Evidence for positive selection in phycoerythrin genes of red algae and cyanobacteria *Prochlorococcus* and *Synechococcus*

S. QIN[*,***], F.Q. ZHAO[*,**], and C.K. TSENG[*]

Institute of Oceanology, Chinese Academy of Sciences[*], Graduate School of Chinese Academy of Sciences[**]
7 Nanhai Road, Qingdao, 266071, P.R. China

Abstract

Because of the shortage of phycoerythrin (PE) gene sequences from rhodophytes, *peBA* encoding β- and α-subunits of PE from three species of red algae (*Ceramium boydenn*, *Halymenia sinensis,* and *Plocamium telfariae*) were cloned and sequenced. Different selection forces have affected the evolution of PE lineages. 8.9 % of the codons were subject to positive selection within the PE lineages (excluding high-irradiance adapted *Prochlorococcus*). More than 40 % of the sites may be under positive selection, and nearly 20 % sites are weakly constraint sites in high-irradiance adapted *Prochlorococcus*. Sites most likely undergoing positive selection were found in the chromophore binding domains, suggesting that these sites have played important roles in environmental adaptation during PE diversification. Moreover, the heterogeneous distribution of positively selected sites along the PE gene was revealed from the comparison of low-irradiance adapted *Prochlorococcus* and marine *Synechococcus*, which firmly suggests that evolutionary patterns of PEs in these two lineages are significantly different.

Additional key words: *Ceramium boydenn*; *Halymenia sinensis*; molecular evolution; phylogenetic analysis; *Plocamium telfariae*; positive selection; *Synechococcus*.

Introduction

Phycobiliproteins are light-harvesting antenna proteins found in cyanobacteria and certain eukaryotic algae belonging to Rhodophyta, Cryptophyta, and Glaucophyta (Zuber 1986, Glazer 1988, Bryant 1991, MacColl 1998). Phycoerythrin (PE) is one kind of phycobiliproteins which is found in all rhodophytes, some species of cyanobacteria and cryptophytes, and *Prochlorococcus* populations (Hess *et al.* 1996). Wilbanks *et al.* (1991) first reported the second type of PE (PE II) derived from the marine cyanobacterium *Synechococcus* sp. WH8020, which carries six bilin chromophores per α,β-unit. Normally, fresh water and terrestrial PEs (PE I) carry only five bilins per α,β-unit. A novel type of PE (PE III) was found in *Prochlorococcus* CCMP1375 (Hess *et al.* 1996), which utilizes chlorophylls *a* and *b* as major photosynthetic pigments. In the high-irradiance (HI) adapted *Prochlorococcus*, PE (HI-PE) mutated at multiple sites and had lost two of the four conserved cysteines for chromophore binding (Steglich *et al.* 2003b). Ting *et al.* (2001) suggest that within the *Prochlorococcus* lineage the selective forces affecting the evolution of the PE gene have not been uniform and PE genes show genetic heterogeneities between *Prochlorococcus* and *Synechococcus* lineages.

The specific goals of this study were to identify whether PE has been subject to positive selection in the process of evolution and to use recently developed likelihood methods to identify specific positively selected sites. By comparing the sites for positive selection between different lineages, it should be possible to gain additional insight into the potential selective mechanisms in PE evolution. The prefixes C-, B-, R- of PEs discussed below only designate the type of source organism, *i.e.* Cyanophyta, Bangiaceae, and Rhodophyceae (or Florideophyceae), respectively.

Received 1 July 2004, *accepted* 11 October 2004.
[***]Corresponding author: Fax: +86-532-2880645, e-mail: sqin@ms.qdio.ac.cn; zhaofangqing@ms.qdio.ac.cn (F.Q. Zhao)
Acknowledgement: We acknowledge Mr. Peng Shi (Kunming Institute of Zoology, Chinese Academy of Sciences) for his comments and suggestions on the manuscript. This work was supported by grants from Key Innovative Project (KZCX3-SW-215) of the Chinese Academy of Sciences.

