APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Sept. 1997, p. 3600–3606 Vol. 63, No. 9
0099-2240/97/$04.00+0

Diversity of the Ribulose Bisphosphate Carboxylase/Oxygenase Form I Gene (*rbcL*) in Natural Phytoplankton Communities

S. L. PICHARD,[1] L. CAMPBELL,[2] AND J. H. PAUL[1]*

Department of Marine Science, University of South Florida, St. Petersburg, Florida 33701,[1] and Department of Oceanography, Texas A&M University, College Station, Texas 77843[2]

Received 25 March 1997/Accepted 19 June 1997

The phytoplankton of the world's oceans play an integral part in global carbon cycling and food webs by conversion of carbon dioxide into organic carbon. They accomplish this task through the action of the Calvin cycle enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO). Here we have investigated the phylogenetic diversity in the form I *rbcL* locus in natural phytoplankton communities of the open ocean and representative clones of marine autotrophic picoplankton by mRNA or DNA amplification and sequencing of a 480 to 483 bp internal fragment of this gene. Five gene sequences were recovered from nucleic acids of natural phytoplankton communities of the Gulf of Mexico. The *rbcL* genes of two *Prochlorococcus* isolates and one *Synechococcus* strain (WH8007) were also sequenced. Sequences were aligned with the database of *rbcL* genes and subjected to both neighbor-joining and parsimony analyses. The five sequences from the natural phytoplankton community spanned nearly the entire diversity of characterized form I *rbcL* genes, with some sequences closely related to isolates such as *Synechococcus* and *Prochlorococcus* (forms IA and I) and prymnesiophyte algae (form ID), while other sequences were deeply rooted. Unexpectedly, the deep euphotic zone contained an organism that possesses a transcriptionally active *rbcL* gene closely related to that of a recently characterized manganese-oxidizing bacterium, suggesting that such chemoautotrophs may contribute to the diversity of carbon-fixing organisms in the marine euphotic zone.

Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) catalyzes the first, rate-limiting step in the Calvin cycle, the primary pathway for photosynthetic carbon reduction in the oceans. This enzyme has been found in two natural forms, which differ in their structures and primary sequences, their ability to fix carbon under varying oxygen tensions, and their evolution (52). Form I is composed of eight large and eight small subunits (L_8S_8), encoded by *rbcL* and *rbcS*, respectively, while form II is composed of large subunits only, usually as a dimer (L_2). The two forms are only ~23% similar based on inferred amino acid sequence data (20, 39). RuBisCO is one of the most well studied enzymes, mainly due to its agricultural importance (1). It has also received much interest as a phylogenetic marker for studies of the evolution of land plants (9). *rbcL* genes have also been recovered by PCR from fossil deposits (17, 19) and from phytoplankton communities of temperate (55), subtropical (31), and Antarctic (6) freshwater lakes. PCR is now considered a standard tool for investigations into microbial diversity in the sea (10, 15, 16, 56).

Recently there has been much interest in the diversity of the marine picoplankton, using studies of both isolates (25, 26) and natural populations (24). Productivity and phytoplankton abundance in the marine euphotic zone are dominated by the picoplankton-size organisms *Synechococcus* (36), *Prochlorococcus* (7), and various chromophytic algae (4, 36). These organisms are now recognized to be genetically diverse (25, 46). The genus *Synechococcus* is loosely defined as a provisional assemblage with an extremely wide range of 39 to 71 mol% G+C (49). Analysis of the RNA polymerase gene *rpoC1* (25) and 16S rRNA (46) has resulted in the division of the marine *Synechococcus* and *Prochlorococcus* organisms into at least two distinct clusters within the cyanobacteria. Studies of the *rbcL* gene in both the marine *Synechococcus* strain WH7803 (51) and the marine *Prochlorococcus* strain GP2 (41) have placed these organisms in the group with form IA *rbcL* genes, which is dominated by autotrophic bacteria. Similarly, the chromophyte algae are a diverse group consisting of an estimated one million extant species representing 13 taxonomic classes (40). This suggests that the marine autotrophic picoplankton may be more genetically diverse than was previously thought.

