fusion protein with IPTG. The expressed PuGGPS with a His tag had a molecular weight of 55.3 kDa (data not shown), indicating that this construct could be used for a color complementation assay. This construct was then

Fig. 1 Color complementation results for PuGGPS in *Escherichia coli*. BL21(DE3) cells harboring different combinations of plasmids were cultivated in LB medium and induced with IPTG for 3 h before the cells were pelleted to show the accumulation of β-carotene. −, pET-32b + pAC-94N as the negative control; *PuGGPS*, pET-PuGGPS + pAC-94N; +, pET-NtGGPS1 + pAC-94N as the positive control
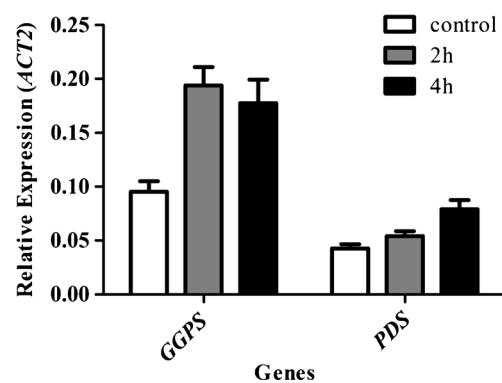

Fig. 3 Quantitative real-time PCR determination of the transcript abundances of *PuGGPS* and *PuPDS* in leafy thalli that were treated with norflurazon for 2 and 4 h. Untreated thalli were used as control. Relative expression was calculated as the ratio between the expression levels of the gene studied and the reference gene *PuAct2*. The data represent the mean±SE (*n*=6)

```
AtGGPS6    MRPR-----------YSLILSAMR-LIRPSNRR--------------------------LSSIASSDSE-------FISYMKNKAKSINKALDNSI   51
AtGGPS4    MEAQNIFLYLLIVFLSLHFVFTTLKGRISPANTRRLIRLLHIPIKSPVAAAIFARKDTREFLDSSIKLVNEEDDFGFSFDFKPYMISKAETINRALDEAI  100
AtGGPS11   MASVTLGSWIVVHHNHHHPSSILTKSRSRSCPITLTKPISFRSKRTVSSSSSIVSSSVVTKEDNLRQSEPS-----SFDFMSYIITKAELVNKALDSAV   95
PuGGPS     MLRTEPGA-----PAASATTDAMPAATVAPTNGV-----------------AGANGAVAPSAGTVALMKF-------LATOKDSVDTALEAAVDASV   68
SeGGPS     MISADP-----------------PQVHP--------------------------AGTFDVKAY-------LKERQ--ALVEAALEASI   36

AtGGPS6    PLCNNFVPLWEPVLEVHKAMRYTLLPGGKRVRPMLCLVACELVGGQESTA--MPAACAVEMIHAASLILDDDLPCMDDDSLRRGKPTNHKVFGEKTSILAS  149
AtGGPS4    PLIEP--------LNIHKAMRYAILAGGKRVRPILCLAACELVGGEERLA--IQACAVEMIHTMSLIKDDLPCMDNDDLRRGKPTIHKVFGESVAIISG  190
