International Journal of Biological Sciences
ISSN 1449-2288 www.biolsci.org 2007 3(7):434-445
©Ivyspring International Publisher. All rights reserved

Research Paper

Phycobilisomes linker family in cyanobacterial genomes: divergence and evolution

Xiangyu Guan[1,*], Song Qin[2,*], Fangqing Zhao[2], Xiaowen Zhang[2], Xuexi Tang[1]

1. College of Marine Life Science, Faculty of Life Science and Technology, Ocean University of China, 266003, Qingdao, P.R. China
2. Institute of Oceanology, Chinese Academy of Sciences, 266071, Qingdao, P.R. China

* Contributed equally to this study.

Correspondence to: Dr. Xuexi Tang, tel.: +86 532 82898500; fax: +86 532 82880645; e-mail: mutushicn@yahoo.com.cn

Received: 2007.03.23; Accepted: 2007.11.06; Published: 2007.011.10

Cyanobacteria are the oldest life form making important contributions to global CO_2 fixation on the Earth. Phycobilisomes (PBSs) are the major light harvesting systems of most cyanobacteria species. Recent availability of the whole genome database of cyanobacteria provides us a global and further view on the complex structural PBSs. A PBSs linker family is crucial in structure and function of major light-harvesting PBSs complexes. Linker polypeptides are considered to have the same ancestor with other phycobiliproteins (PBPs), and might have been diverged and evolved under particularly selective forces together. In this paper, a total of 192 putative linkers including 167 putative PBSs-associated linker genes and 25 Ferredoxin-NADP oxidoreductase (FNR) genes were detected through whole genome analysis of all 25 cyanobacterial genomes (20 finished and 5 in draft state). We compared the PBSs linker family of cyanobacteria in terms of gene structure, chromosome location, conservation domain, and polymorphic variants, and discussed the features and functions of the PBSs linker family. Most of PBSs-associated linkers in PBSs linker family are assembled into gene clusters with PBPs. A phylogenetic analysis based on protein data demonstrates a possibility of six classes of the linker family in cyanobacteria. Emergence, divergence, and disappearance of PBSs linkers among cyanobacterial species were due to speciation, gene duplication, gene transfer, or gene loss, and acclimation to various environmental selective pressures especially light.

Key words: phycobilisomes, cyanobacteria, linker polypeptides, evolution

1. Introduction

Cyanobacteria are prominent constituents of marine biosphere that account for a significant percentage of oceanic primary productivity, and are among the oldest life forms on the Earth capable of doing oxygenic photosynthesis about 3.5 billion years ago, which is similar to the process found in higher plants [1-2]. As the oldest and major light-harvesting antennae, PBSs are highly organized complexes of various PBPs and linker polypeptides (Fig.1.), and are very diverse in structure and pigment composition in cyanobacteria, red algae, and the cryptomonads [3-5]. They function in light harvesting and energy migration toward photosystem II or I reaction centers in thylakoid membrane, except *Gloeobacter violaceus* PCC7421 (Gv) having no thylakoid membrane [6,7].

On one hand, PBSs linkers transfer energy of PBPs to favor a unidirectional flow of excitation energy from the peripheral rod of PBSs to the PBSs core and then from the PBSs core to the photosynthetic reaction center [8]. On the other, PBSs linkers function to stabilize PBSs structure and determine positions of the PBPs within PBSs structure. At the same time, PBSs linkers also interact directly or indirectly with the chromophores to cause the PBSs structure changes that can modulate different PBPs subassemblies and optimize absorbance characteristics [9-11]. The structural function of PBSs linkers in PBSs has allowed cyanobacteria to colonize environments and show a great diversity in terms of light quantity and quality [12,13]. Positions of highly conserved PBPs were determined by the specific linker polypeptides, and it is possible that linker polypeptides somehow interact to form a scaffold-like structure within PBSs [14]. Whether or not this is the case, it is possible to distinguish various PBPs assemblies specifically by their state of aggregation and by their attachment to relevant linker polypeptides [15-17]. Tandeau de Marsac and Cohen-Bazire demonstrated for the first time that several colorless polypeptides that take 12%–15% of the total stainable proteins of the PBS components are accounted for linker polypeptides from eight species of cyanobacteria by SDS-PAGE [18]. The nominated system of linker polypeptides are according to their locations and molecular masses in PBSs. Glazer [19] has provided a system of abbreviations to characterize linker peptides with respect to their locations and molecular masses in PBSs: PBSs rod linker (L_R, 27 to 35 kDa), PBSs rod-core linker (L_{RC}, 25 to 27 kDa), PBSs core linker (L_C, 7.7 to 7.8 kDa), and PBSs core-membrane linker (L_{CM}, 70 to 120 kDa) [16,20]. The importance of linker polypeptide for the assembly of defined complexes and their roles for tuning spectral characteristics of the complexes has been well understood [21,22].

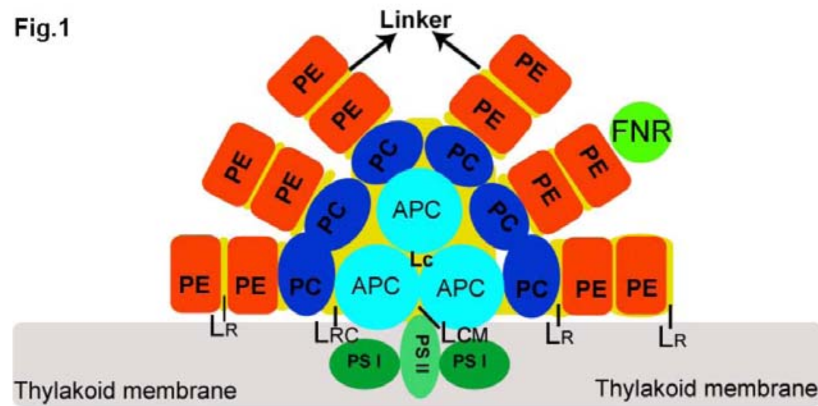

Fig. 1. Structural model of a tricylindrical hemidiscoidal phycobilisome (2, 3). The three sky blue circles represent the tricylindrical core APC, and two bottom cylinders attach to the thylakoid membrane (grey rectangle) with L_{CM}. Six rods are arranged by PC (blue circle), and PE (red circle), and attached FNR (grass green circle) with L_R from inner to outer part. L_{RC} is the linker between core and rod. All linkers are represented by yellow discs located in each rod.

FNR, being also considered as linker polypeptides, transfers electrons from ferredoxin to $NADP^+$ to generate NADPH with an average value of 1.3 FNR per PBS, [23,24]. FNR encodes a protein that is composed of three domains: two C-termainal domains enough to enzymatic activity of FNR and a ~9kDa N-terminal domain generally homologous to the small phycocyanin (PC) rod-linker polypeptide CpcD [23,25]. With the exception of CpcD, it is also reported that there are similarity between FNR and other PBSs linkers' different domains [25,26]. In contrast to other PBSs-associated linkers (cluster with PBPs), the γ subunits serving as phycoerythrin (PE) linker polypeptides are chromophorylated, containing two types of covalently attached linear tetrapyrrole chromophores, phycoerythrobilin (PEB), and phycourobilin (PUB) [27]. Genes of L_{CM} and L_{RC} polypeptides are on the plastid genome, while genes ending the γ subunits are present on the nuclear genome [28,29]. Liu [27] found that no high-degree sequences homology exists between the γ subunits and other linker polypeptides, and suggested that different primary structures in a range of balanced states still perform similar physiological functions [30,31]. In red algae, γ subunits that are also the main chromophorylated components of PBPs and orderly assembled into other PBPs forming a stable complex with α and β subunits of PE [32-34].