Through measurements of levels of *rbcL* mRNA, the *rbcL* gene has been shown to be actively expressed in marine phytoplankton communities (28, 30, 33). Surface water communities and picoplankton isolates have displayed diel rhythms in *rbcL* gene expression (32, 34) that show a high degree of correlation with levels of carbon fixation. The expression of several types of form I *rbcL* sequences was found to correlate with the occurrence of specific phytoplankton types in stratified water columns (35). Therefore, phytoplankton containing diverse *rbcL* sequences reside and express this gene on different temporal and spatial scales in the water column. Hence, it is necessary to study both the pattern of expression and the sequence being expressed in order to understand the molecular dynamics of water column carbon fixation. In this study we examined the types of *rbcL* genes present at select sites in the Gulf of Mexico, and in autotrophic picoplankton isolates, in an attempt to understand the in situ genetic diversity in this locus as it relates to the regulation of carbon fixation in the oceans.

MATERIALS AND METHODS

Sampling sites. Two stations in the eastern Gulf of Mexico were sampled (1,000 ml each), an open-ocean station (GOMst4; August 1994) located at 25°40′N, 84°35′W and a west Florida continental shelf station (GOMst8; September 1993) located at 26°30′N, 83°50′W.

Culturing of photoautotrophic picoplankton. *Prochlorococcus marinus* MED, *Prochlorococcus* sp. Pacific, and *Synechococcus* sp. CCMP836 (WH8007) were grown in either K/10 (−Cu) medium (8) or SN medium (50) as previously described (35) for 14 days. Cultures were harvested by centrifugation at 3,832 × *g* (Sorvall GS-3 rotor) for 20 min. The cell pellet was resuspended in 1/20 volume

* Corresponding author. Phone: (813) 893-9130. Fax: (813) 893-9189. E-mail: jpaul@marine.usf.edu.

of culture medium and recentrifuged at 7,796 × *g* (Sorvall SS-34 rotor). Cell pellets were stored frozen at −80°C until DNA extraction could be performed.

mRNA and DNA isolation. The extraction of *rbcL* mRNA was accomplished by a combination of guanidinium-isocitrate-phenol extraction coupled with bead beating as previously described (29, 33). The diethylpyrocarbonate treatment of water samples prior to filtration was omitted. *rbcL* DNA extractions from environmental samples were performed by boiling lysis in sodium chloride-Tris-EDTA (pH 8) containing 1% sodium dodecyl sulfate as previously described (29, 33). Total DNA from all procaryotic cell pellets was isolated according to the method of Wood and Townsend (54) with ultracentrifugation omitted.

RT-PCR and PCR amplification. Amplification of *rbcL* (480 to 483 bp) targets from mRNA (reverse transcriptase PCR [RT-PCR]) and DNA (PCR) isolated from microbial biomass at stations in the Gulf of Mexico was performed according to the method of Becker-André (2) and Pichard et al. (33) with *Taq* DNA polymerase (Promega Corp., Madison, Wis.). mRNA was first converted to DNA with the following reaction mixture: 50 mM Tris-HCl (pH 7.5), 75 mM KCl, 10 mM dithiothreitol, 3 mM $MgCl_2$, 100 ng of 3′ primer, and 0.5 mM (each) deoxynucleotide triphosphates in a total volume of 20 μl. The reaction mixture was heated to 65°C for 5 min and then cooled to room temperature. RNasin (6 U; Promega Corp.) and reverse transcriptase (40 U; Life Sciences, Inc., St. Petersburg, Fla.) were added, and the reaction was continued for 30 min at 42°C. The reaction mixture volume was then brought to 50 μl with Tris-EDTA (pH 8.0). Five microliters of the cDNA was then amplified by PCR. Extracted DNA was resuspended in 50 μl of Tris-EDTA (pH 8.0) buffer, and 5 μl was used for PCR amplification as previously described (33). Temperature cycling for all PCRs was as follows: 94°C (5 min); 39 cycles of 94°C (1 min), 50°C (2 min), and 72°C (3 min); and 72°C (10 min) final extension. Two sets of degenerate primers (37, 56), designated Chromo and 168/328, were used in amplification. Primers were designed to complement conserved sites on the *rbcL* gene. Each primer contained a restriction enzyme recognition site that was used for cloning and was either GG or CTC clamped on the 5′ end (18). 3′ primers were targeted to the form I conserved amino acid sequence 324-VVGKLE-329, and 5′ primers were targeted to the universally conserved amino acid sequence 172-KPKLGL-178 (sequence positions are as for *rbcL* of *Spinacia oleracea*). Electrophoretic analysis of the PCR was performed through 2% Nusieve-GTG agarose gels (FMC Bioproducts, Rockland, Maine) stained with ethidium bromide at 10 V/cm. The two sets of primers were as follows. The Chromo primer set consisted of Chromo 5′ primer (29-mer [128-fold]; *Bam*HI), 5′-GGGGATCCAAA(AG)CC(TA)AA(AG)(TC)TAGG(TG)(CT)T(AT)TC-3′, and Chromo 3′ primer (29-mer [96-fold degeneracy]; *Mlu*I), 5′-GGACGCGTACC(TC)TC(TC)A(AG)(TC)TTACC(TA)AC(GAT)AC-3′. The 168/328 primer set consisted of D168 5′ primer (30-mer [64 fold]; *Bam*HI), 5′-CTCGGATCCT(AG)T(CT)AAACC(AT)AAA(CT)T(CA)GGT(CT)T-3′, and D328 3′ primer (30-mer [144 fold]; *Eco*RI), 5′-CTCGAATTCTCCTTC(CTA)A(GA)(TC)TT(AG)CC(TG)AC(GTA)AC-3′. (underlined areas are restriction sites [*Bam*HI]).