AtGGPS11   PLREP--------LKIHEAMRYSLLAGGKRVRPVLCIAACELVGGEESTA--MPAACAVEMIHTMSLIHDDLPCMDNDDLRRGKPTNHKVFGEDVAVLAG  185
PuGGPS     APAGPETA------TISDAMKYSLRAGGKRVRPALTLAAASLFGGEAGMAAAMPSAVEMIHTMSLIHDDLPAMDNDDLRRGLPTNHVVYGEDVAILAG  162
SeGGPS     PVAYPEK--------IYDAMRYSLMAGGKRLRPILCLATCELMGGTVEMA--MPTACALEMIHTMSLIHDDLPAMDNDDYRRGKPTNHKVYGEDIAILAG  126
                                                                       FARM

AtGGPS6    NALRSLAVKQTLASTSLGVTS-ERVLFAVQEMARAVGTEGLVAGQAADLAGERMSFKNEDDELRYLELMHVHKTAVLVEAAAVVGAIMGGGSDEEIERLK  248
AtGGPS4    GALIALAFBH---LTEADVSS-KKMVRAVKELAKSIGTKGLVAGQAKDLSSE--GLEQNDVGLEDLEYIHVHKTGSLLEASAVIGAVIGGGIEKEIEKVR  284
AtGGPS11   DALISFAFBHLASATSSDVVSPVRVVRAVGELAKAIGTEGLVAGQVVDISSE--GLDLNDVGLEHLEFIHLHKTAALLEASAVLGAIVGGGSDDEIERLR  283
PuGGPS     DALIALAFBHVAKAIEG-VDP-RRVVSVLGVLGACVGARGLAGGQVMDLESE--GKGDGEVTMETLTWIHTHKTAALLVASAVAGATVAGASDEDVGRLR  258
SeGGPS     DGLLAYAFEYVVEQTKN-VPA-EYLLKIVARLGHAVAATGLVGGQVVDLECE--GQPD--IGLETLHFIHSHKTGALLEASVVSGALLTGAHESDVARLS  220

AtGGPS6    SYARCVGLMFQVMDDVLDETKSSEELGKTAGKDLITGKLTYPKVMGVDNAREYAKRINREAQEHLQGFDSDKVVPLLSLADYIVKRQN  336
AtGGPS4    NFARCIGLLFQVVDDILDETKSSEELGKTAGKDKVAGKLTYPKVIGVEKSKEFVEKIKRDAREHLQGFDSDKVKPLIALTNFIANRNH  372
AtGGPS11   KFARCIGLLFQVVDDILDVTKSSKELGKTAGKDLIADKLTYPKIMGLEKSREFAEKLNREARDQLLGFDSDKVLPILLALANYIAYRQN  371
PuGGPS     TFATKIGLAFQIADDVLDVTSSSTVLGKTSGKDEAVNKVTYPRLLGLDGARARAEELVAEAKACLDVYG-DKAAVLSELADFIIERRN  345
SeGGPS     RYAANIGLAFQIVDDILDITSTRDVLGKTVGKDVAAQKMTYPRIWGLEKSRQEAERIVAEAKAELAVYG-AAAVPLQAIADYITRSH  307
                           SARM
```
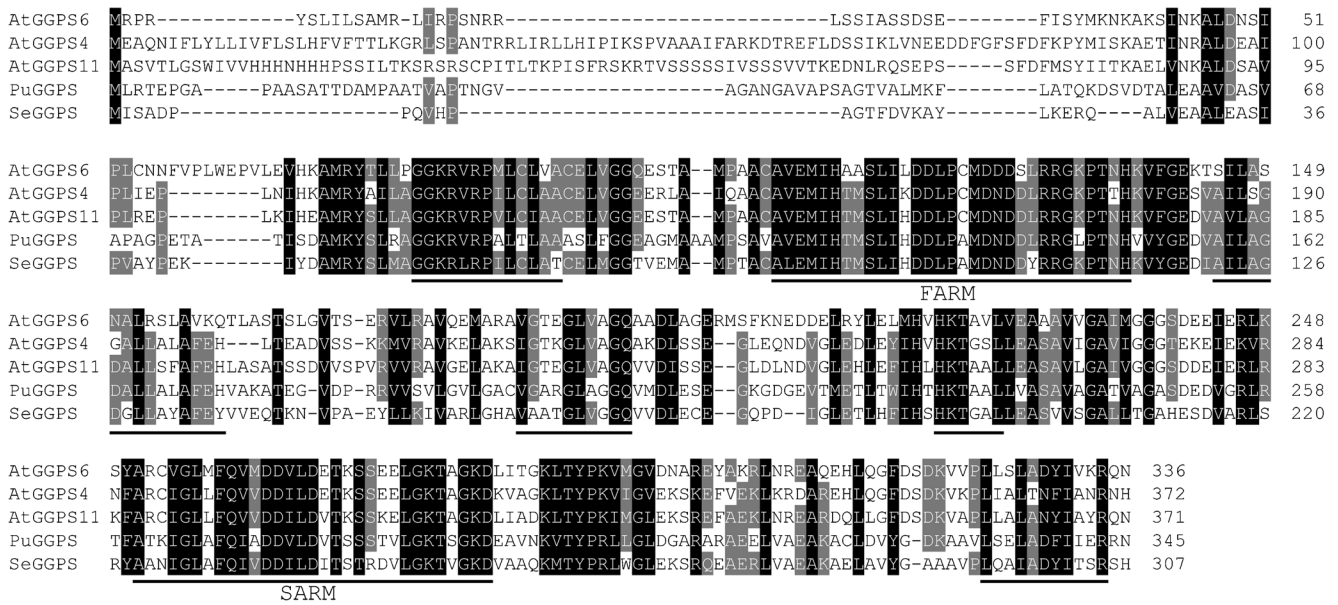
Fig. 4 Alignment of the sequences of GGPS homologs from *Pyropia umbilicalis* (PuGGPS), *Arabidopsis thaliana* (AtGGPS), and *Synechococcus* sp. PCC 6301 (SeGGPS). Conserved motifs are underlined, and the first and second aspartate-rich motifs (FARM and SARM, respectively) are labeled