At present, more and more cyanobacterial genomes' database have brought about a great convenience in search for PBSs linkers using bioinformatic tools. Here, a comparative genomic analysis on all the data sequences of PBSs linkers in the cyanobacteria is presented. Observation on PBSs linker polypeptides was made in 25 cyanobacteria additional to some model strain cyanobacteria with improved method of separation of the PBSs linker family. Besides, evolution of linker polypeptides in the varieties of PBSs was analyzed and specific connections to PBPs or other linkers were performed for better understanding the function of PBSs in different environments.

2. Materials and Methods

Database searching and sequence retrieving

Genomes database were searched at JGI (http://www.jgi.doe.gov/). Protein sequences of the PBSs linkers previously described were used as queries for database. Cyanobacteria species examined included 25 cyanobacterial genomes (20 complete and 5 ongoing): *Anabaena*, *Nostoc*, *Gloeobacter*, *Trichodesmium*, *Crocosphaera*, *Synechocystis*, *Synechococcus* and *Prochlorococcus*. All 25 genome sequences were accessed from IMG (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi) in FASTA format. Each protein in this query dataset was used to search potential novel sequences in all cyanobacteria species with all available genome sequences, by using the BLASTP and TBLASTN programs. Sequences giving better reciprocal BLAST hits were assumed capable of identifying homologous counterparts in these species if they could be aligned up with at least the BLAST-Score > 90 and the E-value < 1E-10. The search was iterated until convergence, examined individually, and then aligned with Clustal X [35]. Sequence identity and similarity were calculated using BioEdit v5.0 [36]. To elucidate the complete genomic structure of the PBSs linkers' genes, all linkers onto the 20 finished cyanobacterial genomes were mapped (Fig. 2).

Similar searches were run with the Pfam domains of the PBSs linker proteins to avoid the exclusion of highly diverged sequences with just limited conserved motifs. PFAM [37] and SMART [38] domain analyses with derived sequences that employed as queries were then carried out to eliminate false positives.

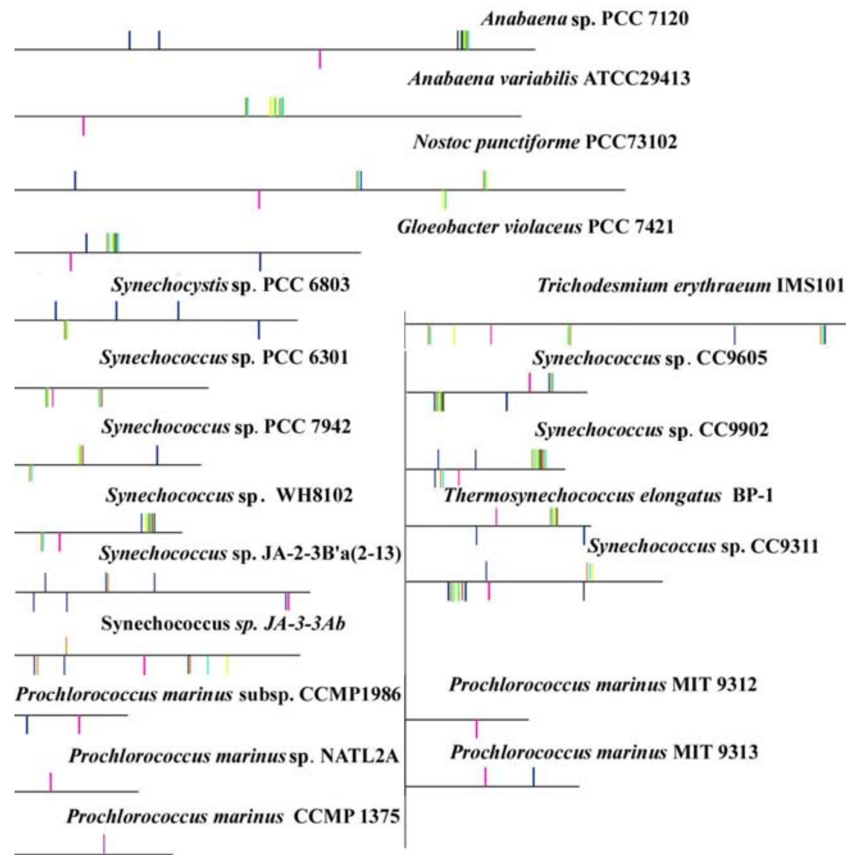

Fig.2. Genomic organization of PBSs linkers in 20 sequenced completely cyanobacterial genomes. Vertical bars show the locations of PBSs linker genes, with FNR in pink and PBSs linkers except FNR in other colors. PBSs linker genes were mapped onto the chromosome evenly. Long horizontal line indicates the chromosome. The vertical bar above horizontal line indicates the transcriptional direction opposite to that below a horizontal line.

Phylogenetic analysis

Multiple sequences were aligned using Clustal X, and then adjusted manually. To maximize the number of sites available for analysis, partial sequences and certain sequences with large deletions were excluded. To understand the evolutionary relationships of all PBSs linkers in these cyanobacteria genomes, neighbor-joining method in MEGA3 [39] and maximum parsimony method in PHYLIP [40] were used to construct the phylogenetic tree, in which confidence levels of each branch were determined by analyzing 1000 bootstrap replicates.

3. Results and discussion

The PBSs linker family in cyanobacteria

From nominated linkers, PBPs-associated linker family comprises five groups of linkers with own locations and molecular masses in the PBSs [19]. Most of the PBSs linkers are clustered with APC (allophycocyanin), and PC or PE. Therefore, they are called APC-associated linker (L_C), PC- and PE-associated linker (L_R) [41,42], while L_{RC} and L_{CM} are involved in attaching peripheral rods to the APC cores. BLASTP and TBLASTN programs analyses show that a total of 192 linkers (159 putative PBPs-associated linker genes, 8 γ subunits, and 25 FNR genes) were obtained and genes ApcC, CpcC, CpcD, CpeC, CpeD, CpeE, MpeC, MpeE, PecC, CpcG, ApcE, and PetH derived from 25 cyanobacterial genomes in this study (Table 1). The species of 21 cyanobacterial strains, their morphologies, main features, and habitats, as well as the abbreviations used at the end of gene names are shown (Table 2). The number of PBPs-associated linkers in these cyanobacteria is from 1 to 13, and there are a maximum number of 13 PBSs linkers in Cw and Tr. γ subunits as special chromophoric linkers were found in some marine *Synechococcus* and three low-light adapted *Prochlorococcus* lineages, while other three sequenced *Prochlorococcus* have only one type linker FNR (Table 1). Although no γ subunit has been found in low-light adapted P13 from database, it may exist because of the same mode of light-harvesting in other low-light adapted *Prochlorococcus*, and γ subunits are difficult to be found with only BLASTP as it is in low sequence homology between γ subunits and other linkers [34]. The light-harvesting structure including γ subunit in low-light adapted *Prochlorococcus* can acclimatize themselves well to same factors especially low-light of environment.

