In silico analysis of phytoene synthase and its promoter reveals hints for regulation mechanisms of carotenogenesis in *Dunaliella bardawil*

Yong-Min Lao, Lan Xiao, Zhi-Wei Ye, Jian-Guo Jiang, and Shi-Shui Zhou

1College of Food and Bioengineering, South China University of Technology, Guangzhou, 510640 and 2School of Biological Science & Engineering, South China University of Technology, Guangzhou, 510006, China

ABSTRACT

Motivation: Previous researches showed that phytoene synthase (Psy) from *Dunaliella bardawil* is the first regulatory point in carotenogenesis. We hypothesize certain interactions between the environmental stress factors and the regulatory sequences of Psy in *D. bardawil* (DbPsy). Consequently, a PCR-based genomic walking approach was performed for isolation of psy promoter and terminator, respectively. The obtained nucleic acid sequences and the corresponding protein structure of DbPsy were analyzed and predicted using various bioinformatics tools. Finally, we presented some hints for the regulation mechanisms of DbPsy at the molecular level according to the computed results.

Results: LA PCR-based genomic walking results showed that the isolated sequences are the promoter and terminator of *psy*, correspondingly. Computational analysis demonstrated several candidate motifs of the promoter exhibiting hypothetic UV-B-, norglurzon- and salt-induced characteristics, as well as some typical domains universally discovered in promoter sequences, such as TATA-box, CCAAT-box and GATA-box, etc. Furthermore, the structure of Psy was also predicted and aligned along with many counterparts at the protein level. Low homology of N-terminus was found in *D. bardawil*, while a relatively conserved C-terminus was predicted to be involved in the catalytic activity and substrate recognition/binding. Phylogenic analysis classified the DbPsy into a cluster with other algae. These results implied that Psy may share similar regulation mechanisms among algae with respect to their C-termini; while the diversity in N-terminus among *Psy* species, along with the predicted inducible motifs in *psy* promoter from *D. bardawil*, may confer the fine tuning differences between *D. bardawil* and other algae.

Conclusion: By means of computer techniques, we found in *D. bardawil* that two interesting conserved motifs of *psy* promoter may involve in UV-B, norglurzon and salt regulation correspondingly; and that the diversity of Psy protein mainly lies in the N-termini among algae. These results indicate some hints for regulation mechanisms of carotenogenesis in *D. bardawil*.

Contact: jgjiang@scut.edu.cn

Received on April 12, 2011; revised on June 2, 2011; accepted on June 12, 2011

© The Author 2011. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com
reactions leading to the formation of β-carotene. The first committed step is a head-to-head condensation of two geranylated diphosphate (GGPP) molecules to produce phytoene by phytoene synthase (Psy) (Salvin et al., 2005). Subsequently, phytoene is converted to lycopene through four desaturation reactions, and β-carotene is finally formed by two cyclization steps. 

We hypothesize certain interactions between the environmental stress factors and the regulatory sequences of Psy in D. bardawil (DbPsy), more specifically, the conserved motifs encompassing the psy gene and the resulting diversity of protein structure caused by the variation of amino acid sequence in comparison to other species. Therefore, we isolated and cloned the promoter and terminator of psy and, then analyzed the obtained nucleic acid sequences and predicted the corresponding protein structure of DbPsy using various bioinformatics tools. Finally, we presented some hints for the regulation mechanisms of DbPsy at the molecular level according to the computed results.

2 METHODS

2.1 Strains and culture conditions

Dunaliella bardawil cells, obtained from the Institute of Hydrobiology, Chinese Academy of Science, were grown in defined medium (Sheffer and Avron, 1986) 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. Escherichia coli TOP10 was used as the host for the multiplication of plasmids.

2.2 Known sequence verification

For validation of the psy sequence deposited at NCBI (GenBank: EU328288.1), two pairs of overlapping primers (Table 1) were designed according to the Psy mRNA sequence (GenBank: EU328287.1).