Materials and methods

DNA extraction, amplification, cloning, and sequencing: Three red algae species *Ceramium boydenn*, *Halymenia sinensis*, and *Plocamium telfariae* were collected from Shilaoren, Qingdao. 2–4 g fresh tissue was frozen with liquid nitrogen and ground to powder with pestle. DNA was extracted from the powdered tissue using a lysis buffer (100 mM Tris-HCl, pH 8.0, 20 mM EDTA, 1.4 M NaCl, 3 % trimethyl ammonium bromide, and 0.2 % (v/v) β-mercaptoethanol. After centrifugation, the top aqueous layer was collected and then mixed with phenol : chloroform : isoamyl alcohol (25 : 24 : 1). Finally, nucleic acids in the extracted aqueous phase were precipitated with two volumes of cold ethanol and 1/10 volume of 3 M Na-acetate at –20 °C for 1 h or overnight. PCR amplification of *peBA* used a forward primer (p1: 5'-A*GAATTC*AATGCTTGAYCW-3', an *Eco*R I enzyme site at 5' end is shown in italics) and a reverse primer (p2: 5'-CCGTTAGSDTARDGMRTTD-3') derived from the reported sequences (Roell and Morse 1993, Bernard *et al.* 1996, Sui and Zhang 2000). Amplification reactions were performed in a *PTC-150 MiniCycler* (*MJ Company*, USA). Reaction parameters were: 94 °C for 5 min, then 94 °C for 1 min→45–55 °C per 1 min→72 °C per 1.5 min for 35 cycles before performing a prolonged incubation at 72 °C for 10 min. PCR products were recovered using low-melting point agarose gel method, and then cloned into the pMD18-7 vector according to the manufacturer's instructions (*TAKARA*, Dalian, China). Plasmids with right inserts were prepared using standard method (Sambrook *et al.* 1989) before attempted sequencing double strands from both 5' and 3' ends.

Sequence alignment and phylogenetic analyses: The following PE genes were retrieved from the *GenBank*: Ane (*Aglaothamnion neglectum*, Z11907), Pbo (*Polysiphonia boldii*, Z14904), Sy1 (*Synechocystis* sp. BO8402, AAF89673), Sy2 (*Synechocystis* sp. PCC 9413, AAF99685), Sy3 (*Synechocystis* sp. PCC6701, P20778), Fdi (*Fremyella diplosiphon*, P05097), So2-1 (*Synechococcus* sp. WH7803, Q08086), So3-1 (*Synechococcus* sp. WH8102, ZP_00116179), So4-1 (*Synechococcus* sp. WH8020, Q02180), So1-2 (*Synechococcus* sp. WH8103, P37721), So3-2 (*Synechococcus* sp. WH8102, ZP_00116179), So4-2 (*Synechococcus* sp. WH8020, P27647), Med4 (*Prochlorococcus mar*. MED4, CAE18764), As (*Prochlorococcus* sp. AS9601, AJ304837), Pmi (*Prochlorococcus marinus* MIT 9313, CAE21857), Pcc (*Prochlorococcus* CCMP1375, NP_874731), Pac1 (*Prochlorococcus* sp. PAC1, CAB75589), and Pac2 (*Prochlorococcus* sp. PAC2, CAB82771). Med4 and As are referred to HI-adapted *Prochlorococcus* strains while Pmi, Pcc, Pac1, and Pac2 are referred to low-irradiance (LI) adapted *Prochlorococcus* strains.

Amino acid alignments of PE were carried out using *CLUSTAL W* software with the default settings (Thompson *et al.* 1994). The nucleotide sequences were then aligned following the same gap patterns. Further analyses were all performed on this set of aligned nucleotide sequences. The gaps in the alignment were excluded and only 165 codons were used in the latter calculation.

All phycobiliproteins are considered to be evolved from the same ancestor through gene duplication (Apt *et al.* 1995), so we use the phycocyanin gene from *Nostoc* sp. PCC7120 as an out-group to examine the relationships among different PEs. A phylogenetic tree made by the maximum parsimony method was constructed using *PROPARS* program from a *PHYLIP* software package (Felsenstein 1989). The reliability of the tree was evaluated by the bootstrap method (Felsenstein 1989) with 1 000 replications.