Table 1. Cyanobacterial genes encoding PBSs linkers.

species\Gene product	APC-associated PBSs core linker(LC)	PC,PE-associated PBSs rod linker(LR)	PBSs rod-core linker(LRC)	PBSs core-membrane linker(LCM)	γ subunit	FNR	NO
Nostoc sp. PCC 7120 (N7) (F)	ApcCN7 asr0023	CpcCN7 alr0530 CpcDN7 asr0531 PecCN7 alr0525	CpcG1N7 alr0534 CpcG2N7 alr0535 CpcG3N7 alr0536 CpcG4N7 alr0537	ApcEN7 alr0020		PetHN7 all4121	10
Anabaena variabilis ATCC29413 (Av) (F)	CpcD3Av Ava2623	CpcD4Av Ava2933 CpcD2Av Ava2932 CpcD1Av Ava2927	Ava2936 Ava2937 Ava2938 Ava2939	Ava2620		PetHAv Ava0782	10
Nostoc punctiforme PCC73102 (Np) (D)	CpcD5Np NpR4840	CpcD1Np NpF0736, CpcD2Np NpF3794 CpcD3Np NpF5291, CpcD4Np NpF5292 CpcD6Np NpF5293	NpF3811 NpF3795	NpR4843		PetHN7 NpR2751	10
Gloeobacter violaceus PCC 7421 (Gv) (F)	ApcCGv gsr1248	CpcC1Gv glr0950, CpcC2Gv gll3219 CpcD1Gv gsr1266, CpcD2Gv gsr1267 CpeCGv glr1263, CpeDGv glr1264 CpeEGv glr1265, glr2806, glr1262		ApcEGv glr1245		PetHGv gll2295	12
Trichodesmium erythraeum IMS101 (Tr) (F)	CpcD2Tr Tery_3647	Tery_4104, Tery_4105 Pec1Tr Tery_4106, Pec2Tr Tery_4107 Tery_0999, Tery_0985 CpcD1Tr Tery_0986	Tery_2486 Tery_3909	Tery_2209 Tery_2210		PetHTr Tery_3658	13
Crocosphaera watsonii WH8501 (Cw) (D) Crocosphaera watsonii WH8501	CpcD7Cw Contig357_or4307	CpcD3Cw Contig361_or5717 CpcD5Cw Contig362_or6341 CpcD6Cw Contig166_or0659 PecC1Cw Contig361_or5719 PecC2Cw Contig361_or5721 Contig315_or2854 CpcD1Cw Contig166_or0658 CpcD2Cw Contig315_or2837 CpcD4Cw Contig361_or5718	Contig362_or6343	Contig207_or1063		PetHCw Contig343_or3658	13
Synechocystis sp. PCC 6803 (S6) (F)	ApcCS6 ssr3383	CpcC1S6 sll1580 CpcC2S6 sll1579 CpcDS6 ssl3093	CpcG1S6 slr2051 CpcG2S6 sll1471	ApcES6 slr0335		PetHS6 slr1643	8
Synechococcus sp. CC9311 (S9) (F)	ApcCS9 sync_2325	CpeD1S9 sync_0511 638114101 sync_0512 CpeCS9 sync_0513 638114105 sync_0516 CpeD2S9 sync_2251	638114104 sync_0515 638114838 sync_1249 CpcG1S9 sync_2488	ApcES9 sync_2321	MpeCS9 sync_0502	PetHS9 sync_1003	12
Synechococcus sp. WH 8102 (S8) (F)	ApcCS8 SYNW0483	MpeES8 (II) SYNW1989 MpeDS8 (II) SYNW2000 CpeCS8 (I) SYNW1999 CpeES8 (I) SYNW2001	CpcG1S8 SYNW0314 CpcG2S8 SYNW1997	ApcES8 SYNW0486	MpeCS8 SYNW2010	PetHS8 SYNW0751	10

Synechococcus sp. CC9605 (S96) (F)	ApcCS96 Syn_cc96052199	CpcCS96(II) Syn_cc96051534 CpcD1S96(II) Syn_cc96050443 CpcD2S96(I) Syn_cc96050444 Syn_cc96050442	Syn_cc96050446 Syn_cc96052287 CpcGS96 Syn_cc96052579	ApcES96 Syn_cc96052196	Syn_cc96050433	PetHS96 Syn_cc96051917	11
Synechococcus sp. CC9902 (S99) (F)	CpcD1S99 Syn_cc99020477	CpcD2S99 Syn_cc99021899 Syn_cc99021871, Syn_cc99021885 Syn_cc99021883, Syn_cc99020444	Syn_cc99021881 Syn_cc99021003 Syn_cc99020399	Syn_cc99020480	Syn_cc99021895	PetHS99 Syn_cc99020749	12
Synechococcus elongatus PCC 7942 (S79) (F)	CpcD1S79 Syn_pcc79420325	403100330 Syn_pcc79421049 403100340 Syn_pcc79421050 CpcD2S79 Syn_pcc79421051	403110230 Syn_pcc79422030	403092970 Syn_pcc79420328		PetHS79 Syn_pcc79420978	7
Synechococcus sp. PCC 6301 (S63) (F)	ApcCS63 syc1188_d	CpcC1S63 syc0498_c CpcC2S63 syc0499_c CpcDS63 syc0497_c	CpcGS63 syc2065_d	ApcES63 syc1185_d		PetHS63 syc0566_c	7
Thermosynechococcus elongatus BP-1 (Te) (F)	ApcCTe tsl0955	CpcCTe tlr1959 CpcDTe tsr1960	CpcG1Te tlr1963 CpcG2Te tlr1964 CpcG4Te tlr1965	ApcETe tll2365		PetHTe tlr1211	8
Synechococcus sp. WH 7805 (S78) (D)	639019614 WH7805_12498	639020074 WH7805_06646 639020076 WH7805_06656 639020077 WH7805_06661	639019440 WH7805_11638 639020072 WH7805_06636	639019618 WH7805_12518		PetHS78 WH7805_04581	8
Synechococcus sp. WH 5701 (S57) (D)	638958495 WH5701_15296	638958186 WH5701_05910 638958190 WH5701_05930 638959531 WH5701_08859	638958192 WH5701_05940 638958614 WH5701_15881	638958492 WH5701_15281	638961018 WH5701_00450	PetHS57 WH5701_10210	9
Synechococcus sp. RS9917 (SRS) (D)	638963552 RS9917_08310	638963041 RS9917_02873 638963045 RS9917_02893	638963039 RS9917_02863 638963429 RS9917_07710	638963555 RS9917_08325		PetHSRS RS9917_01102	7
Synechococcus sp. JA-3-3Ab (SJAb) (F)	ApcCSJAb CYA_2225	CpcDSJAb CYA_0218 637872096JAb CYA_0506 637872115JAb CYA_0528 CpcCSJAb CYA_2041	CpcG1SJAb CYA_0215	637873357JAb CYA_1814 637873394JAb CYA_1851		PetHSJAb CYA_1257	9
Synechococcus sp. JA-2-3Ba (2-13) (SJBa) (F)	ApcCSJBa CYB_1440	CpcD1SJBa CYB_0941 637874979 CYB_0568 CpcCSJBa CYB_2737	CpcG1SJBa CYB_0944	637874843 CYB_0431		PetHSJBa CYB_2882	7
Prochlorococcus marinus str. MIT 9313 (P93) (F)						PetHP13 PMT1101	1
Prochlorococcus marinus sp. NATL2A (Pn) (F)					MpeCPn PMN12a1678	PetHPn PMN12a0675	2
Prochlorococcus marinus str. MIT 9312 (P12) (F)						PetHP12 Pmt93121086	1
Prochlorococcus marinus subsp. CCMP1986 (P86) (F)						PetHP86 PMM1075	1
Prochlorococcus marinus str. CCMP 1375 (P75) (F)					PpeCP75 Pro0345	PetHP75 Pro1123	2
Prochlorococcus marinus MIT 9211 (P92)					638824638 P9211_07152	PetHP92 P9211_03182	2
Number in all	19	79	40	21	8	25	192

Words in first () of species line are abbreviations; Words in second () of species line are "Genome Completion: [F]inished, [D]raft".

Table 2. The species names of the 21 cyanobacterial strains, morphologies, main features, and habitats.