Cloning of PCR amplification products. PCR products were cloned into one of three different vectors for DNA sequencing (M13mp18 and M13mp19, pSPORT [Life Technologies, Gaithersburg, Md.], or TA cloned into pCR [Invitrogen, Carlsbad, Calif.]). TA-cloned products were subcloned into M13mp18 and M13mp19. For cloning into M13 or pSPORT, three PCRs were pooled and extracted once with phenol-chloroform, then with chloroform, and then precipitated with 2 volumes of cold 100% ethanol and 1/10 volume of sodium acetate (pH 4.8). The PCR products were then resuspended in distilled H_2O and cut with the appropriate restriction enzymes. Restricted PCR products were then gel purified through 1% low-melting-point agarose and cloned into M13 or pSPORT according to instructions supplied by the manufacturer. TA cloning was performed per the manufacturer's instructions.

Variant screening and sequencing. Cloned *rbcL* amplicons were screened for variation by use of tetrameric restriction endonucleases. *Hae*III (GGCC), *Alu*I (AGCT), *Tru*9I (TTAA), and *Mbo*I (GATC) were used for screening sequence variants (21). Sequencing was performed from single-stranded templates by the Sanger dideoxy method with the Sequenase version 2.0 DNA sequencing kit (Amersham Life Sciences, Inc., Arlington Heights, Ill.). For sequencing clones from the pSPORT phagemids, single-stranded DNA was produced according to the protocol provided by the manufacturer (Life Technologies, Inc.) with the helper phage M13K07.

Phylogenetic reconstruction. Deduced amino acid sequences were aligned with Clustal W, and phylogenetic analysis was performed with programs (Seqboot, Protdist, Neighbor, Protpars, and Consense) in PHYLIP version 3.5c (13, 14). Trees were constructed with the tree-drawing program Treeview version 1.1 (23). Analyses were performed with inferred amino acid sequences (11) to compensate for site saturation with the deep phylogenetic divergences involved (53) and lineage-specific compositional (G+C) bias (43).