Estimation of d_N/d_S ratios and selected sites: The non-synonymous/synonymous substitution rate ratio ($\omega = d_N/d_S$) provides an effective tool to identify selective pressure at the protein level. The relative rate of non-synonymous of substitutions (d_N) to synonymous substitutions (d_S) provides an indication of the presence and type of selection acting on a protein-coding sequence. For the majority of proteins over the majority of evolutionary time, d_N/d_S is expected to be below one, indicating that a gene evolves with constraint on amino acid replacements and is therefore functional. On the other hand, a d_N/d_S value greater than one is commonly taken as evidence that the protein is undergoing diversification selection for increased amino acid diversity. Finally, a d_N/d_S equal to one is consistent with neutral evolution. This research uses the site-specific models (Nielsen and Yang 1998, Yang *et al.* 2000) to calculate ω detecting positive selection at individual sites. This method compares a null model with a more general one.

To reduce calculation time, two sequences from each group of PE were randomly chosen to reconstruct the phylogenetic trees for further calculation (Fig. 1*B*). In this study we compared model M0 (one-ratio) with model M3 (discrete) and model M7 (beta) with model M8 (beta&ω) (Nielsen and Yang 1998, Yang *et al.* 2000). However, this method only allows the ω ratio to vary among sites but not among different lineages. Consequently, we firstly performed the calculation with all lineages of PEs, and then certain lineage (HI-PEs, As, and Med4) was excluded to reduce the level of sequence divergence in the site-specific analyses. A new model, *i.e.* the branch-site model developed by Yang and Nielsen (2002) which allows the ω ratio to vary both among sites and among lineages was also used in this study to detect molecular adaptation along certain branches based on the phylogenetic tree. The computation was completed using the *PAML* software package version *3.14* (Yang 1997) and the posterior probability that a particular codon site is

positively selected can be estimated using the empirical Bayes' approach described by the software.

Results

Amplification of PE genes from the three species of red algae: The primers contained redundancies and had theoretical annealing temperature 45–55 °C. Optimal PCR conditions were found at a primer-template annealing temperature at 47 °C in amplifying *peBA* from *Ceramium boydenn*. A single DNA fragment with an approximate size expected for *peBA* was amplified (data not shown). Using the similar PCR conditions, *peBA* from genomic DNAs of *Halymenia sinensis* and *Plocamium telfariae* were amplified. Sequences have been deposited to *GenBank* (AF526383, AY502053, AY502054).