Species	Morphology	Genome size	Linkers (%)	LHC	Features
Prochlorococcus marinus subsp. CCMP1986	Unicellular	1760	0.57	Chl a_2/b_2	Marine; HH
Prochlorococcus marinus str. MIT 9312	Unicellular	1853	0.54	Chl a_2/b_2	Marine; HH
Prochlorococcus marinus sp. NATL2A	Unicellular	1937	1.03	Chl a_2/b_2	Marine; LH
Prochlorococcus marinus str. CCMP 1375	Unicellular	1926	1.04	Chl a_2/b_2	Marine; LH
Prochlorococcus marinus str. MIT 9313	Unicellular	2327	0.43	Chl a_2/b_2	Marine; LH
Synechococcus sp. CC9311	Unicellular	2942	4.08	PBSs	Marine
Synechococcus sp. WH 8102	Unicellular	2580	3.88	PBSs	Marine
Synechococcus sp. CC9902	Unicellular	2358	5.09	PBSs	Marine
Synechococcus sp. CC9605	Unicellular	2753	4.00	PBSs	Marine
Synechococcus elongatus PCC 7942	Unicellular	2712	2.58	PBSs	Freshwater
Synechococcus elongatus PCC 6301	Unicellular	2578	2.72	PBSs	Freshwater
Crocosphaera watsonii WH8501	Unicellular	5996	2.17	PBSs	Nitrogen-fixing
Synechocystis sp. PCC 6803	Unicellular	3618	2.21	PBSs	Freshwater
Trichodesmium erythraeum IMS101	Filamentous	7750	1.68	PBSs	Nitrogen-fixing
Nostoc punctiforme PCC73102	Filamentous	7672	1.30	PBSs	Heterocystous
Anabaena variabilis ATCC29413	Filamentous	5760	1.74	PBSs	Heterocystous
Nostoc sp. PCC 7120	Filamentous	6210	1.61	PBSs	Heterocystous
Thermosynechococcus elongatus BP-1	Unicellular	2521	3.17	PBSs	Thermophilic
Gloeobacter violaceus PCC 7421	Unicellular	4478	2.68	PBSs	No thylakoid membranes
Synechococcus sp. JA-3-3Ab	Unicellular	2813	3.20	PBSs	Thermophilic
Synechococcus sp. JA-2-3B'a(2-13)	Unicellular	2913	2.75	PBSs	Thermophilic

In overall, the basic architecture of PBS is widely conserved, while PBPs, core structure, and PBSs linkers diversified greatly across different strains of cyanobacteria [43]. Moreover, for a single strain, it depends upon the environmental conditions, such as nutrient availability, temperature, light quality, and light intensity [14]. L_C coexisting with L_{CM} is a single-gene in 19 cyanobacterial genomes except for five sequenced *Prochlorococcus*, and L_R and L_{RC} can be found in different amounts in these 20 cyanobacteria. All 25 cyanobacteria including Gv without thylakoids have an FNR. Multiple copies of L_{RC} were identified in most of referred cyanobacterial species in this study. However, only one L_{RC} was found in Cw, S79, S63, SJAb, and SJBa (Table 1), and there is no such linker in Gv and *Prochlorococcus*. The numbers of PC and PE-associated rod linkers are also diverse among these 25 cyanobacteria. The composition of PBS in cyanobacteria and red algae vary in response to environmental changes in light intensity, light quality (only cyanobacteria), and nutrient availability [43]. Differences in chromophore composition of phycobiliproteins result in wavelength-specific difference in light absorption among species of cyanobacteria. The assembly of the PBS is mediated by linker polypeptides, and each trimeric or hexameric subassembly of PBS contains at least one specific linker polypeptide, which determines the type, location, and aggregation state of the PBP within the rod and also modulates the spectroscopic properties [34]. Light quality and quantity are among the major factors affecting the composition of PBSs. In some cyanobacteria, the relative proportion of PC and PE can vary within the PBS rods in response to a change in light climate [5,17], but such complementary chromatic adaptation is rare among marine *Synechococcus* [9,41]. Changes in photon fluxes also have an effect on the structure of PBS. Marine cyanobacteria can resist high light stress by decreasing the content of PBS in cell [44,45], due to a reduction in surface of thylakoid membranes [46]. *Prochlorococcus* and *Synechococcus* are abundant unicellular cyanobacteria and major participants in global carbon cycles. Although *Prochlorococcus* and *Synechococcus* are closely related to each other and often cohabit, they possess very different photosynthetic light-harvesting antennas [45,47,48,49]. *Synechococcus* and the majority of cyanobacteria use PBSs having 7-12 linkers, while *Prochlorococcus* uses a chlorophyll a2 /b2 light-harvesting complex. Differences in absorption properties and cellular costs between chlorophyll a2/b2 and PBS antennas differentiate them with own ecological niche in the ocean [45,50]. *Prochlorococcus* is a unicellular cyanobacterium that lacks PBS and contains chlorophyll b as major accessory pigment, which enables it to absorb blue light efficiently at low-light intensity and blue wavelengths characteristic in deep euphotic zone [50]. P86 and P13 are representatives of high- and low-light adapted ecotypes. The low light-adapted strain has significantly more genes than its high light counterpart such as γ subunits, but neither has PBSs-associated linkers. As transitional light-harvesting antennae, γ subunits appeared more recently than other PBSs-associated linkers and the linkers (FNR and γ subunits) have some compensatory function.

Evolution in genus *Prochlorococcus* would have evolved towards genome reduction [45]. Specific genome amplification and diversification have taken certain place during adaptation of the latter to their specific environments [45]. The PBSs linker genes should be one of the reduced genes along with *Prochlorococcus* genomes reduction, whereas all *Prochlorococcus* strains evolved to use (divinyl-) chlorophyll a/b-protein complexes as the major

antenna system [50]. In reverse, there are many PBSs linkers in the genomes of *Synechococcus* that is also considered as genome reduction. Whether such a reduced genome is a derived state resulting from progressive gene loss or is an ancestral state are unclear. The PBSs of open-ocean *Synechococcus* cyanobacteria are among the most complex ones described so far since they possess four types of constitutive PBPs: APC, PC, and two forms of phycoerythrin: PEI and PEII. The PE-associated linkers are divergent to PEI- and PEII-associated linkers along with PE divergence. In previous studies, the PBSs and linkers were probably co-evolved from very early stage [51]. Some PBSs rods have one combination of three PE-associated linkers (e.g., CpeC, MpeD, and MpeC), while others would have another combination (e.g., CpeE, MpeD, and MpeE) [26]. Indeed, PBSs rods are in fact more compact than we assumed. Cyanobacteria with short PBSs rods may have a heterogeneous linker composition, which allow these cyanobacteria to colonize a variety of light-quantity and light-quality environments. Ting et al. [2] presented a scenario to explain how *Prochlorococcus* antenna evolved in an ancestral cyanobacterium in iron-limited oceans, resulting in diversification in *Prochlorococcus* and marine *Synechococcus* lineages from a common PBS-containing ancestor.

Genomic distribution and sequence analysis of PBSs linker genes

In many cases, the PBSs-associated linkers and PBPs as well as enzymes that involved in biosynthesis or binding of phycobilins are directly adjacent to each other on chromosome and may form up an operon that is regulated by a same activator [52,53]. From these data, 36 gene clusters ranging from 1.5 to 13.2 kb were found in these cyanobacterial genomes (Fig.3). In overall, there are 15 APC-associated gene clusters, 12 PC-associated gene clusters, 4 PE with PC or ambiguous PBPs-associated gene clusters, and 5 ambiguous PBPs-associated gene clusters (Fig.3). Most PBSs-associated linkers and PBPs often clustered and transcribed in the same direction, but in the reverse direction with some enzymes' genes. The other linkers' genes such as FNR are arranged randomly in the genome. There are also many linkers contiguous to each other in these genomes such as P99 and Cw genomic sequences, and they do not assemble into an operon. Although there are γ subunits in some *Prochlorococcus* and *Synechococcus*, but they do not form gene clusters with other PBSs linker genes. APC-associated linker gene clusters exist universally, and the largest putative operon makes up of PE with PBPs and enzymes' genes in S93. Genes encoding PC or PEC subunits are typically followed by genes encoding APC-associated linker polypeptides and/or the genes for chromophore attachment to the alpha subunit. A L_{RC} gene always follows the above operon and forms a separate transcription unit. The *apcC* gene, encoding small linker polypeptide $Lc^{8.9}$, lies downstream from *apcB* and *apcE* genes, and upstream from *apcA* gene. *apcC* gene with *apcA/B* genes locate together on a transcriptional unit. As one type of the terminal acceptors of excitation energy within the PBSs, *apcC* is also found in the attachment of PBSs to the membrane.

