EUKARYOTIC CELL, Mar. 2008, p. 550–553 Vol. 7, No. 3
1535-9778/08/$08.00+0 doi:10.1128/EC.00396-07

Evolutionary Dynamics of Light-Independent Protochlorophyllide Oxidoreductase Genes in the Secondary Plastids of Cryptophyte Algae[▽]

Anna Fong and John M. Archibald*

The Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada

Received 29 October 2007/Accepted 24 December 2007

Plastid genes encoding light-independent protochlorophyllide oxidoreductase (LIPOR) subunits were isolated from cryptophyte algae, the first example of such genes in plastids of secondary endosymbiotic origin. The presence of functional and nonfunctional copies of LIPOR genes in cryptophytes suggests that light-independent chlorophyll biosynthesis is a nonessential pathway in these organisms.

The biosynthesis of chlorophyll is an indispensable process in photosynthetic prokaryotes and eukaryotes. In plants and algae, chlorophyll synthesis occurs in the plastid (chloroplast) and involves two evolutionarily distinct enzymes, both of which catalyze the reduction of protochlorophyllide to chlorophyllide (30). Light-dependent protochlorophyllide oxidoreductase is ubiquitous among eukaryotic phototrophs and is comprised of nucleus-encoded subunits that are targeted to the plastid post-translationally (3). In contrast, light-independent (or "dark active") protochlorophyllide oxidoreductase (LIPOR or DPOR) synthesizes chlorophyll in the absence of light and is not universally distributed (33, 35). When present, LIPOR is comprised of three distinct plastid-encoded subunits, ChlL, ChlB, and ChlN. Eukaryotic light-dependent protochlorophyllide oxidoreductase and LIPOR are both derived from prokaryotic enzymes and entered the eukaryotic cell via the cyanobacterial progenitor of the plastid (many, though not all, modern-day cyanobacteria possess both systems [34, 37]). Genes for LIPOR subunits have been characterized for a number of "primary" plastid-containing eukaryotes, including green algae (9, 31), mosses (e.g., a *Marchantia* sp. [15]), the glaucophyte *Cyanophora paradoxa* (32), and the red alga *Porphyra purpurea* (29), but are absent in the plastid genome of the red alga *Cyanidioschyzon merolae* (23). Angiosperms have also lost LIPOR genes, although some species are capable of synthesizing chlorophyll in the dark using an as yet unknown mechanism (1). This patchy distribution raises questions about the timing and frequency of LIPOR gene/enzyme loss during the evolution of photosynthetic eukaryotes.

Little is known about chlorophyll biosynthesis in "secondary" plastid-containing algae, such as haptophytes, heterokonts (e.g., diatoms and kelp), dinoflagellates, and cryptophytes, lineages that acquired photosynthesis by the endosymbiotic uptake of a red alga (2, 13, 25). Interestingly, the plastid genomes of three diatoms (*Thalassiosira pseudonana*, *Phaeodactylum tricornutum*, and *Odontella sinensis* [16, 24]), the haptophyte *Emiliania huxleyi* (28), and the cryptophyte *Guillardia theta* (10) all lack LIPOR genes. We were thus surprised when *chlL*, *chlB*, and *chlN* pseudogenes were found in the plastid genome of another cryptophyte, *Rhodomonas salina* CCMP1319 (14). Random sequencing of a genomic library from another cryptophyte, *Hemiselmis andersenii* CCMP644, uncovered a plastid-derived clone containing a fragment of *chlL*. To better understand the origin and distribution of LIPOR genes in cryptophytes, we used PCR to search for the presence of *chlL*, *chlB*, and *chlN* genes in a broad range of cryptophyte species.