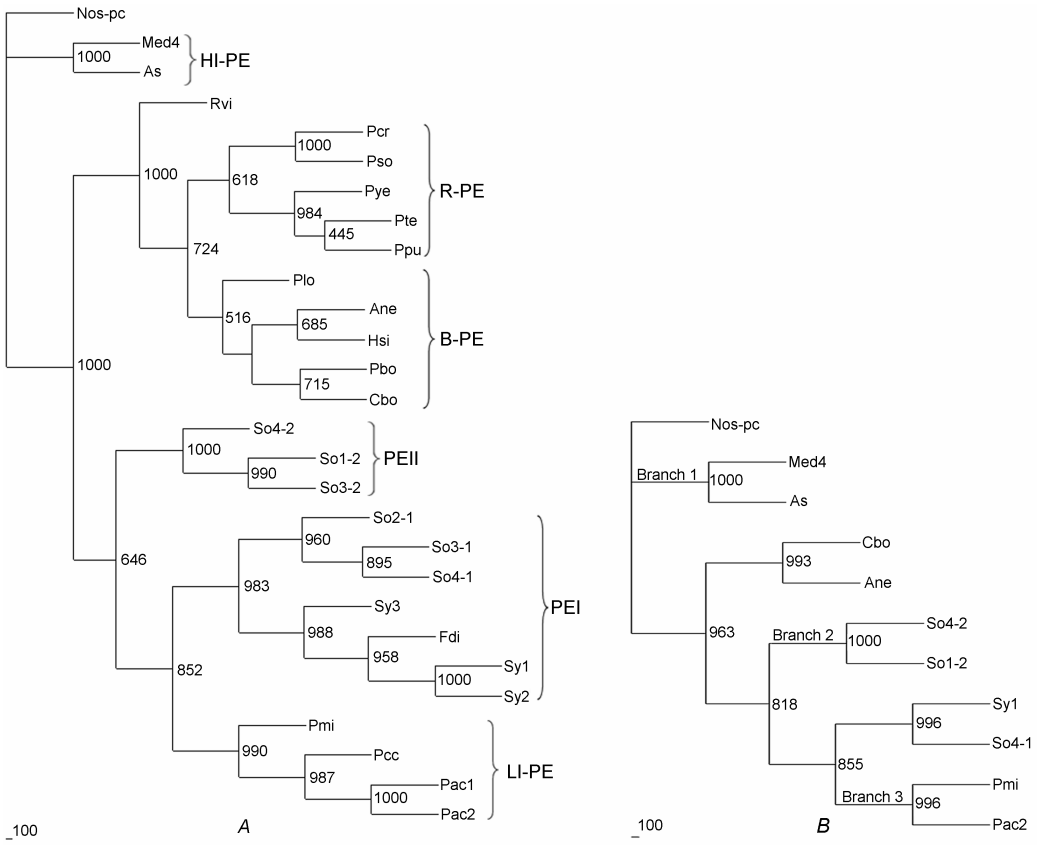

Fig. 1. Phylogenetic trees inferred from phycoerythrin amino acid sequences using maximum parsimony method and *Nostoc* sp. PCC 7120 (accession: NC_003272) as an out-group. *Numbers at the nodes* indicate bootstrap values. *Scale bars* represent level of sequence divergence. (*A*) The tree used for phylogenetic analyses. (*B*) The tree used for d_N/d_S calculation.

Multiple sequence alignment and phylogenetic analyses: Results of amino acid sequence alignment of PE genes (data not shown) are in accordance with former studies (Apt *et al.* 1995, Ting *et al.* 2001). Several residues known for their functional importance, such as residues related to chromophore attachment, are conserved in the alignment.

To examine the relationships among different types of PE genes, phylogenetic trees were constructed using amino acid sequences of phycocyanin genes from *Nostoc* sp. PCC7120 (Cyanophyta) as an out-group. Maximum parsimony method (Fig. 1*A*) and maximum likelihood method (data not shown) were used. The sub-topologies of the evolutionary trees obtained by these two methods were nearly identical and the topologies were reliable by the criteria of bootstrap. Maximum parsimony analyses of the 27 sequences identified two monophyletic groups with 100 % bootstrap values: one consisted of Med4 and As (HI-PE from HI-adapted *Prochlorococcus*), and the other includes C-PE I, C-PE II, LI-PE, and R (B)-PE. Within the second group, the high level of sequence conservation observed in comparisons among R (B)-PEs was reflected in the tight clustering in the phylogenetic tree (with 100 % bootstrap value). Obviously, PEs in red algae form a monophyletic family, which comprises two separate clades, *i.e*. Florideophyceae and Bangiophyceae (bootstrap support >70 %). Moreover, species in Porphyridiales, an order usually designated as the ancestor in red

algae (Gabrielson and Garbary 1986, 1987), disperse on different branches. Rvi strain of Porphyridiales (red algae) is basal to the red algae clade, while Pso and Pcr strains of the same order are in the more recently evolved branch (bootstrap support 100 %). This finding suggests that Porphyridiales should be polyphyletic, which is consistent with previous studies (Oliveira and Bhattacharya 2000).

Table 1. Parameter estimates for the phycoerythrin gene (branch 1). p is the number of free parameters for the ω distribution. Sites potentially under positive selection are identified using *Aglaothamnion neglectum* sequence as the reference.