The structure of PBSs linker gene clusters varies among species. Almost of PBSs-associated linkers form up gene clusters with PBPs, but the number of CpcG and CpeC, CpeD, and CpeE that are divergent from a same ancestor is uncertain in gene clusters. In N7, there are a PC gene cluster constituting of four CpcG1-4 followed by cpcB/A (C-PC β and chain), CpcD (PC-associated rod linker), and CpcE/F (Phycocyanobilin lyase α and β subunit), loaded in a PE gene cluster adjacently, while in Te a gene cluster make up of three CpcG1, 2, 4 with CpcB/A and CpcE/F. Similarly, an obvious diversity of a PE-cluster existed in Gv, and has three PE-associated linkers CpeC, CpeD, and CpeE. In S63 and S79, the duplicated PC genes arrange a tandem repeat unit with three rod linkers' genes between $cpcB_1/A_1$ and downstream $cpcB_2/A_2$, $cpcE/F$ set. The occurrence of several different gene sets in the same type of PBS component is apparently the result of adaptation of these organisms to different environmental conditions, such as light quality and nutrient availability. Some linkers have been diverged from a common ancestor along with PBPs divergence and may be caused by gene duplication or horizontal gene transfer.

With known morphology of PBSs linker clusters to construct PBSs, divergence and evolution in arrangement of PBSs cannot be excluded. For example, some PBSs might have one combination of APC-, PC-, and PE-associated linker clusters, while others would have different combinations (only one or two PBPs-associated linker clusters). It is possible that PBSs linker clusters are in fact more compact than we thought, with short rods having a heterogeneous linker composition adapted well to changing environments [46]. In another case, a model strain *S*81 consists of an APC core and rods that made of one type of PC and two types of PE (I and II), and gather into a complicated operon. PEI and PEII can bind both PUB and PEB in different proportions to light acclimation [46]. In some cyanobacteria, PBSs rod of only PC is simpler since it possesses only one complete set of α and β subunits and two PBSs rod-core linkers (CpcG1 and CpcG2), indicating probably a heterogeneous rod linker composition [54]. Six [46] hypothesized that PEII-associated linker would firmly anchor the proximal PEII disk to the rest of the rod, whereas the short C-terminus of another PEII-associated linker makes it more susceptible to release/breakage during photo acclimation processes.

In S81, S99, and S96, *cpeR* occurs downstream of PBPs-associated linkers and genes related with PBSs. The structure is similar to the operon structure of *F. diplosiphon*, and the genes of operon are regulated by the same activator such as CpeR. CpeR is transcribed as a part of the *cpeCDE* operon on an extended transcript, and required for expression of the *cpeB/A* operon. Therefore, it is proposed that at onset of green light, operons *cpeCDESTR* and *cpeB/A* are expressed in series as a genetic cascade [53]. Maybe the clusters in S81, S99, and S96 work in the same fashion to that of the operon in *F. diplosiphon*. According to known genes, ambiguous genes function in clusters of S81, S99, and S96 can be inferred comparatively. This method can deduce the genes' function such as *cpeS*, and *cpcT* [55,56], and provide information for validation in experiment.

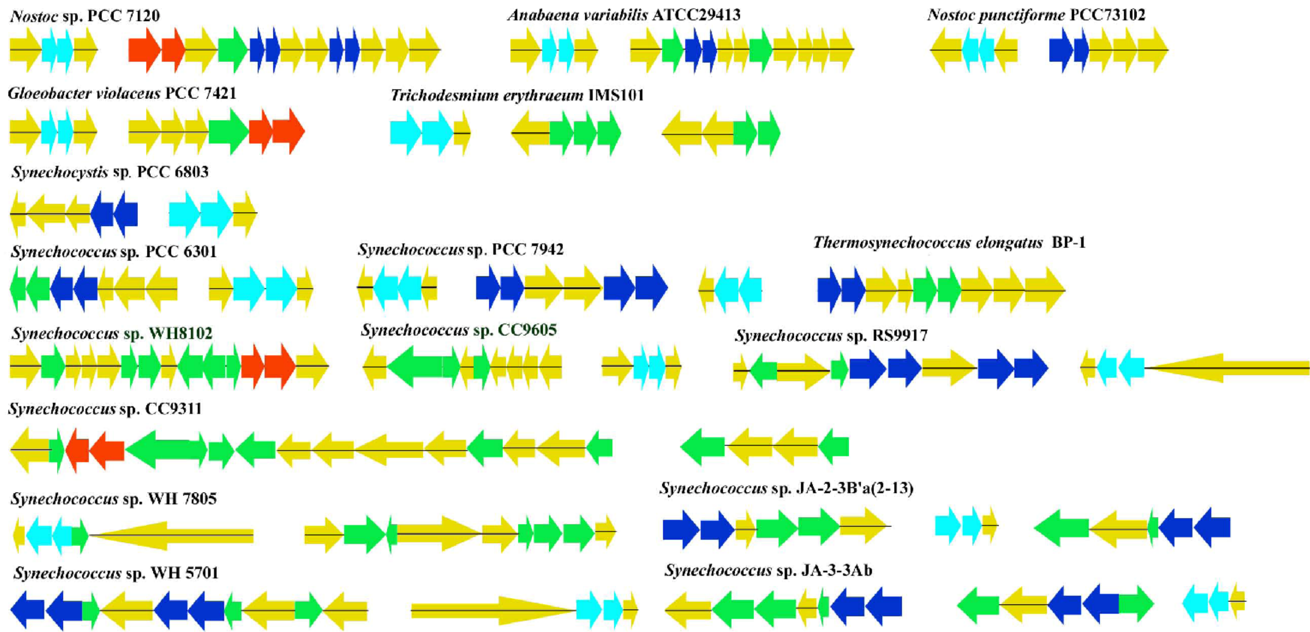

Fig.3. Organization of the gene clusters encoding PBSs and PBSs-associated linkers of 17 cyanobacteria. Arrows represent the direction of translation, and the relative sizes of operon deduced from analysis of the amino acid sequence. The cyanobacteria names are given on top of the corresponding region. Yellow arrows indicate the PBSs-associated linker; sky blue, blue, and red arrows represent APC, PC, and PE, respectively; green arrows mean ambiguous PBPs or PBBs linkers.