Nucleotide sequence accession numbers. Form I *rbcL* sequences used in the analysis were retrieved from the National Center for Biotechnology Information and GenBank under the following accession numbers: *Rhodobacter sphaeroides*, M64624; *Xanthobacter flavus*, X17252; Mn-oxidizing bacterium, L32182; *Cryptomonas* sp. Φ, X62349; *Emiliania huxleyi*, D45845; *Pleurochrysis carterae*, D11140; *Chrysochromulina hirta*, D45846; *Odontella sinensis*, Z67753; *Cylindrotheca* sp., M59080; *Ectocarpus siliculosus*, X52503; *Heterosigma akashiwo*, X61918; *Nitrobacter vulgaris*, L22885; *Hydrogenovibrio marinus*, D43622; *Prochlorococcus* sp. GP2, D21822; *Chromatium vinosum*, M26396; *Synechococcus* sp. PCC6301; J01536; *Anabaena* sp. PCC7120, L02520; *Spinacia oleracea*, J01443; *Micromonas pusilla*, U30276; *Chlorella ellipsoidea*, M20655; *Tetraselmis* sp., U30284; *Prochloron* sp., D21834; *Synechococcus* sp. PCC7002, D13971; and *Prochlorothrix hollandica*, X57359. *Porphyridium aerugineum rbcL* was taken from Valentin and Zetsche (47). The eight new *rbcL* clones presented in this study are available from GenBank under the following accession numbers: GOMst8 mRNA 5m, U93860; GOMst4 mRNA 80m, U93861; GOMst4 mRNA 50m, U93856; GOMst4 DNA1 5m, U93290; GOMst4 DNA2 5m, U93855; *Prochlorococcus* sp. Pacific, U93858; *Prochlorococcus marinus* MED, U93857; and *Synechococcus* sp. WH8007, U93859.

RESULTS

A search of the GenBank nucleotide database on 25 February 1997 yielded 2,252 *rbcL* sequences, most of which were of terrestrial plant origin (1,785). The second-largest grouping of sequences was from the marine rhodophyta (312), with other marine *rbcL* genes comprising ~1% of the total. Selected full-length and near-full-length *rbcL* genes were aligned, and phylogenetic relationships were constructed by neighbor joining (Fig. 1) or maximum parsimony (data not shown). These form I *rbcL* sequences clustered into four major groups designated forms IA to -D (44). Form IA contains mainly proteobacteria, although it also includes a marine *Prochlorococcus* sequence from a Pacific Ocean isolate (41). Form IB contains the cyanobacteria, the chlorophyte algae, and higher plants. Form IC represents another proteobacterial group, which experiences a loss of RuBisCO activity during photosynthesis due to the noncatalytic binding of RuBP and synthesis of the inhibitor xylulose-1,5-bisphosphate (XuBP), a condition known as fallover (45). Fallover also occurs with some form IB and ID enzymes. Form ID includes the chromophyte algae, which exhibit extremely high specificity factors favoring carboxylation. In cases where their properties have been examined, these different types display distinct and unique enzymatic properties, such as different affinities for their CO_2 substrates and differential specificity for carboxylation over the oxygenation reaction (38, 44).

Using two sets of primers (Chromo and 168/328), we recovered *rbcL* genes from water samples from several stations and depths in the Gulf of Mexico as well as from *Prochlorococcus* and *Synechococcus* clones collected in the Pacific Ocean and the Mediterranean Sea and from the Atlantic Ocean, respectively. PCR appeared to recover a limited number of operational taxonomic units (OTUs) as determined by tetrameric restriction endonuclease restriction fragment length polymorphisms of the amplified genes. Amplifications yielded at the most one or two OTUs. In most cases two clones were sequenced from each OTU. These were found to be identical, and the results of the screening are presented in Table 1. The percent identities between full-length and PCR-length (160 to 161 residues) sequences are presented in Table 2. The G+C content of the recovered clones ranged from 37 to 56 mol%.

Station GOMst4 yielded two sequences amplified by RT-PCR from mRNA. A deep euphotic zone sequence from 80 m was most closely aligned (87% identical) with a recently discovered manganese-oxidizing bacterium (5), while a midwater sequence from 50 m clustered (92% identical [Fig. 2]) with a marine *Prochlorococcus* sp. GP2 form IA sequence from the Pacific Ocean (42). GOMst4 also yielded two sequences amplified from DNA collected from surface water (5 m). One sequence clustered closely among the form IB *Synechococcus* sequences (88 to 95% identical). The second GOMst4 surface water (5 m) sequence also clustered in the form IA clade and was most closely related to the *Hydrogenovibrio* and *Prochlorococcus* GP2 sequences. Station GOMst8 yielded only one sequence amplified from mRNA, which clustered with form ID sequences from the prymnesiophytes *Pleurochrysis carterae*, *Chrysochromulina hirta*, and *Emiliania huxleyi* (91 to 92% identical [Fig. 2]).