Total RNA was extracted from 6 ml of D. bardawil cells grown at the late log phase using E.Z.N.A. Total RNA Kit II (OMEGA) following conditions recommended by the manufacturer. First strand cDNA was synthesized with Oligo dT-Adaptor Primer in a total volume of 10 μl, according to the Psy mRNA sequence (GenBank: EU328287.1). First strand cDNA was synthesized using an RNA PCR Kit (AMV) ver. 3.0 (TaKaRa) and then analyzed the obtained nucleic acid sequences and predicted the corresponding protein structure of DbPsy using various bioinformatics tools. Finally, we presented some hints for the regulation mechanisms of DbPsy at the molecular level according to the computed results.

2.3 Genomic walking for isolation of psy promoter and terminator

Based on the sequences of the psy gene of D. bardawil (GenBank: EU328288.1), three gene-specific primers for cloning of psy promoter, pSP1, pSP2 and pSP3, were designed (Table 1). Likewise, the primers for isolation of psy terminator, tSP1, tSP2 and tSP3, are also shown in Table 1. Genomic DNA of D. bardawil was extracted following the method described by Yang et al. (2000). Using Genomic Walking Kit

Table 1. Primers used in this study (5′→3′)

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Primer</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>Dbpsy-F</td>
<td>ATGGCAGACGCCGAAACAGCAAATCTT</td>
</tr>
<tr>
<td></td>
<td>Dbpsy-inR</td>
<td>GCCCTGAGGCGGTCATTGTA</td>
</tr>
<tr>
<td></td>
<td>Dbpsy-inF</td>
<td>GCACTCAAGAAGTACGCTCTAG</td>
</tr>
<tr>
<td></td>
<td>Dbpsy-R</td>
<td>TTAATTTGCTCTTGGGCACCAAG</td>
</tr>
<tr>
<td>Genomic walking</td>
<td>pSP1</td>
<td>CGAGGTAGAAGGTCTGGCAG</td>
</tr>
<tr>
<td></td>
<td>pSP2</td>
<td>GGCACTAGTGGAGCGGTCCTT</td>
</tr>
<tr>
<td></td>
<td>pSP3</td>
<td>CTGCTCTGCGACTACCTCCTC</td>
</tr>
<tr>
<td></td>
<td>tSP1</td>
<td>AATGCCCAAAAAAGGGAAGAAGG</td>
</tr>
<tr>
<td></td>
<td>tSP2</td>
<td>CAAACTACCTTGGGTGATTC</td>
</tr>
<tr>
<td></td>
<td>tSP3</td>
<td>AGACTGGCTGCGGCTTTTGC</td>
</tr>
<tr>
<td>Sequence validation</td>
<td>Ptest For</td>
<td>ACAATACTGCAACCAATACCC</td>
</tr>
<tr>
<td></td>
<td>Test Rev</td>
<td>CGACCGATGACTGATTACCC</td>
</tr>
</tbody>
</table>

Dbpsy-inR and Dbpsy-inF are overlapping primers. Complement sequences of the DbPsy cDNA are in bold and italic.

(TaKaRa). LA PCR-based genomic walking was performed to obtain the psy promoter and terminator, respectively. The primary nested PCR products were diluted to 1:100 with distilled water for subsequent nested PCR reactions; all manipulations were corresponding with the manufacturer’s protocol. Validation of putative psy promoter and terminator were performed by using primers Ptest For (Table 1) and pSP3 or primers tSP3 and Test Rev (Table 1), respectively. Sequence analysis was performed using Blat Software (http://blat.ncbi.nlm.nih.gov/). Protein predictions software (http://plantpan.mcb.nctu.edu.tw/).