Conservation domain analysis of PBSs linker

Ancestral PBPs are probably associated with same precursor of linker polypeptides. It may be possible that the linker polypeptides developed from an earlier (possibly non-globin) ancestor of PBPs [51,57]. Two additional unique β residues interact directly with the linker polypeptides, and linker polypeptides are conserved with β residues locating in F' and F helices [18]. In N7, we chose a cluster including PBSs-associated linkers, PBPs, PBPs lyase, and FNR, which have conversion domains shown in Fig.4. The amino acids of conversion domains are generally hydrophobic to form β-sheet. The sequence alignment of identified extensions from CpcG (1-4) protein shows that these extensions have 37%-54% identity and 59%-71% similarity, while the identity between CpcGs and PBPs, CpcGs and PBPs lyase are both <10%. The PC-associated linker and PE-associated linker have also a lower identity of about 20%. The amino acid sequences have high sequence identity with each other ranging from 43% to 99% (Fig. 4). The CpcD-like domain of FNR is more frequently found at the C-terminus of CpcD that encodes the unique rod-terminating linker protein $L_R^{8.9\ PC}$ [26]. Evidenced by this domain's presence at the alignment between $L_R^{8.9\ PC}$ and FNR in N7, the identity and similarity is only 10% and 14%, respectively. However, a CpcD-like domain is consistent with that localization, which is at the N-terminus of the CpcC proteins and a C-terminus domain, which is more similar to the PC-associated linker protein in Gv [26]. The sequence analysis shows that the identity and similarity between N-region of CpcA approximately 70 amino acids and CpcD are 16% and 44% in N7, respectively. Especially in high-light-grown cells, the *cpcD* gene apparently did not undergo gene duplication. Therefore, it may assume that $L_R^{8.9}$ functions as rod-terminating linker polypeptide for ending rods with PC [14]. The CpcG3 protein exhibited a strong resemblance to those of CpcG4 at 51% in identity and 72% in similarity, for CpcG1, a 38% identity, and a 59% similarity. CpcG1 with CpcG2 have 53% in identity and 69% in similarity. Conversion domains massed on the N-terminal was also observed among these series of CpcG genes. The sequence alignment in S6 shows that a large part of N-terminus of CpcG2 is almost identical to that of CpcG1, while the C-terminal part is highly diverged from each other [58]. With above values of alignment, CpcG1 and CpcG2 came from a common ancestor (I), and CpcG3 and CpcG4 came from another ancestor (II) in N7. Ancestor I and II associated with each other, and co-evolved in own distinct roles since the very beginning.

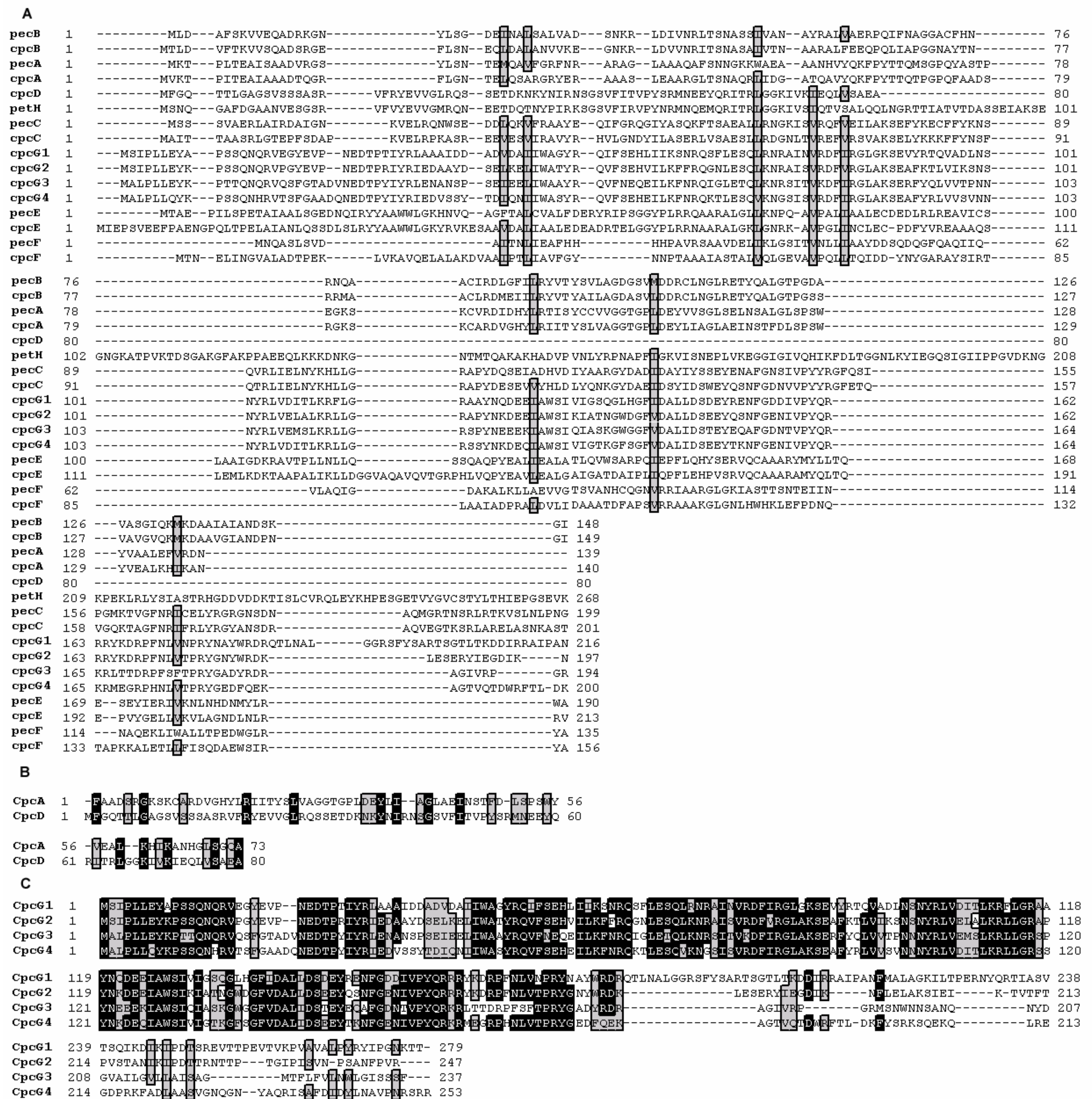

Fig.4. Multiple amino acid sequence alignment of the PBSs linkers with PBPs in *Nostoc* sp. PCC 7120. (A). Sequence alignments for an opera including pecB, pecC, pecE, pecF, cpcB, cpcA, cpcC, cpcD, cpcE, cpcF, and cpcG1-4 in N7. (B). Sequence alignments for N-terminal domain of FNR in N7. (C). Sequence alignments for cpcG1-4 in N7. The positions of similarity and identity are marked in gray and black, respectively.

Phylogenetic analysis

We further investigated the relationship among these PBSs linkers of these 25 cyanobacteria by generating an alignment of 192 identified PBSs linker amino acid sequences followed by generation of an MP phylogenetic tree (Supplementary Material). The resultant tree depicts 6 phylogenetic classes of APC-associated L_C, PC, PE-associated L_R, L_{RC}, L_{CM}, γ subunit, and FNR. Linkers of APC-associated L_C, L_{RC}, and L_{CM} were assembled in monophyletic group distinctly according to their sequences and functional

characteristics. Some PE-associated rod linkers are assembled into two groups in different branches, and others disperse throughout the rod linker cluster with PC-associated linkers. A cluster of PE-associated core linkers has a close relationship with γ subunits and the other with L_{CM}. Further, APC-associated L_C may share a recent common ancestor with CpcD (L_R). The phylogenetic tree reveals that the genetic diversification of all groups involves in a more complex pattern in gene degeneration and duplication. The L_{RC} (CpcG) of N7, Av, SJAb, SJBa, and Te form another cluster, and in two branches of that CpcG1 and CpcG2, CpcG3 and CpcG4 (except CpcG1 of SJAb and SJBa) are in at least 91% and 98% bootstrap values, respectively. These four copies (CpcG1-4) most likely have evolved into a recent duplication, whereas evolution of CpcG among these species is more complex with horizontal gene transfers.