co-transformed into *E. coli* BL21(DE3) cells with pAC-94N, which harbors the full set of genes needed to synthesize β-carotene when GGPP is supplied. When we cultivated the transformed bacteria in LB medium and induced gene expression with IPTG for 3 h, the *E. coli* cells co-transformed with pET-PuGGPS and pAC-94N changed from white to orange yellow, indicating β-carotene accumulation and demonstrating that the expressed PuGGPS was able to produce GGPP for the subsequent metabolic processes driven by the enzymes expressed by pAC-94N. Bacterial cells transformed with both pET-NtGGPS1, which expresses a functionally confirmed tobacco GGPS, and pAC-94N also accumulated β-carotene, as expected. Bacterial cells harboring a pET-32b empty vector and pAC-94N did not turn orange yellow (Fig. 1).

The pigments in the cultivated *E. coli* cells were extracted and separated by HPLC using previously reported methods (Yang et al. 2014). Our results showed there was no carotenoid accumulation in the negative control harboring pET-32b and pAC-94N (Fig. 2a). However, cells harboring pAC-94N with either pET-PuGGPS or pET-NtGGPS1 accumulated β-carotene together with a low amount of lycopene (Fig. 2b, c). This result indicates that PuGGPS is capable of synthesizing GGPP in vivo for the biosynthesis of carotenoids, such as NtGGPS1.

The herbicide norflurazon is an inhibitor of phytoene desaturase (PDS), which is a key enzyme in the biosynthesis of carotenoids (Chamovitz et al. 1993). To determine whether PuGGPS is functionally involved in carotenoid biosynthesis *in planta*, we quantified the expression of the *PuGGPS* and *PuPDS* genes following the perturbation of carotenoid metabolism by this inhibitor. Our results showed that after