3 RESULTS

3.1 Isolation and sequence analysis of the full-length DbPsy cDNA

The Psy template deposited at NCBI comes from direct submission without any support of report hitherto, which necessitates the validation work before genomic walking. To this end, RT-PCR
Three candidate promoter fragments of the third nested PCR were respectively, corresponding to pSP3 and tSP3, respectively. Two segments, length in 3458 bp and 1099 bp, promoter and terminator using testing primers Ptest For and Test Rev with fragment shares 184 bp identical nucleotides with the 3′ end of the shorter one. Likewise, a putative terminator of 1100 bp recovered, with an 184 bp region identical to the downstream region of 3′-end of pSP3 and tSP3, respectively. M, 500 bp DNA marker. (Fig. 1c). The AGE and sequencing results further indicated that the isolated promoter possesses many conserved motifs of plant (Fig. 2a), e.g. GATA box at position +149, CCAAT boxes widely found in eukaryotic genes, CGCG boxes with conserved sequence TGTCTC at position +121. Accordingly, some regulatory elements were also found, such as AFRAT involving in IAA regulation with consensus sequence TGTCCT at position +341 (Goda et al., 2004), the light-regulated elements G box (position −28) and ASF-1 binding site (position +102), etc. Interestingly, two remarkable motifs were also found in the promoter: box-L-like core sequences (position −491, also named BOXLCOREDCPAL, in Fig. 2a) and GT-1 motifs (position −85 and −2606). Previous researches has implied a valuable application of UV irradiation for accumulating carotenoids (Mogedas et al., 2009; Salguero et al., 2009; Merino, 2005) failed to detect any canonical poly (A) signal in the hypothetic terminator. While a reverse translation attenuator might function in D. bardawil, the pyl terminator may possess no canonical poly (A) signal.

Fig. 1. Isolation of promoter and terminator of pyl gene in D. bardawil.
(a) Validation of pyl template. M, 100 bp DNA marker; lane 1 and 2, Psy cDNA fragments with 21 bp overlapped. (b) Promoter isolation by LA-PCR. Figure 2a, for lanes at 370 nt upstream of the initiator codon ATG. The shortest one with 689 bp nucleotides did not have any homology with the upstream of pyl, while the 3763 bp fragment was supposed probably to be the 5′-UTR of pyl. M, DL5000 DNA marker; lane 1–3, products of 1st nested PCR, 2nd nested PCR and 3rd nested PCR, correspondingly. (c) Isolated fragment assumed to be the terminator of pyl. The brightest fragment of 1100 bp recovered, with an 184 bp region identical to the downstream of pyl, was inferred to be the pyl terminator. M, DL2000 DNA marker; lane 1–3, products of 1st nested PCR, 2nd nested PCR and 3rd nested PCR, respectively; M, 500 bp DNA marker. (d) Validation of hypothetic pyl promoter and terminator using testing primers Pest For and Test Rev with pSP3 and tSP3, respectively. Two segments, length in 3458 bp and 1099 bp, respectively, corresponding to pyl promoter and terminator, were obtained.

3.2 Isolation of the promoter and terminator of DbPsy
Three candidate promoter fragments of the third nested PCR were purified and sequenced (Fig. 1b). Sequence analysis by BlastN found that both the 3763 bp and 2975 bp fragments possess identical regions (130 bp in length) as expected with the 5′-end of pyl. Comparison of the sequences showed that the 3763 bp fragment is the extension of the shorter one. Likewise, a putative terminator fragment shares 184 bp identical nucleotides with the 3′-end of pyl as intended (Fig. 1c). The AGE and sequencing results further validate the hypothetic promoter and terminator (Fig. 1d). These results suggested that the obtained sequences are the upstream and downstream UTR of pyl, correspondingly.