Some ambiguous genes can be divided into these six classes in high bootstrap values with known PBSs linkers by phylogenetic analysis. For example, some innominate PBSs linkers are γ subunits (Syn_cc96050433, Syn_cc99021895, WH5701_00450), and they are high homologous with known MpeC (SYNW2010) in S81. These innominate PBSs linkers may be homological and have the same functions with known PBSs linkers such as MpeC. The ambiguous function genes should be validated in the future.

Although *Prochlorococcus* has no PBS, the vestiges of a CpcD-like domain which sequence has diverged and become shorter and extensively modified are still recognizable. FNR sequence is possible to be a potentially interesting evolutionary marker for both ancient and recent cyanobacteria [14]. The 16s rDNA sequences from 21 sequenced cyanobacteria were retrieved from IMG, and FNR was the only linker in all these cyanobacteria. Therefore, FNR as a PBSs linker identified in 21 cyanobacteria was also used to construct the phylogeny to discern the evolutionary history of the PBSs linker family (Supplementary Material graph B). Both 16s rDNA and FNR phylogenetic trees can be divided into two major unbalanced clades and separated into several monophyletic clusters with strong bootstrap support. In phylogenetic tree of FNR, clade I contains 18 FNR proteins of these 21 cyanobacteria, while clade II contains only 3 FNR proteins. In clade I, *Prochlorococcus* and marine *Synechococcus* form a cluster and share a common PBP-containing ancestor, and may have both diversified at a similar point in evolution [48]. FNR of S63 and S79 are identical, and are highly homologous with the branch of FNR in Av, Np, and N7. In clade II, SJAb sequences share a more recent common ancestor with SJBa than with other *Synechococcus*, and in a group with Gv at the bottom of both trees except for Te that clusters together in clade II of 16s rDNA. It is found that these groups of FNR that correspond to their 16s rDNA phylogeny are mostly based on the tree topology, but S79, S63, and Te distribute in different clades from the 16s rDNA phylogeny. Therefore, most FNR sequences are highly conservative. The appearance of cyanobacterial FNR might be due to the speciation, and did not diversify under selective pressures.

4. Conclusion

The current work on the PBSs linker family facilitates our understanding on biological functions and complicated interactions between linker polypeptides and the PBPs from comparative analysis on 25 cyanobacteria genomes. 192 putative PBSs linker genes have been identified from 25 species of cyanobacteria using BLASTP, TBLASTN and ClustalX. The gene clusters of 36 PBSs linkers ranging from 1.5 kb to 13.2 kb were found in these cyanobacterial genomes. The possibilities of six classes of the linker family were demonstrated. Gene duplication, loss, shuffling, and/or horizontal transfer appear to have played important roles during the evolution and divergence of cyanobacterial PBSs linker polypeptides. Various environmental factors especially light acclimation were the primary selective pressures. Future research will drive the field towards a deeper understanding on evolutional mechanisms of photosynthetic light-harvesting complexes in cyanobacteria and red algae.

Supplementary Material

Unrooted MP tree for PBSs linkers in 25 cyanobacteria and unrooted NJ trees for 16s rDNA and FNR in 21 cyanobacteria [http://www.biolsci.org/v03p0434s1.pdf]

Acknowledgement

This work was supported by the Key Project (KZCX-2-YW-209) of Knowledge Innovation Program of Chinese Academy of Sciences and Hi-Tech Research and Development Program (2006AA090303) of China (863).

Conflict of interest

The authors have declared that no conflict of interest exists.

References

1. Gralnick J., Webb E., Beck B., and Downs DM. Iron Stress in Open-Ocean Cyanobacteria (Synechococcus, Trichodesmium, and Crocosphaera spp): Identification of the IdiA Protein. Appl. Environ. Microbiol. 2000; 67: 5444-5452
2. Ting C.S., Rocap G., King J., Chisholm S.W. Cyanobacterial Photosynthesis in the Oceans: Origins and Significance of Divergent Light-Harvesting Strategies. Trends in Microbiology 2002;10: 134-142
3. Sidler W., Nutt H., Kumpf B., et al. The complete amino-acid sequence and the phylogenetic origin of phycocyanin-645 from the cryptophytan alga Chroomonas sp. Biol Chem Hoppe Seyler. 1990;371(7):537-47
4. Grossman A., Bhaya D., Apt K., Kehoe D. Light-harvesting complexes in oxygenic photosynthesis: diversity, control, and evolution. Annu. Rev. Genet. 1995;29: 231-288
5. Talarico L., Maranzana G. Light and adaptive responses in red macroalgae: an overview. J. Photochem. Photobiol. B Biol. 2000;56: 1-11
6. Rippka R., Waterbury J and Cohen-Bazire G. A cyanobacterium which lacks thylakoids. Arch. Microbiol. 1974;100: 419-436
7. Fujita Y. A study on the dynamic features of photosystem stoichiometry: accomplishments and problems for future studies. Photosynth. Res. 1997;53: 83-93
8. MacColl R. Cyanobacterial phycobilisomes. J Struct Biol. 1998;124: 311-334
9. Gottschalk L., Fischer R., Lottspeich F., Scherer H. Origin of the red shifted absorption in phycocyanin 632 from Mastigocladus