Alignments of ChlL, ChlB, and ChlN proteins covering the known breadth of prokaryotic and eukaryotic diversity were constructed, and degenerate PCR primers were designed against evolutionarily conserved regions near the amino and carboxy termini of each protein (primer sequences available upon request). PCR products corresponding to some or all of the *chl* genes were successfully amplified from total cellular DNA, cloned, and sequenced as described previously (18, 19) for the species listed in Table 1 (between 3 and 10 independent clones were sequenced per amplicon). Interestingly, the *chlN*

TABLE 1. LIPOR genes in cryptophyte algae

Organism	Presence of indicated LIPOR gene[a]		
	chlB	*chlL*	*chlN*
Rhodomonas salina CCMP1319	ψ[b]	ψ[b]	ψ[b]
Hemiselmis andersenii CCMP644	+[c]	+[d]	+
Hemiselmis tepida CCMP443	?	+	+
Hemiselmis cryptochromatica CCMP1181	?	?	?
Chroomonas pauciplastida nom. prov. CCMP268	+[c]	+	+
Chroomonas mesostigmatica CCMP1168	?	+	ψ

[a] GenBank accession numbers for the new sequences are EU233747 to EU233756. +, present; ?, unable to amplify using PCR; ψ, pseudogene.
[b] Pseudogenes are present in the complete plastid genome sequences (see the text).
[c] The sequences contain group II introns.
[d] The coding sequence possesses an alternate start codon (GUG).

* Corresponding author. Mailing address: Dalhousie University, Department of Biochemistry and Molecular Biology, 5850 College Street, Sir Charles Tupper Medical Building, Halifax B3H 1X5, Canada. Phone: (902) 494-2536. Fax: (902) 494-1355. E-mail: john.archibald@dal.ca.
▽ Published ahead of print on 4 January 2008.