P. umbilicalis leafy thalli were treated with norflurazon for 2 h, the expression of *PuPDS* was slightly increased and that of *PuGGPS* was increased by approximately twofold compared with the control levels. Following treatment with
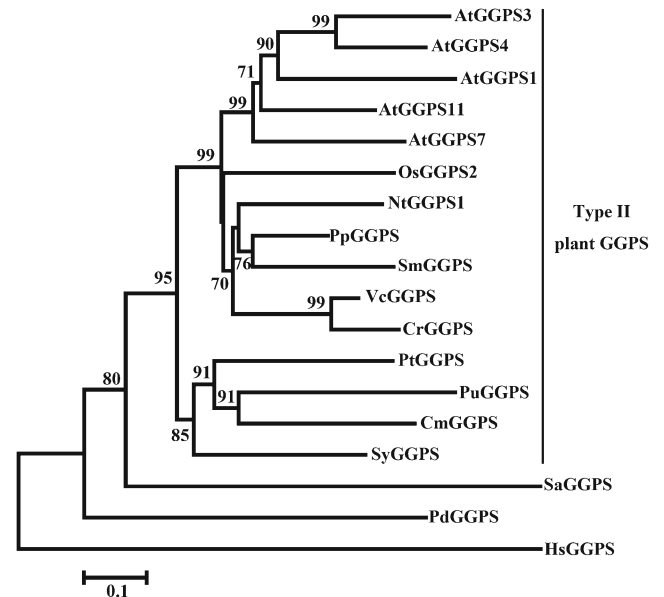
Fig. 5 Phylogenetic analysis of PuGGPS and its homologs from *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Nicotiana tabacum* (Nt), *Physcomitrella patens* (Pp), *Selaginella moellendorffii* (Sm), *Chlamydomonas reinhardtii* (Cr), *Volvox carteri* (Vc), *Synechocystis* sp. PCC 6803 (Sy), *Cyanidioschyzon merolae* (Cm), *Phaeodactylum tricornutum* (Pt), *Pantoea dispersa* (Pd), *Homo sapiens* (Hs), and *Sulfolobus acidocaldarius* (Sa). A bootstrap (1000 replicates) neighbor-joining unrooted phylogenetic tree was generated using MEGA 5. The bootstrap values are labeled beside the branches, and the *scale bar* indicates 10 % sequence divergence

norflurazon for 4 h, the expression of *PuPDS* was increased by twofold, whereas that of *PuGGPS* remained at the same level observed at 2 h (Fig. 3).

In higher plants, GGPSs are encoded by multiple members of a gene family. However, in this primitive red alga, we could not identify other transcript homologs sharing reasonable sequence similarity to suggest a possible GGPS function. For a better understanding of PuGGPS, we blasted its sequence against the non-redundant protein database using the blastp algorithm in NCBI. Most of the hits were GGPSs from different organisms. We aligned the sequences of GGPS11, GGPS4, and GGPS6 from *A. thaliana*, which are known to be localized to different organelles, and that of the GGPS from *Synechococcus* sp. PCC 6301 with the PuGGPS sequence. The multiple alignment showed that PuGGPS shares seven typical conserved motifs with other GGPSs from both cyanobacteria and higher plants, including the first aspartate-rich motif (FARM) and the second aspartate-rich motif (SARM), which are signature sequences for GGPSs (Fig. 4).

GGPSs are categorized into three different types (I, II, and III) based on the sequences of the chain-length determination (CLD) regions, including the FARM and several of its

upstream amino acid residues (Hemmi et al. 2003). PuGGPS is a typeType II GGPS, of which the FARM is DXXXXD (X is any amino acid), instead of DDXXD in the FARM of the other two types (Hemmi et al. 2003). A phylogenetic tree based on GGPS protein sequences of different types (Wang and Ohnuma 2000; Vandermotern et al. 2009; van Schie et al. 2013) indicated that PuGPPS was closely clustered with GGPSs from the cyanobacterium *Synechocystis* sp. PCC 6803, the unicellular red alga *Cyanidioschyzon merolae*, and the diatom *Phaeodactylum tricornutum* (Fig. 5). GGPSs from the green algae *Chlamydomonas reinhardtii* and *Volvox carteri* were more closely clustered with those from higher plants (Fig. 5).

To ascertain whether the tertiary structure of PuGGPS resembles that of GGPSs for which crystal structures have been elucidated, the tertiary structure of PuGGPS was constructed using homology modeling (Levit et al. 2012). The results showed that the constructed three-dimensional model of PuGGPS resembled the crystal structures of GGPSs from *Sinapis alba* (PDB ID: 2J1P) and *Lactobacillus brevis* (3M9U) (Fig. 6a). There were 11 α-helices with no β-sheets in the constructed model, which is a structure typical of prenyl diphosphate synthase called the "terpenoid synthase