3.3 Computational analysis of promoter and terminator of DbPsy
BlastN search found that the obtained sequences have great homology with the 5′- and 3′-end of pyl gene, respectively. CorePromoter detection suggested that the TSS, ‘A’ marked in Figure 2a, located at 370 nt upstream of the initiator codon ATG. Five species (Arabidopsis, Maize, Rice, Tomato and other) were selected when conducting transcriptional factor binding sites search at PlantPAN server (Chang et al., 2008). The retrieval results showed that the isolated promoter possesses many conserved motifs of plant (Fig. 2a), e.g. GATA box at position +149, CCAAT boxes widely found in eukaryotic genes, CGCG boxes with conserved sequences CGCGGC and ACGCGC at position −344 and +121 correspondingly. In addition, some regulatory elements were also found, such as AFRAT involving in IAA regulation with consensus sequence TGTCCT at position +341 (Goda et al., 2004), the light-regulated elements G box (position −28) and ASF-1 binding site (position +102), etc. Interestingly, two remarkable motifs were also found in the promoter: box-L-like core sequences (position −491, also named BOXLCOREDCPAL, in Fig. 2a) and GT-1 motifs (position −85 and −2606). Previous researches has implied a valuable application of UV irradiation for accumulating carotenoids (Mogedas et al., 2009; Salguero et al., 2009; White and Jahnke, 2002; Xue et al., 2005). Maeda et al. (2005) showed that BOXLCOREDCPAL involved in UV-B irradiation regulation in Daucus carota, therefore, such motif in D. bardawil might function in a similar way. The GT-1 box (also called GTIMOTIFCAM4, Fig. 2a) might contribute to pathogen- and NaCl-induced expression of the promoter (Park et al., 2004). Both the Signal Miner and RibEx programs (Abreu-Goeder and Merino, 2005) failed to detect any canonical poly (A) signal in the hypothetic terminator. While a reverse translation attenuator inside the last intron of pyl was identified by RibEx at 322 bp upstream of the TAA terminator codon, as is shown in Figure 2c. Such disappointing results prompted us to clone and analyze pyl terminator sequence again, resulting in the same outcome (data not shown). These results suggest that we may fail to isolate the genuine terminator of pyl in D. bardawil, or the pyl terminator may possess no canonical poly (A) signal.

3.4 Analysis of the genomic structure of pyl
To further validate and elucidate the genomic structure of pyl, NCBI Splign (Kapustin et al., 2008) and GENSCAN (Burge and Karlin, 1997) were performed. Both servers returned a consistent result that...
Table 2. Components of psy gene in D. bardawil

<table>
<thead>
<tr>
<th></th>
<th>Start</th>
<th>End</th>
<th>Length (bp)</th>
<th>GC content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>555</td>
<td>555</td>
<td>60.54</td>
</tr>
<tr>
<td>2</td>
<td>810</td>
<td>1070</td>
<td>261</td>
<td>55.17</td>
</tr>
<tr>
<td>3</td>
<td>2090</td>
<td>2245</td>
<td>156</td>
<td>56.41</td>
</tr>
<tr>
<td>4</td>
<td>2614</td>
<td>2765</td>
<td>152</td>
<td>55.92</td>
</tr>
<tr>
<td>5</td>
<td>3150</td>
<td>3300</td>
<td>151</td>
<td>49.67</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>1275</td>
<td>52.85</td>
</tr>
<tr>
<td>Intron</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>556</td>
<td>809</td>
<td>254</td>
<td>49.61</td>
</tr>
<tr>
<td>2</td>
<td>1071</td>
<td>2089</td>
<td>1019</td>
<td>47.2</td>
</tr>
<tr>
<td>3</td>
<td>2246</td>
<td>2613</td>
<td>368</td>
<td>53.8</td>
</tr>
<tr>
<td>4</td>
<td>2766</td>
<td>3149</td>
<td>384</td>
<td>54.95</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>2025</td>
<td>50.33</td>
</tr>
<tr>
<td>mRNA</td>
<td>1</td>
<td>1275</td>
<td>1275</td>
<td>47.15</td>
</tr>
</tbody>
</table>

Psy contains six exons separated by five introns (Fig. 2b). As shown in Table 2, the GC content (%) of all six exons varies from 49.67% to 60.54%, whereas the GC content (%) of the five introns ranges from 47.2% to 54.95%. Although the introns of psy are flanked by conventional 5′ splice donor GT and 3′ splice acceptor AG, they appear to lack of A+T-rich region, which characterizes a canonical intron.

3.5 Physical and chemical characteristics of DbPsy

The 424 amino acid peptide possesses a computed molecular weight of 48.30 kDa, a theoretical isoelectric point (pI) of 9.16, and exhibits an aliphatic index of 76.65. It contains 46 negatively charged residues (Asp + Glu) and 57 positively charged residues (Arg + Lys) with several hydrophilic and hydrophobic regions, as determined by ExPaSy (Gasteiger et al., 2003). In addition, it is classified as an unstable protein with a considered methionine N-terminus. WoLF PSORT showed that DbPsy may be situated in the chloroplast (Horton et al., 2007).