Models	Branches	p	LnL	Estimates of parameters	Positively selected sites
Site-specific models	M0: one ratio	1	-3289.09	$\omega = 0.059$	None
	M3: discrete (K = 2)	3	-3225.00	$p_0 = 0.671$, ($p_1 = 0.329$) $\omega_0 = 0.016$, $\omega_1 = 0.171$	None
	M3: discrete (K = 3)	5	-3221.06	$p_0 = 0.156$, $p_1 = 0.578$, ($p_2 = 0.266$) $\omega_0 = 0.000$, $\omega_1 = 0.028$, $\omega_2 = 0.203$	None
	M7: beta	2	-3222.45	$p = 0.560$, $q = 7.106$	
	M8: beta&ω	4	-3222.45	$p_0 = 1.000$, $p = 0.560$, $q = 7.103$ ($p_1 = 0.000$), $\omega = 0.753$	None
Branch-site model B	Branch 1	5	-3219.28	$p_0 = 0.397$, $p_1 = 0.197$, ($p_3+p_4 = 0.406$) $\omega_0 = 0.016$, $\omega_1 = 0.167$, $\omega_2 = 5.164$	5F 86Y 90A 118T 147C (at $p>0.95$) 44C 65G 71R 101D 120S 157Y (at $p>0.90$)

Table 2. Parameter estimates for the phycoerythrin gene (branches 2 and 3). p is the number of free parameters for the ω distribution. Sites potentially under positive selection are identified using *Aglaothamnion neglectum* sequence as the reference.

Models	Branches	p	LnL	Estimates of parameters	Positively selected sites
Site-specific models	M0: one ratio	1	-3191.51	$\omega = 0.021$	None
	M3: discrete (K = 2)	3	-3155.05	$p_0 = 0.911$, ($p_1 = 0.089$) $\omega_0 = 0.012$, $\omega_1 = 21.08$	8V 11Q 26S 119N 144A 160K (at $p>0.99$) 43S 149A 163S (at $p>0.95$)
Branch-site model B	Branch 2	5	-3094.06	$p_0 = 0.623$, $p_1 = 0.244$, ($p_3+p_4 = 0.133$) $\omega_0 = 0.006$, $\omega_1 = 0.087$, $\omega_2 = 999.00$	54I 64G 66N 67C 72R (at $p>0.99$) 9V 15K 53M (at $p>0.90$)
	Branch 3	5	-3097.10	$p_0 = 0.608$, $p_1 = 0.264$, ($p_3+p_4 = 0.128$) $\omega_0 = 0.006$, $\omega_1 = 0.080$, $\omega_2 = 715.78$	35V 40S 120S 148S (at $p>0.99$) 23T 28G 32L 46V 47S 142Q (at $p>0.95$)

Positive selection of PE gene: Three branches that separate HI-PE (Med4 and As), PE II (So1-2, So4-2), and LI-PE (Pmi, Pac2) from other PE lineages were labelled in Fig. 1*B*. Two pairs of models: M0 (one-ratio) *vs*. M3 (discrete), and M7 (beta) *vs*. M8 (beta&ω) were used to estimate d_N/d_S ratios but results showed that none of the calculation results (Table 1) suggests the presence of a site with ω>1. Result of one-ratio model (M0) gave an estimate ω = 0.02-0.06 (Tables 1 and 2), which indicated that the PE gene family is under strong selective constraints in evolution. When HI-PEs were excluded from the test, result of model M3 (K = 2) suggested that 8.9 % of sites are under positive selection with ω = 21.08 (Table 2). Result of model M3 (K = 2) was significantly better than that of the model M0, with $2\Delta LnL = 2\times36.46 = 72.92$, $p<<1$ %, df = 2. Branch-site model is an extension to the M3 (discrete) model with K = 2 site classes (Yang and Nielsen 2002). Results using branch-site model showed that 40.6 % of sites are under positive selection along branch 1 (Table 1, $\omega_2 = 5.164$) and that $2\Delta LnL = 2\times5.72 = 11.44$, with $p<0.01$ and df = 2. Table 2 shows that about 13.3 % of the codons along branch 2 were subject to positive selection, with $2\Delta LnL = 2\times61 = 122$, and $p<<0.01$. The estimates of the proportion of neutral or weakly constraint sites and sites subject to strong constraint were 24.7 and 60.5 %, respectively. LI-PEs along branch 3 also exhibit strong positive selection among 12.8 % of sites, and four amino acid sites were identified at the 99 % cut-off in this research.