Fig. 6 Homology modeling of PuGGPS. **a** 3D structures of PuGGPS (*red*) and homologs from different organisms, showing their similarities. **b** Docking of GGPP (*yellow*) to PuGGPS and the location of Mg^{2+} (*green*). **c** Ramachandran analysis of the 3D structure of PuGGPS. The *red spots* show abnormal intersection angles, and the *green spots* show normal angles. **d** Locations of the amino acids with abnormal intersection angle (indicated with the stick ball)
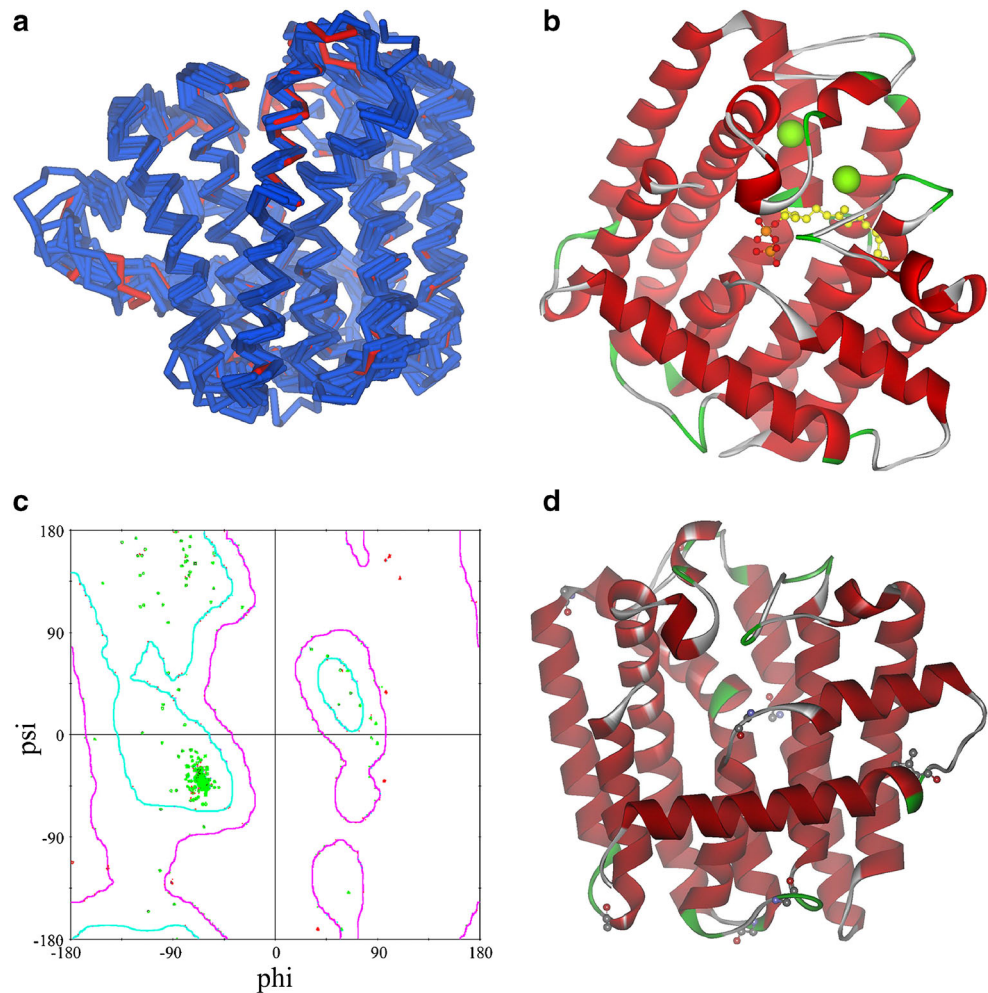
fold." Several of the α-helices form the active cavity, and the product GGPP can precisely dock into this cavity (Fig. 6b). The FARM and SARM signatures are located on the two sides of the cleft to facilitate the chelation of Mg^{2+}, which is required for the activity of GGPS. The reliability of the constructed model was tested using a Ramachandran plot, and we found most of the intersection angles between the amino acids to be reasonable, with a few exceptions. However, abnormal intersection angles were noted between amino acids in the loops or at the corners, which may be caused by the attraction of amino acids during normal folding (Fig. 6c, d).