3.6 Conserved domain and motifs in DbPsy

Through the NCBI Conserved Domain Search (Marchler-Bauer et al., 2009), the deduced amino acid sequence was specifically hit by the trans-Isoprenyl Diphosphate Synthases (Trans_IPPS) which included squalene and phytoene synthases. The prediction showed that DbPsy can be classified into the superfamily of IPPS and Class I terpene cyclases. Six regions together forming Trans_IPPS were also predicted in DbPsy (Fig. 3): substrate binding pocket, Mg2+ binding site, active site lid residues, catalytic residues and two aspartate-rich sites (171-DELVD-175, 297-DVGED-301). In rat, the catalytic site consists of a large central cavity formed by mostly antiparallel alpha helices with two aspartate-rich regions (DXXXD) (Gu et al., 1998). The catalytic residues involved in the reactions of forming squalene. However, unlike squalene synthase, Psy does not require NADPH for its activity (Gu et al., 1998). The two aspartate-rich motifs, along with Tyr and Phe rich in the rest conserved regions, may involve in condensation reactions in DbPsy (Tansey and Shechter, 2000).

3.7 The advanced structures of DbPsy

The secondary structure prediction run at the PSIPred server (McGuffin et al., 2000) showed that the DbPsy consists of 72.64% α-helix (20 helices; 308 residues), 1.42% extended-beta (3 strands; 6 residues) and 25.94% random coil (23 coils; 110 residues) configurations (Fig. 4). MEMSAT3 and MEMSAT-SVM program detected a transmembrane region between amino acids 250 and 265. The carboxy-terminus is possible in the cytoplasm; while the amino-terminus, followed by a 37 amino acid signal peptide, locates in the extracellular fluid (Fig. 4).

Moreover, the tertiary structure of DbPsy was also constructed by 3D-JIGSAW (Contreras-Moreira and Bates, 2002). The calculated result shown in Figure 5 revealed that the single polypeptide chain model covers 322 residues from 103 to 424 amino acids, which is composed of 209 H-bonds, 20 helix elements and 32 turns. However, unlike the secondary structure retrieved from PSIPred server, there is no strand element predicted using the 3D-JIGSAW server.

3.8 Homologous alignment and phylogenetic analysis of DbPsy

The gene family alignment of amino acid sequences obtained by ClustalX analysis revealed that the overall structure of the
Regulation mechanisms of carotenogenesis in Dunaliella bardawil

Fig. 5. The tertiary structure of DbPsy. Comparative modeling was performed using 3D-JIGSAW. The structure was presented as ribbons (the secondary structure elements) and wireframe using RasMol version 2.7.2.1. The α-helix and β-turn regions of the putative protein were indicated with cylinders (red) and bands (blue), respectively. The two aspartate-rich motifs were colored by yellow, and the distinctive Asp residues (171, 175, 297 and 301) were labeled and highlighted as spacefill. The N-terminus and C-terminus were also shown.

Fig. 6. The hypothetic Trans_IPPS_HH domain found in DbPsy. Comparison of deduced amino acid sequence of DbPsy with its homologs indicated a functional conserved domain Trans_IPPS_HH, which lays in the C-terminus from 132 to 424 amino acids. The active site lid residues (141 aa to 145 aa) are marked with red line above; the highly conserved catalytic residue Try164 is marked with red asterisk; two aspartate-rich regions are boxed in blue. Gaps introduced to maximize sequence homology are indicated by dashes. Amino acid residues identical in all sequences are highlighted in white type on a black background; less conserved residues appear on a light gray background.

Trans_IPPS_HH domains has been conserved from bacteria to higher plant (Figs 3 and 6), though all of the homologs aligned varied considerably in their N-termini (data not shown). The functional conserved domain may cover from 132 to 424 amino acids.

The BlastP search demonstrated that DbPsy shares the highest homology with *D.salina* at the protein level (79% identity and 87% similarity), 68% identity and 80% similarity with the *Haematococcus pluvialis* counterpart, 69% and 81% with the *Chlamydomonas reinhardtii* counterpart, 55% and 69% with the *Arabidopsis thaliana* counterpart. A lower homology of DbPsy was found with the prokaryote counterparts, with the exception of cyanobacterial Psys (40–42% identity).