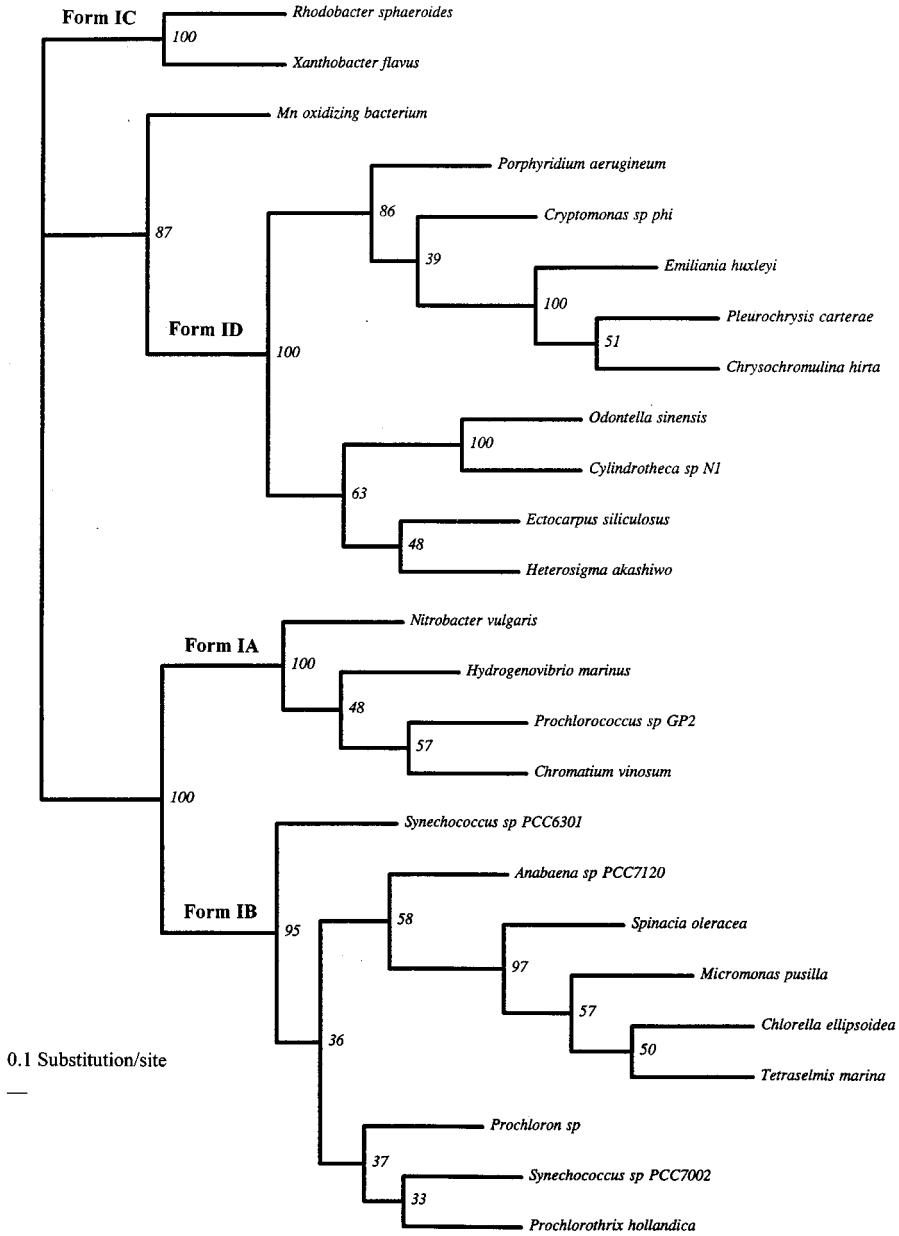

FIG. 1. Consensus neighbor-joining (distance) phylogenetic tree of full-length or nearly full-length deduced amino acid *rbcL* sequences from a variety of autotrophic taxa which utilize the Calvin cycle pathway of carbon dioxide fixation. Bootstrap values are shown inside each node (percentages of 500 bootstraps).

rbcL sequences amplified from Mediterranean and Pacific *Prochlorococcus* strains and an Atlantic Ocean coastal *Synechococcus* strain, WH8007, all clustered together with that of the freshwater cyanobacterium *Synechococcus* strain PCC6301 (nearly 97% identical). A maximum-parsimony analysis of the data in Fig. 2 was also performed (data not shown). The results of the maximum-parsimony (not shown) and neighbor-joining (Fig. 2) analyses of the partial-sequence data were similar to the results of neighbor-joining analysis (Fig. 1) of the full-sequence data for previously published *rbcL* genes.