The distribution of the inferred positively selected sites (Fig. 2) shows that some codons along PE genes

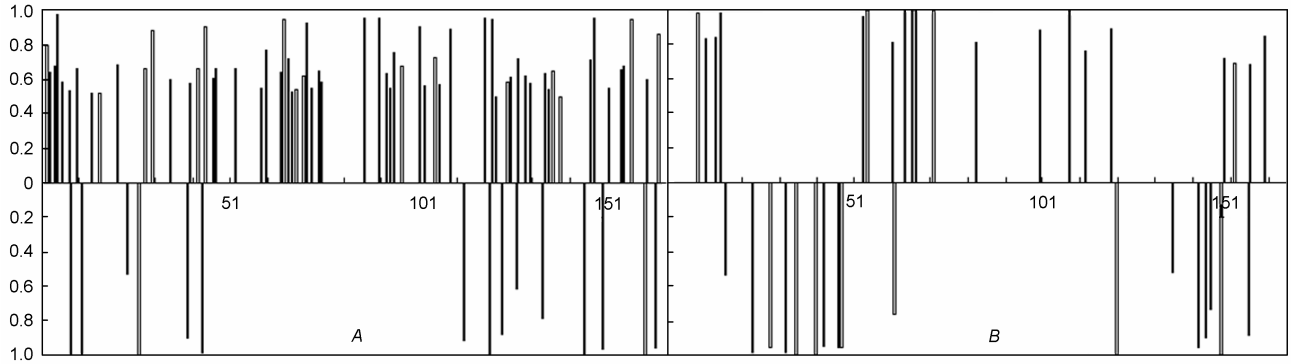

Fig. 2. Posterior probabilities (>50 %) for sites under positive selection. *X-axis* denotes position in the amino acid alignment. *Y-axis* denotes posterior probability of sites under positive selection. (*A*) Sites inferred from branch 1 (above X-axis) and sites inferred from site-specific model (below X-axis). (*B*) Sites inferred from branch 2 (above X-axis) and branch 3 (below X-axis).

have undertaken positive selection during evolution. These sites are mostly located in the chromophore binding domains. However, PE genes in HI-adapted *Prochlorococcus* lineage show a different evolutionary pattern with the other lineages, with 40.6 % of positively selected sites distributing more evenly along the whole sequence. Similarly, heterogeneous distribution of positively selected sites was found from the comparison of LI-adapted *Prochlorococcus* and marine *Synechococcus* (Fig. 2*B*), although several functionally important amino acids are still conserved in both lineages.

Discussion

The types of PEs used in this paper belong to five categories: PE I, PE II, HI-PE, LI-PE, and R-PE (from red algae). We demonstrated that specific sites of the PE gene have evolved rapidly under positive selection, and that PE genes derived from *Prochlorococcus* and marine *Synechococcus* showed different evolutionary patterns.

PE gene has been used as a sensitive molecular marker to identify the natural populations of *Prochlorococcus* (Steglich *et al.* 2003b), and some other phycobiliprotein genes (such as *cpcAB*) have also been used to investigate relationships between closely related picocyanobacteria isolates (Robertson *et al.* 2001, Janson and Graneli 2002). In our studies, PEs from red algae form a monophyletic family, and comprise two separate clades, Florideophyceae and Bangiophyceae, with the bootstrap support value >70 %. In addition, the nucleotide sequences of PE, which are more divergent than their amino acid sequences, can be also used to evaluate the phylogenetic relationships. The above observations on the phylogeny of *peBA*, as more information is gathered, will contribute to systematics of red algae.

Our results firmly support the inference that the selective pressures affecting the evolution of PE genes have not been uniform in *Prochlorococcus* strains, as was put forward by Ting *et al.* (2001). As shown in Table 1, more than 40 % of sites may be under positive selection, and nearly 20 % are weakly constraint sites in HI-adapted *Prochlorococcus* strains. PE genes in LI-adapted *Prochlorococcus*, however, are still under strong purified selection with 12.8 % of positively selected sites.