Discussion

GGPP, synthesized by GGPS, is a precursor for the biosynthesis of many biologically important metabolites, including carotenoids and chlorophylls. The gene that we cloned and identified from the transcriptome of *P. umbilicalis* encodes a protein of 345 amino acids with an estimated molecular weight of 35 kDa, which is consistent with previously reported GGPSs (Engprasert et al. 2004; Thabet et al. 2012). This protein has high sequence similarity with GGPSs from cyanobacteria, unicellular red algae, and diatoms, supporting an early evolution of carotenoid metabolism from cyanobacteria to heterokonts (Yang et al. 2014). Furthermore, PuGGPS shares seven conserved motifs, including the FARM and SARM signatures (Fig. 4), and has a three-dimensional structure that is very similar to those of other GGPSs (Fig. 6a). Phylogenetic analysis also showed that it belongs to the type II GGPSs of eubacteria and plants (Fig. 5) (Wang and Ohnuma 2000; Vandermotern et al. 2009). These results indicate that PuGGPS is a member of the GGPS protein family. This function was further supported by a color complementation assay, HPLC analysis, and three-dimensional modeling (Figs. 1, 2, and 6b).

The application of norflurazon perturbs metabolic flux because of its competition with a cofactor (plastoquinone) required by PDS (Norris et al. 1995; Breitenbach et al. 2001). In our studies, the treatment of *P. umbilicalis* leafy thalli with norflurazon resulted in the prompt induction of *PuGGPS* expression (Fig. 3), suggesting that PuGGPS is likely to be involved in the biosynthesis of carotenoids and that it accumulates in response to the decreased catalytic activity of PuPDS. However, the transcript abundance of *PuGGPS* did not increase with prolonged norflurazon treatment (Fig. 3). One possible reason for this finding is that the norflurazon supplied in the medium was not sufficient to compete with plastoquinone for PuPDS. The decreased catalytic activity of PuPDS was then compensated for by the induced expression of *PuPDS* and the accumulation of the enzyme; thus, it was unnecessary to express *PuGGPS* at a higher level. This agrees with a previous report in higher plants that the promoter activity of *PDS* is activated by

end-product regulation (Corona et al. 1996). Additionally, because the transcriptome of *P. umbilicalis* has been sequenced under different conditions and is thought to include a near-complete repertoire of genes (Chan et al. 2012), PuGGPS is most likely the only enzyme that supplies GGPP for the biosynthesis of not only carotenoids but also chlorophylls and other diterpenoid derivatives; thus, its expression is not regulated solely by the carotenoid metabolic pathway.

Previous studies have shown that carotenoid metabolism in red algae is not as complicated as that in higher plants (Schubert et al. 2006; Cunningham et al. 2007). Epoxycarotenoids, such as antheraxanthin and violaxanthin, have not been identified in either *P. umbilicalis* or other related species (Gantt et al. 2010; Schubert et al. 2006; Yang et al. 2014). Therefore, carotenoids in *P. umbilicalis* are unlikely to be utilized in a xanthophyll cycle to confer photoprotection (Yang et al. 2014). Both cyanobacteria and red algae lack transmembrane light harvesting complex II (LHCII), and it has been reported that carotenoid-binding protein is involved in energy quenching in cyanobacteria (Kirilovsky 2007; Rakhimberdieva et al. 2007). The functions of carotenoids in primitive red algal species remain unclear. The detailed characterization of the entire metabolic pathway under different environmental conditions should shed light on the regulatory mechanisms governing different metabolic branches beyond GGPP and also on the early evolution of carotenoid metabolism.

References

Armstrong GA, Alberti M, Hearst JE (1990a) Conserved enzymes mediate the early reactions of carotenoid biosynthesis in nonphotosynthetic and photosynthetic prokaryotes. Proc Natl Acad Sci U S A 87:9975–9979

Armstrong GA, Schmidt A, Sandmann G, Hearst JE (1990b) Genetic and biochemical characterization of carotenoid biosynthesis mutants of *Rhodobacter capsulatus*. J Biol Chem 265:8329–8338

Beck G, Coman D, Herren E, Ruiz-Sola MA, Rodríguez-Concepción M, Gruissem W, Vranová E (2013) Characterization of the GGPP synthase gene family in *Arabidopsis thaliana*. Plant Mol Biol 82:393–416