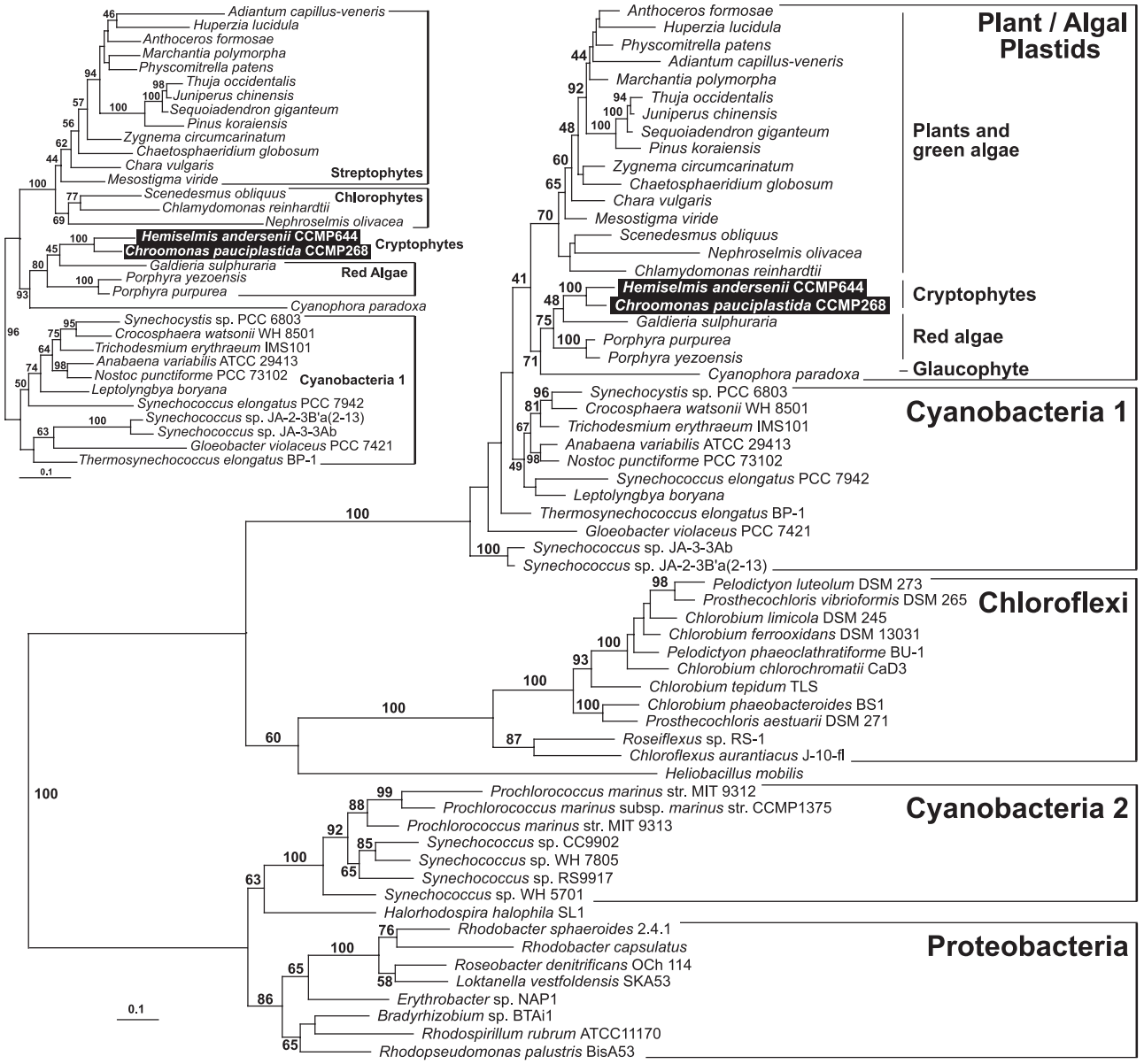

FIG. 1. Maximum likelihood phylogenetic tree of LIPOR proteins. Concatenated ChlL, ChlB, and ChlN proteins from plants and algae were analyzed together with their bacterial counterparts (BchL, BchB, and BchN). The alignment contained 61 sequences and 777 unambiguously aligned amino acid positions. The tree was constructed using IQPNNI (see the text) and is shown arbitrarily rooted with proteobacteria and their closest cyanobacterial homologs. The cryptophyte sequences determined in this study are highlighted. PhyML bootstrap values (100 replicates) are provided where they are >40%. The smaller phylogeny (upper left) was produced from a set of 33 sequences with the plant/algal sequences rooted only with the most closely related cluster of cyanobacterial sequences. This analysis included 847 amino acid positions. Scale bars indicate the inferred number of amino acid substitutions per site.

gene fragment from the cryptophyte *Chroomonas mesostigmatica* possesses numerous stop codons and frameshifts and thus appears to be a pseudogene (see below), as seen previously in *Rhodomonas salina* (14). The *H. andersenii* and *Chroomonas pauciplastida chlB* genes each possess group II introns but in different locations. The structures and biochemical properties of these introns will be presented elsewhere.

In order to elucidate the origin of the cryptophyte LIPOR genes and, more generally, to infer the evolutionary history of the LIPOR gene family as a whole, plant and algal ChlL, ChlB, and ChlN protein sequences were analyzed individually (data not shown) and in concatenation (Fig. 1) in the context of diverse bacterial homologs (BchL, BchB, and BchN) using two maximum-likelihood phylogenetic method programs, PhyML (12) and IQPNNI (36). Plastid-encoded ChlL, ChlB, and ChlN proteins typically form a monophyletic group in such analyses, albeit with low statistical support when the proteins are analyzed individually (data not shown). Two different forms of BchL, BchB, and BchN exist in cyanobacteria, with the plant and algal sequences nested within the "cyanobacteria 1" clade (Fig. 1), a group that includes *Nostoc*, *Trichodesmium*, *Thermosynechococcus*, and *Gloeobacter* species. The second cluster

of cyanobacterial sequences, very distinct from the first, is comprised of a variety of *Prochlorococcus* and marine *Synechococcus* species. Additional prokaryotic sequences fall mainly into two clusters, the Chloroflexi and a diverse group of proteobacteria. Sequences from *Heliobacillus mobilis* and *Halorhodospira halophila* showed no strong affinity for any particular group (Fig. 1), though they are clearly well separated from "cyanobacteria 1" and plant/algal sequences. The origin of photosynthesis (and LIPOR genes) in noncyanobacterial prokaryotes is believed to be the result of lateral gene transfer, although the exact details remain contentious (see references 8, 22, and 37 and the references therein for a discussion). The significance of the deep divergence seen in the cyanobacterial LIPOR genes is unclear at present, although it could be the result of an ancient paralogy followed by differential gene loss. However, it is significant that intra- and interphylum lateral gene transfer has been a major factor in the evolution of cyanobacterial genomes (6, 38), with some genes in *Prochlorococcus* and *Synechococcus* being most closely related to proteobacteria (38). Additional support for the uniqueness of the "cyanobacteria 2" clade comes from the recent discovery of a hyperconserved, lineage-specific protein of unknown function shared between members of this group (39).