Some marine *Synechococcus* species contain PE II besides PE I. Table 2 shows that about 13.3 % of the PE II codons are detected by d_N/d_S analyses, which indicates that positive selection has occurred along the marine *Synechococcus* lineage right after PE gene duplication. The newly replicated PE in marine *Synechococcus* can be strongly adapted for the absorption of blue-green radiation, with the combination of phycourobilin (PUB) chromophores (Wilbanks *et al.* 1991). The PUB-attached peptide is located from site-40 to site-70 along the alignment as suggested by Swanson *et al.* (1991). Our results (Table 2) suggest that five amino acids (53M, 54I, 64G, 66N, and 67C), adjacent to PUB binding site, are under strong positive selection in PE II. These possible positively selected sites may be related to environmental adaptation. Interestingly, the distributions of possible positively selected sites between marine *Synechococcus* and *Prochlorococcus* are different. A possible reason may be that PE genes in these two lineages evolve in two different ways because of the diverse selection forces. Marine *Synechococcus* lineages modified their phycobiliproteins for environmental adaptation, with the duplication of PE genes and the presence of R-phycocyanin. *Prochlorococcus*, however, possesses chlorophyll *b* as the principal light-harvesting pigment and lost most of its phycobiliproteins (Hess *et al.* 1996). Although experimental evidence suggests that PE in *Prochlorococcus marinus* CCMP1375 could transfer energy to chlorophylls (Lokstein *et al.* 1999), the effective contribution to total light-harvesting capacity was less than 1.8 % and its accurate role remains unknown (Steglich *et al.* 2003a). We argue that adaptive changes in different lineages of PE

have been important for their functional divergence and the selection forces exerted on chromophore binding domains should be mostly irradiation-related factors, such as irradiance and quality of radiation.

Our results revealed that 8.9 % of the codons are subject to positive selection within the PE lineages (excluding HI-adapted *Prochlorococcus*). The possible selected sites and the different evolutionary patterns found in this research offer us clues to research for the molecular basis of adaptation of organisms to their specific environments.

References

Apt, K.E., Collier, J.L., Grossman, A.R.: Evolution of the phycobiliproteins. – J. mol. Biol. **248**: 79-96, 1995.

Bernard, C., Etienne, A., Thomas, J.C.: Synthesis and binding of phycoerythrin and its associated linkers to the phycobilisomes in *Rhodella violacea* (Rhodophyta): compared effects of high light and translation inhibitors. – J. Phycol. **32**: 265-271, 1996.

Bryant, D.A.: Cyanobacterial phycobilisomes: Progress toward complete structural and functional analysis *via* molecular genetics. – In: Bogorad, L., Vasil, K. (ed.): The Photosynthetic Apparatus: Molecular Biology and Operation. Pp. 257-300. Academic Press, San Diego – New York – Boston – London – Sydney – Tokyo – Toronto 1991.

Felsenstein, J.: Phylogeny inference package (version 3.2) – Cladistics **5**: 164-166, 1989.

Gabrielson, P.W., Garbary, D.J.: Systematics of red algae (Rhodophyta). – CRC crit. Rev. Plant Sci. **3**: 325-366, 1986.

Gabrielson, P.W., Garbary, D.J.: A cladistic analysis of Rhodophyta: Florideophycidean orders. – Brit. Phycol. J. **22**: 125-138, 1987.

Glazer, A.N.: Phycobilisomes. – In: Abelson, L.N., Simon, M.I. (ed.): Methods in Enzymology. Vol. **167**. Pp. 304-312. Academic Press, San Diego – New York – Berekeley – Boston – London – Sydney – Tokyo – Toronto 1988.

Hess, W.R., Partensky, F., Van der Staay, G.W.M., Garcia-Fernandez, J.M., Börner, T., Vaulot, D.: Coexistence of phycoerythrin and a chlorophyll *a/b* antenna in a marine prokaryote. – Proc. nat. Acad. Sci. USA **93**: 11126-11130, 1996.