Breitenbach J, Zhu C, Sandmann G (2001) Bleaching herbicide norflurazon inhibits phytoene desaturase by competition with the cofactors. J Agric Food Chem 49:5270–5272

Chamovitz D, Sandmann G, Hirschberg J (1993) Molecular and biochemical characterization of herbicide-resistant mutants of cyanobacteria reveals that phytoene desaturation is a rate-limiting step in carotenoid biosynthesis. J Biol Chem 268:17348–17353

Chan CX, Blouin NA, Zhuang Y, Zäuner S, Prochnik SE, Lindquist E, Lin S, Benning C, Lohr M, Yarish C, Gantt E, Grossman AR, Lu S, Müller K, Stiller JW, Brawley SH, Bhattacharya D (2012) *Porphyra* (Bangiophyceae) transcriptomes provide insights into red algal development and metabolism. J Phycol 48:1328–1342

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 31:3497–3500

Corona V, Aracri B, Kosturkova G, Bartley G, Pitto L, Giorgetti L, Scolnik PA, Giuliano G (1996) Regulation of a carotenoid biosynthesis gene promoter during plant development. Plant J 9:505–512

Cunningham FX Jr, Gantt E (2007) A portfolio of plasmids for identification and analysis of carotenoid pathway enzymes: *Adonis aestivalis* as a case study. Photosynth Res 92:245–259

Cunningham FX Jr, Lee H, Gantt E (2007) Carotenoid biosynthesis in the primitive red alga *Cyanidioschyzon merolae*. Eukaryot Cell 6:533–545

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

Engprasert S, Taura F, Kawamukai M, Shoyama Y (2004) Molecular cloning and functional expression of geranylgeranyl pyrophosphate synthase from *Coleus forskohlii* Briq. BMC Plant Biol 4:18–25

Gantt E, Berg GM, Bhattacharya D, Blouin NA, Brodie JA, Chan CX, Collén J Jr, Cunningham FX, Gross J, Grossman AR, Karpowicz S, Kitade Y, Klein AS, Levine IA, Lin S, Lu S, Lynch M, Minocha SC, Müller K, Neefus CD, Oliveira MCD, Rymarquis L, Smith A, Stiller JW, Wu W-K, Yarish C, Zhuang Y, Brawley SH (2010) *Porphyra*: complex life histories in a harsh environment: *P. umbilicalis* an intertidal red alga for genomic analysis. In: Seckbach J, Chapman DJ (eds) Red algae in the genomic age. Springer, New York, pp 129–148

Hefner J, Ketchum REB, Croteau R (1998) Cloning and functional expression of a cDNA encoding geranylgeranyl diphosphate synthase from *Taxus canadensis* and assessment of the role of this prenyltransferase in cell induced for taxol production. Arch Biochem Biophys 306:62–74

Hemmi H, Noike M, Nakayama T, Nishino T (2003) An alternative mechanism of product chain-length determination in type III geranylgeranyl diphosphate synthase. Eur J Biochem 270:2186–2194

Kirilovsky D (2007) Photoprotection in cyanobacteria: the orange carotenoid protein (OCP)-related non-photochemical-quenching mechanism. Photosynth Res 93:7–16

Kitade Y, Nakamura M, Uji T, Fukuda S, Endo H, Saga N (2008) Structural features and gene-expression profiles of actin homologs in *Porphyra yezoensis* (Rhodophyta). Gene 423:79–84

Lange BM, Ghassemian M (2003) Genome organization in *Arabidopsis thaliana*: a survey for genes involved in isoprenoid and chlorophyll metabolism. Plant Mol Biol 51:925–948

Levit A, Barak D, Behrens M, Meyerhof W, Niv MY (2012) Homolog model-assisted elucidation of binding sites in GPCRS. Methods Mol Biol 914:179–205

Mergaert J, Verdonck L, Kersters K (1993) Transfer of *Erwinia ananas* (synonym, *Erwinia uredovora*) and *Erwinia stewartii* to the genus *Pantoea* emend. as *Pantoea ananas* (Serrano 1928) comb. nov., and *Pantoea stewartii* (Smith 1898) comb. nov., respectively, and description of *Pantoea stewartii* subsp. *indologenes* subsp. nov. Int J Syst Bacteriol 43:162–173