Concatenated phylogenies of LIPOR proteins reveal that the cryptophyte sequences are most closely related to those of red algae (Fig. 1), as predicted based on the inferred origin of the plastid in this lineage (2, 7, 10, 14). Bootstrap support for the common branch uniting the red algal and cryptophyte sequences was moderate: 75% using the full data set and 80% using a taxon-restricted alignment with more sites and only the "cyanobacteria 1" sequences as outgroups (Fig. 1). The glaucophyte alga *Cyanophora paradoxa* branched next to the cryptophyte/red algal sequences to the exclusion of the plant/green algal clade in both analyses and with both methods.

From a molecular evolutionary perspective, the LIPOR enzymes represent an interesting paradox. They are extraordinarily highly conserved (e.g., the cryptophyte ChlB proteins share >70% amino acid identity over >500 amino acids with their closest cyanobacterial homologs), yet they appear to be "optional" in photosynthetic eukaryotes. A comparative analysis of complete plastid genome sequences indicates that LIPOR genes were lost from plastids at some point during the early evolution of angiosperms (20, 21), and an analysis of LIPOR proteins in species of conifer shows evidence of relaxed selection and loss of enzyme activity (17). The discovery of genes for LIPOR subunits in the plastid genomes of diverse cryptophyte algae suggests that these genes have been lost relatively recently in some members of this lineage (e.g., *Guillardia theta*) but retained in others. Indeed, the presence of *chl* pseudogenes in *Rhodomonas salina* CCMP1319 (14) and *Chroomonas mesostigmatica* (Table 1) suggests that while these genes/proteins were very likely inherited directly from the red algal plastid that gave rise to the cryptophyte organelle, they are currently in a state of flux, presumably allowing some cryptophytes to synthesize chlorophyll in the dark but not others. Thus, it will be important to search for the presence of LIPOR enzymes in additional cryptophyte species, as well as in members of the haptophytes and heterokonts, once more plastid genome sequences become available. Given that recent phylogenomic analyses indicate that cryptophytes and haptophytes are each other's closest relatives (26), then it can be inferred that if haptophytes and heterokonts do indeed lack LIPOR genes, the two groups lost them independently of one another.

Interestingly, recent expressed sequence tag surveys have shown the presence of expressed and presumably functional copies of *chlL* genes in the nuclear genomes of the red algal secondary plastid-containing dinoflagellates *Heterocapsa triquetra* (27) and *Amphidinium carterae* (4). Although truncated, both sequences have obvious N-terminal extensions, with the *A. carterae* extension predicted to be a chloroplast transit peptide using ChloroP (11) and iPSORT (5). We used Southern hybridization to explore the possibility that some of the PCR-amplified cryptophyte *chl* genes presented in this study are in fact nucleus borne. Using a *chlL* gene probe against chromosomes separated by pulsed-field gel electrophoresis yielded hybridization patterns suggesting that copies of this gene exist in both the plastid and nuclear genomes of *Chroomonas pauciplastida*, as well as in three members of the genus *Cryptomonas* (data not shown). No evidence for a nuclear copy of *chlL* was found in *G. theta* or *H. andersenii*. Thus, it is possible that some of the sequences presented in Table 1, including the *chlN* pseudogene of *C. mesostigmatica*, reside in the nuclear genome. If so, these genes have been transferred very recently, as their A+T contents and codon biases are indistinguishable from those of bona fide cryptophyte plastid genes. Collectively, these observations underscore the dynamic nature of LIPOR gene evolution in cryptophytes and other secondary plastid-containing algae. Combined with laboratory experiments aimed at determining the presence/absence of light-independent chlorophyll biosynthesis in cryptophytes, complete plastid and nuclear genome sequences from more diverse algal species should improve our understanding of the distribution and biochemical significance of LIPOR genes even further.

Nucleotide accession numbers. The new sequences have been assigned GenBank accession numbers EU233747 to EU233756.

We thank O. Zhaxybayava, C. Lane, H. Khan, and two anonymous reviewers for helpful comments on an earlier version of the paper. We also appreciate the assistance of Catherine Kozera and Sharen Bowman.

This work was supported by Genome Atlantic and an NSERC discovery grant (283335-2004) awarded to J.M.A. J.M.A. is a scholar of the Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity.