The phylogenetic tree was constructed using Neighbor Joining method by MEGA 4.0.2 software (Tamura et al., 2007) (Fig. 7). Three defined clusters can be found: Clusters I, II and III. Cluster I contains Psys from higher plants, cluster II consists of the algae Psys and cluster III comprises the bacterial homologs. Such phylogenetic dendrogram showed similar patterns with BlastP search and DNA sequence alignment results (data not shown); which indicated that Psys isolated from green algae can be classified into one group (Fig. 7). Accordingly, these results indicated an evolutionary relationship among green algae.

4 DISCUSSION

*Dunaliella bardawil* and *D.salina* are recognized as the most promising natural sources to produce β-carotene in biotechnology commercial-production applications, due to their ability of massive accumulation of β-carotene under stress conditions, such as high light intensity and salinity, stress temperature, privative nutrients, etc. Plenty of studies have been performed to investigate the carotenoid biosynthesis pathway for the ultimate goal to produce as much β-carotene as possible (Megdas et al., 2009; Salguero et al., 2005; Xue et al., 2005). Nevertheless, most of these researches mainly focused on culture conditions and extraction approaches, or other aspects except carotenogenic enzyme genes. So far, a small number of enzymes directly involving in carotenoid biosynthetic pathway have been studied in *Dunaliella*, especially in *D.bardawil* (Yan et al., 2005; Zhu et al., 2005). Psy is one of the few which were cloned in *D.salina* and *D.bardawil*. However, it has not yet been experienced deeply functional validation up to now in *D.bardawil*. The attempts to obtain psy regulatory elements from *D.bardawil* would facilitate interpretation of the regulation mechanisms for Psy expression and accumulation of β-carotene. Additionally, works on elaborating the Psy structure at the protein level would help further presentation of the differentiations of Psy regulation even the
whole carotenoid metabolic pathway between species. We strongly speculate, based on previous intensive efforts (Ben-Amotz et al., 1987; León et al., 2005; White and Jahnke, 2002), that regulation of Psy and carotenogenesis is fulfilled at multiple levels from DNA structure to post-translational control. More specifically, the regulatory elements of psy gene and its protein structure may confer different regulation mechanisms in D.bardawil from other algae.

As shown previously, massive accumulation of β-carotene is triggered by environmental stress (Mogedas et al., 2009; White and Jahnke, 2002; Ye et al., 2008), implying a regulatory expression pattern of many carotenogenic genes. Among these environmental stresses, UV interference executes different effects on β-carotene accumulation through photosynthetic reactions (Mogedas et al., 2009; White and Jahnke, 2002). One of the protective mechanisms against UV-induced photosynthetic damage involves in massive accumulation of β-carotene in D.bardawil exposed to UV-A rather than UV-B, while D.salina lacks such ability (White and Jahnke, 2002). UV-B even executes contrasting effects on β-carotene biosynthesis: decreases β-carotene but enriches other carotenoids including phytoene. However, this UV-A-induced protection is lost in cells treated with the bleaching herbicide norflurazon (Ben-Amotz et al., 1987). In such cells phytocenese synthesis is enhanced significantly, with a decreased level of β-carotene, so that León et al. (2005) applied this bleaching herbicide to product phytoene in D.bardawil. Our study found that the promoter of DbPsy possesses a conserved motif BOXLCOREDPCPAL (Fig. 3) homologous with DcPAL1 counterpart in D. carota. Such motif participates in up-regulation responses of DcPAL1 to UV-B irradiation in D. carota mediated by DcMYB1 (Maeda et al., 2005). These proofs suggest that, in accordance with the investigations discussed above (Ben-Amotz et al., 1987; Mogedas et al., 2009; White and Jahnke, 2002), DbPsy may be regulated by UV-B and norflurazon with respect to similar photobleaching effects to accumulate phytoene, rather than β-carotene, through its promoter region. We conclude that UV-B and norflurzon may induce photobleaching through the BOXLCOREDPCPAL situated within the promoter of DbPsy.