Misawa N, Nakagawa M, Kobayashi K, Yamano S, Izawa Y, Nakamura K, Harashima K (1990) Elucidation of the *Erwinia uredovora*

carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. J Bacteriol 172:6704–6712

Norris SR, Barrette TR, DellaPenna D (1995) Genetic dissection of carotenoid synthesis in *Arabidopsis* defines plastoquinone as an essential component of phytoene desaturation. Plant Cell 7:2139–2149

Okada K, Saito T, Nakagawa T, Kawamukai M, Kamiya Y (2000) Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in *Arabidopsis*. Plant Physiol 122:1045–1056

Orlova I, Nagegowda DA, Kish CM, Gutensohn M, Maeda H, Varbanova M, Fridman E, Yamaguchi S, Hanada A, Kamiya Y, Krichevsky A, Citovsky V, Pichersky E, Dudareva N (2009) The small subunit of snapdragon geranyl diphosphate synthase modifies the chain length specificity of tobacco geranylgeranyl diphosphate synthase in planta. Plant Cell 21:4002–4017

Rakhimberdieva MG, Vavilin DV, Vermaas WFJ, Elanskaya IV, Karapetyan NV (2007) Phycobilin/chlorophyll excitation equilibration upon carotenoid-induced non-photochemical fluorescence quenching in phycobilisomes of the *Synechocystis* sp. PCC 6803. Biochim Biophys Acta Bioenerg 1767:757–765

Ruiz-sola MÁ, Rodríguez-Concepción M (2012) Carotenoid biosynthesis in *Arabidopsis*: a colorful pathway. Arabidopsis Book 10, e0158

Schubert N, García-Mendoza E, Pacheco-Ruiz I (2006) Carotenoid composition of marine red algae. J Phycol 42:1208–1216

Scolnik PA, Bartley GE (1994) Nucleotide sequence of an *Arabidopsis* cDNA for geranylgeranyl pyrophosphate synthase. Plant Physiol 104:1469–1470

Starr RC, Zeikus JA (1993) UTEX—the culture collection of algae at the University of Texas at Austin. J Phycol 29:1–106

Sun T-H, Liu C-Q, Hui Y-Y, Wu W-K, Zhou Z-G, Lu S (2010) Coordinated regulation of gene expression for carotenoid metabolism in *Chlamydomonas reinhardtii*. J Integr Plant Biol 52:868–878

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Thabet I, Guirimand G, Guihur A, Lanoue A, Courdavault V, Papon N, Bouzid S, Giglioli-Guivarc'h N, Simkin AJ, Clastre M (2012) Characterization and subcellular localization of geranylgeranyl diphosphate synthase from *Catharanthus roseus*. Mol Biol Rep 39: 3235–3243

van Schie CCN, Haring MA, Schuurink RC (2013) Prenyldiphosphate synthases and gibberellin biosynthesis. In: Bach TJ, Rhomer M (eds) Isoprenoid synthesis in plants and microorganisms: new concepts and experimental approaches. Springer, New York, pp 213–232

Vandermotern S, Haubruge E, Cusson M (2009) New insights into short-chain prenyltransferases: structural features, evolutionary history and potential for selective inhibition. Cell Mol Life Sci 66:3685–3695

Wang KC, Ohnuma S (2000) Isoprenyl diphosphate synthases. Biochim Biophys Acta 1529:33–48

Yang LE, Jin QP, Xiao Y, Xu P, Lu S (2013) Improved methods for basic molecular manipulation of the red alga *Porphyra umbilicalis* (Rhodophyta: Bangiales). J Appl Phycol 25:245–252

Yang LE, Huang XQ, Hang Y, Deng YY, Lu QQ, Lu S (2014) The P450-type carotene hydroxylase PuCHY1 from *Porphyra* suggested the evolution of carotenoid metabolism in red algae. J Integr Plant Biol 56:902–915