REFERENCES

1. **Adamson, H. Y., R. G. Hiller, and J. Walmsley.** 1997. Protochlorophyllide reduction and greening in angiosperms: an evolutionary perspective. J. Photochem. Photobiol. B **41:**201–221.
2. **Archibald, J. M., and P. J. Keeling.** 2002. Recycled plastids: a green movement in eukaryotic evolution. Trends Genet. **18:**577–584.
3. **Aronsson, H., C. Sundqvist, and C. Dahlin.** 2003. POR hits the road: import and assembly of a plastid protein. Plant Mol. Biol. **51:**1–7.
4. **Bachvaroff, T. R., G. T. Concepcion, C. R. Rogers, E. M. Herman, and C. F. Delwiche.** 2004. Dinoflagellate expressed sequence tag data indicate massive transfer of chloroplast genes to the nuclear genome. Protist **155:**65–78.
5. **Bannai, H., Y. Tamada, O. Maruyama, K. Nakai, and S. Miyano.** 2002. Extensive feature detection of N-terminal protein sorting signals. Bioinformatics **18:**298–305.
6. **Beiko, R. G., T. J. Harlow, and M. A. Ragan.** 2005. Highways of gene sharing in prokaryotes. Proc. Natl. Acad. Sci. USA **102:**14332–14337.
7. **Bhattacharya, D., H. S. Yoon, and J. D. Hackett.** 2003. Photosynthetic eukaryotes unite: endosymbiosis connects the dots. Bioessays **26:**50–60.
8. **Bryant, D. A., and N.-U. Frigaard.** 2006. Prokaryotic photosynthesis and phototrophy illuminated. Trends Microbiol. **14:**488–496.