Janson, S., Graneli, E.: Phylogenetic analyses of nitrogen-fixing cyanobacteria from the Baltic Sea reveal sequence anomalies in the phycocyanin operon. – Int. J. system. evol. Microbiol. **52**: 1397-1404, 2002.

Lokstein, H., Steglich, C., Hess, W.R.: Light-harvesting antenna function of phycoerythrin in *Prochlorococcus marinus*. – Biochim. biophys. Acta **1410**: 97-98, 1999.

MacColl, R.: Cyanobacterial phycobilisomes. – J. struct. Biol. **124**: 311-334, 1998.

Nielsen, R., Yang, Z.H.: Likelihood models for detecting positively selected amino acid sites and applications to HIV-1 envelope gene. – Genetics **148**: 929-936, 1998.

Oliveira, M.C., Bhattacharya, D.: Phylogeny of the Bangiophycidae (Rhodophyta) and the second endosymbiotic origin of algal plastids. – Amer. J. Bot. **87**: 482-492, 2000.

Robertson, B.R., Tezuka, N.R., Watanabe, M.M.: Phylogenetic analyses of *Synechococcus* strains (cyanobacteria) using sequences of 16s rDNA and part of the phycocyanin operon reveal multiple evolutionary lines and reflect phycobilin content. – Int. J. system. evol. Microbiol. **51**: 861-871, 2001.

Roell, M.K., Morse, D.E.: Organization, expression and nucleotide sequence of the operon encoding R-phycoerythrin α and β subunits from the red alga *Polysiphonia boldii*. – Plant mol. Biol. **21**: 47-58, 1993.

Sambrook, J., Fritsch, E., Mantiatis, T.: Molecular Cloning: A Laboratory Manual. – Cold Spring Harbor Laboratory Press, Cold Spring Harbor – New York 1989.

Steglich, C., Mullineaux, C.W., Teuchner, K., Hess, W.R., Lokstein, H.: Photophysical properties of *Prochlorococcus marinus* SS120 divinyl chlorophylls and phycoerythrin *in vitro* and *in vivo*. – FEBS Lett. **553**: 79-84, 2003a.

Steglich, C., Psot, A.F., Hess, W.R.: Analysis of natural populations of *Prochlorococcus* spp. in the northern Red Sea using phycoerythrin gene sequence. – Environ. Microbiol. **5**: 681-690, 2003b.

Sui, Z.H., Zhang, X.C.: Cloning and analysis phycoerythrin genes in *Gracilaria lemaneiformis* (Rhodophyceae). – Chin. J. Oceanol. Liminol. **18**: 42-46, 2000.

Swanson, R.V., Ong, L.J., Wilbanks, S.M., Glazer, A.N.: Phycoerythrins of marine unicellular cynanobacteria. II. Characterization of phycobiliproteins with unusually high phycourobilin content. – J. biol. Chem. **266**: 9528-9534, 1991.

Thompson, J.D., Higgins, D.G., Gibson, T.J.: CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. – Nucl. Acids Res. **22**: 4673-4680, 1994.

Ting, S.C., Rocap, G., King, J., Chisholm, S.W.: Phycobiliprotein genes of the marine photosynthetic prokaryote *Prochlorococcus*: evidence for rapid evolution of genetic heterogeneity. – Microbiology **147**: 3171-3182, 2001.

Wilbanks, S.M., de Lorimier, R., Glazer, A.N.: Phycoerythrins of marine unicellular cyanobacteria. III. Sequence of a class II phycoerythrin. – J. biol. Chem. **266**: 9535-9539, 1991.

Yang, Z.H.: PAML: a program package for phylogenetic analysis by maximum likelihood. – CABIOS **13**: 555-556, 1997.

Yang, Z.H., Nielsen, R.: Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. – Mol. biol. Evolut. **19**: 908-917, 2002.

Yang, Z.H., Nielsen, R., Goldman, N., Pedersen, A.-M.K.: Codon-substitution models for heterogeneous selection pressure at amino acid sites. – Genetics **155**: 431-449, 2000.

Zuber, H.: Structure of light-harvesting antenna complexes of photosynthetic bacteria and red algae. – Trends biochem. Sci. **11**: 414-419, 1986.