laminosus. Photochem. Photobiol. 1991;54: 283-288
10. Gottschalk L., Lottspeich F., Scherer H. Reconstitution of allophycocyanin from Mastigocladus laminosus with isolated linker polypeptide. Photochem. Photobiol. 1993;58: 761-767
11. Reuter W., Wiegand G., Huber R., Than M.E. Structural analysis at 2.2A of orthorhombic crystals presents the asymmetry of the allophycocyanin-linker complex, Ap.LC7.8 from phycobilisomes of Mastigocladus laminosus. Proc Natl Acad Sci USA. 1999;96: 1363-1368
12. Grossman A.R., Schaefer MR., Chiang G.G., and Collier J.L. The phycobilisome, a light-harvesting complex responsive to environmental conditions. Microbiol. Rev. 1993;57: 725-749
13. Allen J.F., and Matthijs HCP. Complementary adaptations, photosynthesis and phytochrome. Trends Plant Sci. 1997;2: 41-43
14. Sidler WA. Phycobilisome and Phycobiliprotein Structure. In: Bryant DA editor. The Molecular Biology of Cyanobacteria. The Netherlands: Kluwer Academic Publishers. 1994:139-216
15. Bryant D.A., Cohen-Bazire G. Effects of chromatic illumination on cyanobacterial phycobilisomes. Evidence for the specific induction of a second pair of phycocyanin subunits in Pseudanabaena 7409 grown in red light. Eur J Biochem. 1981;119: 415-424
16. de Lorimier R., Bryant D.A., Stevens S.E. Genetic analysis of a 9kDa phycocyanin-associated linker polypeptide. Biochim Biophys Acta. 1990;1019: 29-41
17. Bryant D.A. Cyanobacterial phycobilisomes: progress towards complete structural and functional analysis via molecular genetics. In: Bogorad L, Vasil IK, editors. Cell Culture and Somatic Cell Genetics of Plants. San Diego: Academic Press. 1991:257-300
18. Gantt E. Structure and function of phycobilisomes: light harvesting pigment complexes in red and blue-green algae. Int. Rev. Cytol. 1998;66: 45- 80
19. Glazer A.N. Light harvesting by phycobilisomes. Annu. Rev. BBC. 1985;14: 47- 77
20. Zilinskas B.A., Greenwald L.J. Phycobilisome structure and function. Photosynth Res. 1986;10: 7 - 35
21. Man D., Wang W., Sabehi G, et al. Diversification and spectral tuning in marine proteorhodopsins. EMBO J 2003;22: 1725-1731
22. Steglich C., et al. A green light-absorbing phycoerythrin is present in the high-light-adapted marine cyanobacterium Prochlorococcus sp MEDY. Environ Microbiol. 2005;7: 1611-1618
23. Van Thor J.J., Gruters O.W., Matthijs H.C and Hellingwerf K.J. Localization and function of ferredoxin: NADP (+) reductase bound to the phycobilisomes of Synechocystis. EMBO J. 1999;18: 4128-4136
24. Gomez-Lojero C., Perez-Gomez B., Shen G., et al. Interaction of ferredoxin:NADP+ oxidoreductase with phycobilisomes and phycobilisome substructures of the cyanobacterium Synechococcus sp.strain PCC 7002. Biochemistry. 2003; 42(47):13800-13811
25. Fillat M.F., Flores E, and Gomez-Moreno C. Homology of the N-terminal domain of the petH gene product from Anabaena sp. PCC 7119 to the CpcD phycobilisome linker polypeptide. Plant Mol Biol. 1993;22: 725-729
26. Krogmann D.W., Gómez-Lojero C. The phycocyanin-associated rod linker proteins of the phycobilisomes of Gloeobacter violaceus PCC 7421 contain unusually located rod-capping domains. Biochim. Biophys. Acta. 2006;1757: 130-134
27. Liu L.N., Chen XL., Zhang Y.Z., Zhou B.C. Characterization, structure and function of linker polypeptides in phycobilisomes of cyanobacteria and red algae: An overview. Biochim. Biophys. Acta. 2005;1708: 133-142
28. Apt K.E., Hoffman N.E., Grossman A.R. The gamma subunit of R phycoerythrin and its possible mode of transport into the plastid of red algae. J. Biol. Chem. 1993;268: 16208-16215
29. Egelhoff T and Grossman A. Cytoplasmic and Chloroplast Synthesis of Phycobilisome Polypeptides. Proc. Natl. Acad. Sci. U.S.A. 1983;80: 3339-3343
30. Glauser M., Stirewalt V.L., Bryant D.A., Sidler W., Zuber H. Structure of the genes encoding the rod-core linker polypeptides of Mastigocladus laminosus phycobilisomes and functional aspects of the phycobiliprotein/ linker- polypeptide interactions. Eur J Biochem. 1992;205: 927- 937
31. Zhang J.P., Wan Z.L., et al. Isolation, crystallization and preliminary crystallographic analysis of R-phycocyanin from Polysiphonia urceolata. Acta. Biophys. Sinica. 1995;11: 481-484
32. Mörschel E., Wehrmeyer W and Koller KP. Biliprotein assembly in the disc-shaped phycobilisomes of Rhodella violacea. Electron microscopical and biochemical analysis of B-phycoerythrin-C-phycocyanin aggregates. Eur. J. Cell Biol. 1980;21: 319-327
33. Glazer A.N. Directional energy transfer in a photosynthetic antenna. J. Biol. Chem. 1989;264: 1 - 4
34. Apt KE., Metzner S., Grossman AR. The γ subunits of phycoerythrin from a red alga: position in phycobilisomes and sequence characterization. J Phycol, 2001;37: 64-70
35. Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G. The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 1997;25: 4876- 4882
36. Hall T.A. BioEdit A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 1999;41: 95- 98
37. Sonnhammer ELL, Eddy SR., Birney E., Bateman A., and Durbin R. Pfam: Multiplesequence alignments and HMM-profiles of protein domains. Nucleic Acids Res. 1998;26: 320-322
38. Schultz J., Milpetz F., Bork P., and Ponting C.P. SMART, a simple modular architecture research tool: Identification of signaling domains. Proc Natl Acad Sci USA. 1998;95: 5857-5864
39. Kumar S, Tamura K, and Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 2004;5: 150-163
40. Felsenstein J. PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics 1989;5: 164-166
41. Glazer A.N. Phycobilisomes a macromolecular complex optimized for light energy transfer. Biochim. Biophys. Acta. 1984;768: 29- 51
42. Parbel A., Scheer H. Model for the phycobilisome rod with interlocking disks based on domain-weighted linker polypeptide sequence homologies of Mastigocladus laminosus. Int. J. Photoenergy. 2000;2: 31-40
43. Kondo K., Ochiai Y., Katayama M., and Ikeuchi M. The Membrane- Associated CpcG2-Phycobilisome in Synechocystis: A New Photosystem I Antenna. Plant Physiology, 2007; 144(2): 1200 - 1210
44. Piven I., Ajlani G and Sokolenko A. Phycobilisome linker proteins are phosphorylated in Synechocystis sp PCC6803. J Biol Chem. 1995;280(22): 21667-72
45. Rocap, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N.A., et al. Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation. Nature 2003;424: 1042–1047
46. Six C., Thomas JC, Thion L, Lemoine Y., Zal F., Partensky F. Two Novel Phycoerythrin-Associated Linker Proteins in the Marine Cyanobacterium Synechococcus sp Strain WH8102. J Bacteriol. 2005;187: 1685-1694
47. Palenik B and Haselkorn R. Multiple evolutionary origins of prochlorophytes, the chlorophyll b-containing prokaryotes. Nature 1992;355: 265-267
48. Urbach E., Scanlan D.J., Distel DL, et al. Rapid diversification of marine picophytoplankton with dissimilar light harvesting structures inferred from sequences of Prochlorococcus and Synechococcus (cyanobacteria). J. Mol. Evol., 1998;46: 188-201

49. Rocap G., Distel D.L., Waterbury J.B., and Chisholm S.W. Resolution of Prochlorococcus and Synechococcus ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. Appl Environ Microbiol 2002;68: 1180–1191
50. Hess W.R., Rocap G., Ting C., Larimer F., Lamerdin J., Stilwagon S., Chisholm SW. The photosynthetic apparatus of Prochlorococcus: Insights through comparative genomics. Photosynthesis Res Minireview. 2001;70: 53-71
51. Apt K.E., Collier J.L., Grossman A.R. Evolution of the phycobiliproteins. J Mol Biol. 1995;248: 79-96
52. Wilbanks SM., and Glazer AN. Rod structure of a phycoerythrin II-containing phycobilisome. I. Organization and sequence of the gene cluster encoding the major phycobiliprotein rod components in the genome. J Biol Chem. 1993;268: 1226-1235
53. Cobley J.G, et al. CpeR is an activator required for expression of the phycoerythrin operon (cpeBA) in the chromatically-adapting cyanobacterium, Fremyella diplosiphon. Molecular Microbiology. 2002;44: 1517-1531
54. Nakamura Y., Kaneko T., Sato S, et al. Complete genome structure of Gloeobacter violaceus PCC 7421, a cyanobacterium that lacks thylakoids. DNA Res. 2003;10: 137-145
55. Shen G., et al. Identification and Characterization of a New Class of Bilin Lyase: THE cpcT GENE ENCODES A BILIN LYASE RESPONSIBLE FOR ATTACHMENT OF PHYCOCYANOBILIN TO CYS-153 ON THE beta-SUBUNIT OF PHYCOCYANIN IN SYNECHOCOCCUS SP. PCC 7002. J Biol Chem. 2006;281: 17768-17778
56. Zhao KH., et al. Chromophore Attachment to Phycobiliprotein beta-Subunits phycocyanobilin: cysteine-beta 84 phycobiliprotein lyase activity of CpeS-like protrin from Anabaena sp. PCC7120. J Biol Chem. 2006;281: 8573-8581
57. Zhao F and Qin S. Evolutionary Analysis of Phycobiliproteins: Implications for Their Structural and Functional Relationships. J Mol Evol. 2006;63: 330–340
58. Kondo K., Geng X.X., Katayama M., Ikeuchi M. Distinct roles of CpcG1 and CpcG2 in phycobilisome assembly in the cyanobacterium Synechocystis sp PCC 6803. Photosynth Res 2005;84: 269-273