9. **Choquet, Y., M. Rahire, J. Girard-Bascou, J. Erickson, and J. D. Rochaix.** 1992. A chloroplast gene is required for the light-independent accumulation of chlorophyll in *Chlamydomonas reinhardtii*. EMBO J. **11:**1697–1704.
10. **Douglas, S. E., and S. L. Penny.** 1999. The plastid genome from the cryptomonad alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J. Mol. Evol. **48:**236–244.
11. **Emanuelsson, O., H. Nielsen, and G. von Heijne.** 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. **8:**978–984.
12. **Guindon, S., and O. Gascuel.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52:**696–704.
13. **Keeling, P.** 2004. A brief history of plastids and their hosts. Protist **155:**3–7.
14. **Khan, H., N. Parks, C. Kozera, B. A. Curtis, B. J. Parsons, S. Bowman, and J. M. Archibald.** 2007. Plastid genome sequence of the cryptophyte alga *Rhodomonas salina* CCMP1319: lateral transfer of putative DNA replication machinery and a test of chromist plastid phylogeny. Mol. Biol. Evol. **24:**1832–1842.
15. **Kohchi, T., H. Shirai, H. Fukuzawa, T. Sano, T. Komano, K. Umesono, H. Inokuchi, H. Ozeki, and K. Ohyama.** 1988. Structure and organization of *Marchantia polymorpha* chloroplast genome. IV. Inverted repeat and small single copy regions. J. Mol. Biol. **203:**353–372.
16. **Kowallik, K. V., B. Stoebe, I. Schaffran, P. Kroth-Pancic, and U. Freier.** 1995. The chloroplast genome of a chlorophyll a+c-containing alga, *Odontella sinensis*. Plant Mol. Biol. Rep. **13:**336–342.
17. **Kusumi, J., A. Sato, and H. Tachidi.** 2006. Relaxation of functional constraint on light-independent protochlorophyllide reductase in *Thuja*. Mol. Biol. Evol. **23:**941–948.
18. **Lane, C. E., and J. M. Archibald.** 2006. Novel nucleomorph genome architecture in the cryptomonad genus *Hemiselmis*. J. Eukaryot. Microbiol. **53:** 515–521.
19. **Lane, C. E., H. Khan, M. MacKinnon, A. Fong, S. Theophilou, and J. M. Archibald.** 2006. Insight into the diversity and evolution of the cryptomonad nucleomorph genome. Mol. Biol. Evol. **23:**856–865.
20. **Martin, W., T. Rujan, E. Richly, A. Hansen, S. Cornelsen, T. Lins, D. Leister, B. Stoebe, M. Hasegawa, and D. Penny.** 2002. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc. Natl. Acad. Sci. USA **99:**12246–12251.
21. **Martin, W., B. Stoebe, V. Goremykin, S. Hansmann, M. Hasegawa, and K. V. Kowallik.** 1998. Gene transfer to the nucleus and the evolution of chloroplasts. Nature **393:**162–165.
22. **Mulkidjanian, A. Y., E. V. Koonin, K. S. Makarova, S. L. Mekhedov, A. Sorokin, Y. I. Wolf, A. Dufresne, F. Partensky, H. Burd, D. Kaznadzey, R. Haselkorn, and M. Y. Galperin.** 2006. The cyanobacterial genome core and the origin of photosynthesis. Proc. Natl. Acad. Sci. USA **103:**13126–13131.
23. **Ohta, N., M. Matsuzaki, O. Misumi, S. Y. Miyagishima, H. Nozaki, K. Tanaka, I. T. Shin, Y. Kohara, and T. Kuroiwa.** 2003. Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*. DNA Res. **10:**67–77.
24. **Oudot-Le Secq, M.-P., J. Grimwood, H. Shapiro, E. V. Armbrust, C. Bowler, and B. R. Green.** 2007. Chloroplast genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*: comparison with other plastid genomes of the red lineage. Mol. Genet. Genomics **277:**427–439.
25. **Palmer, J. D.** 2003. The symbiotic birth and spread of plastids: how many times and whodunit? J. Phycol. **39:**4–12.
26. **Patron, N. J., Y. Inagaki, and P. J. Keeling.** 2007. Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. Curr. Biol. **17:**887–891.
27. **Patron, N. J., R. F. Waller, J. M. Archibald, and P. J. Keeling.** 2005. Complex protein targeting to dinoflagellate plastids. J. Mol. Biol. **348:**1015–1024.
28. **Puerta, M. V., T. R. Bachvaroff, and C. F. Delwiche.** 2005. The complete plastid genome sequence of the haptophyte *Emiliania huxleyi*: a comparison to other plastid genomes. DNA Res. **12:**151–156.
29. **Reith, M. E., and J. Munholland.** 1995. Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Mol. Biol. Rep. **13:**333–335.
30. **Schoefs, B., and F. Franck.** 2003. Protochlorophyllide reduction: mechanisms and evolution. Photochem. Photobiol. **78:**543–557.
31. **Shi, C., and X. Shi.** 2006. Characterization of three genes encoding the subunits of light-independent protochlorophyllide reductase in *Chlorella prototothecoides* CS-41. Biotechnol. Prog. **22:**1050–1055.
32. **Stirewalt, V. L., C. B. Michalowski, W. Loffelhardt, H. J. Bohnert, and D. A. Bryant.** 1995. Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*. Plant Mol. Biol. Rep. **13:**327–332.
33. **Suzuki, J. Y., and C. E. Bauer.** 1992. Light-independent chlorophyll biosynthesis: involvement of the chloroplast gene chlL (frxC). Plant Cell **4:**929–940.
34. **Suzuki, J. Y., and C. E. Bauer.** 1995. A prokaryotic origin for the light-dependent chlorophyll biosynthesis of plants. Proc. Natl. Acad. Sci. USA **92:**3749–3753.
35. **Suzuki, J. Y., and C. E. Bauer.** 1997. Genetic analysis of chlorophyll synthesis. Annu. Rev. Genet. **31:**61–89.
36. **Vinh, L. S., and A. Von Haeseler.** 2004. IQPNNI: moving fast through tree space and stopping in time. Mol. Biol. Evol. **21:**1565–1571.
37. **Xiong, J., W. M. Fischer, K. Inoue, M. Nakahara, and C. E. Bauer.** 2000. Molecular evidence for the early evolution of photosynthesis. Science **289:** 1724–1730.
38. **Zhaxybayeva, O., J. P. Gogarten, R. L. Charlebois, W. F. Doolittle, and R. T. Papke.** 2006. Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. Genome Res. **16:**1099–1108.
39. **Zhaxybayeva, O., J. P. Gogarten, and W. F. Doolittle.** 2007. A hyperconserved protein in *Prochlorococcus* and marine *Synechococcus*. FEMS Microbiol. Lett. **274:**30–34.