Moreover, a salt- and light-induced pattern of β-carotene should be concerned with certain regulatory domains in carotenogenic genes. Here, we found two GT-1-like boxes deposited inside the isolated promoter, which were widely discovered in many light-regulated genes (Terzaghi and Cashmore, 1995). With respect to GT1CONSENSUS (GAAAAA, position −354), we speculate an intermediary role on interaction between TFIIA and GT-1-like factors in D.bardawil (Terzaghi and Cashmore, 1995). TFIIA indirectly interacts with GT-1-like factors through GT1CONSENSUS to accumulate β-carotene induced by light. Another GT-1-like box (Fig. 3, GT1GMSLM4) involves in NaCl-induced up-regulated expression of SCaM-4 in Glycine max (Park et al., 2004), consequently, GT1GSCM4 found in DbPsy promoter may fulfill similar function as its counterpart in Glycine max: up-regulates its host gene psy in order to enhance β-carotene synthesis. UV-B even executes contrasting effects on UV-A and UV-B exclusively in D.bardawil (White and Jahnke, 2002) suggest UV-B may also up-regulate the expression of Psy through these motifs. Such unique characteristics imply fine tuning of Psy (even the whole carotenogenesis) related to diversity of DNA-protein interactions between D.bardawil and other algae.

Psy has been regarded as the first key enzyme to catalyze a head to head combination reaction, generating colorless phytoene in carotenoid biosynthesis. It is supposed to be the regulatory point controlling the flux of carbon source towards carotenoids (Shewmaker et al., 1999). To further illuminate the regulation mechanisms of DbPsy, we performed a series of in silico analysis at the protein level. Advanced protein structure analysis manifests a relatively conserved C-terminus containing putative substrate binding and catalytic domains from around 100 to 400 amino acids. Phylogenetic analysis indicated that Psyss isolated from algae occupy one cluster (Fig. 7). These clues imply that they may share some similar regulation mechanisms distinctively with regard to the C-terminus of Psy in algae. Furthermore, alignments of the deduced amino acid sequence showed the sequence diversity...
between species primarily in the N-terminus. Tran et al. (2009) suggested the N-terminal regions, unessential to the enzymatic function, uniquely account for the major differences of the two Psy classes found in some algae. Moreover, the N-terminal region of plant Psy is generally longer than bacterial enzymes (Kim et al., 2003). This region is crucial for protein targeting and dimerization that are critical for enzyme stability and activity (Cunningham and Gantt, 1998; Sun et al., 1996). The N-terminus diversity between D. bardawil and other algae may affect Psy turnover and catalytic activity in certain similar patterns. In addition, increased level of Psy does not directly correlate with increased carotenoids content (Kim et al., 2003); the N-terminal variety may render different capabilities between different Psy of inter and intraspecies with regard to protein targeting, processing or assembly into a fully functional complex. As discussed above, we speculate the sequence differences of Psy at the protein and DNA levels may be responsible for the fine tuning differences between D. bardawil and D. salina, we also conducted alignments between these two algae, which returned similar diversity of N-terminus (or the 5′-end of the Psy mRNA CDS) and conserved C-terminus (or the 3′-end of the Psy mRNA CDS) (data not shown). Accordingly, we speculate that the fine tuning mechanisms of Psys between algae may lie in the variation of their N-terminus. Studies on the N-terminal diversity as well as other differential regions between species would be helpful to clarify the distinct regulation mechanisms for Psy in D. bardawil.

An in-depth understanding of the regulation mechanisms of DbPsy still requires further intensive investigations on the promoter function involving in the putative motifs discussed above and other related conserved sequences. Moreover, the identification of psy promoter in D. salina and subsequent comparison with the one in D. bardawil will shed light on the fine tuning differences between the two algae at the DNA level. Besides, the protein structure is also a key point to aid the presentation of these mechanisms, especially its N-terminus. Such understanding would also contribute to bioengineering using DbPsy promoter to produce foreign proteins.

**Funding:** National Natural Foundation of China (grant 30870025).

**Conflict of Interest:** none declared.

**REFERENCES**