DISCUSSION

In this study RT-PCR and PCR were used to recover a 480- to 483-bp section of the RuBisCO large subunit gene (*rbcL*) from natural phytoplankton communities and autotrophic picoplankton isolates (*Prochlorococcus* and *Synechococcus*). We recovered a total of eight clones, three of which were clones from mRNA, indicative of active expression of these genes in the marine water column at Gulf of Mexico offshore and continental shelf stations. The phylogenies built upon PCR-length sequences (483 bp) were found to be equivalent to the phylogenies built with full-length sequences. However, bootstrap confidence in the PCR-length phylogenies was low within certain clades while still supporting the basic division of the *rbcL* tree into four distinct form I types. Therefore, shorter *rbcL* sequences appear to accurately reflect the relationships among organisms that can be inferred from full-length sequence information. Taken together (mRNA and DNA), *rbcL* sequences were recovered that represent three of the four major types of form I RuBisCO. These data show that the water

TABLE 1. Results of genetic analysis of environmental and photoautotrophic picoplankton sequences

Station and sequence	Primer set	Vector	No. of clones		OTU	No. sequenced[b]	Forward/ reverse sequence[c]	Clone name	Phylogenetic affiliate
			Total	RE-screened[a]					
GOMst4									
mRNA									
50 m	Chromo	pSPORT	22	22	1	2	Y/N	GOMst4 mRNA 50m	*Prochlorococcus* GP2
80 m	Chromo	pSPORT	32	32	1	2	Y/N	GOMst4 mRNA 80m	Mn-oxidizing bacterium
DNA									
5 m	168/328	M13mp18 or -19	15	15	2	2	Y/Y	GOMst4 DNA1 5m	*Synechococcus* 6301 (OTU 1)
								GOMst4 DNA2 5m	*Prochlorococcus* GP2 (OTU 2)
GOMst8									
mRNA									
5 m	168/328	pCR/M13	5	ND[d]	ND	2	Y/Y	GOMst8 mRNA 5m	*Pleurochrysis carterae*
Prochlorococcus									
Pacific	168/328	M13mp18 or -19	44	ND	ND	2	Y/Y	NA[e]	*Synechococcus* 6301
Mediterranean	168/328	M13mp18 or -19	16	ND	ND	1	Y/Y	NA	*Synechococcus* 6301
Synechococcus									
WH8007	168/328	M13mp18 or -19	8	ND	ND	1	Y/N	NA	*Synechococcus* 6301

[a] Tetrameric restriction enzyme (RE) screening.
[b] Number of clones sequenced from each OTU.
[c] Y, yes; N, no.
[d] ND, not determined.
[e] NA, not applicable.

column autotrophic community is comprised of a diversity of organisms. The clones represent just a sample of the diversity contained in the Gulf of Mexico sites and appear to reflect the most abundant members at any one depth as determined by a combination of flow cytometry and high-pressure liquid chromatography pigment analysis (35). This is more likely to be the case under stratifying conditions in the water column, where peak abundances in various phytoplankton taxa are displayed at different depths through the euphotic zone.

At the GOMst4 station, the picophytoplankton community was stratified with depth (35). The *Synechococcus* population exhibited its maximum abundance (4.2×10^4 cells/ml) in the surface water and was approximately equal in abundance to the *Prochlorococcus* population. *Prochlorococcus* was found in maximum abundance (1.9×10^5 cells/ml) at 50 m, and the subsurface chlorophyll maximum at 80 m contained the greatest abundance (6.2×10^3 cells/ml) of picoeucaryote algae, mainly of the chromophytic type. The surface water *Synechococcus* and deep water chromophytic algal populations at this station have been recognized as actively transcribing their *rbcL* genes (35). The retrieval of *rbcL* sequences reflects this stratification, as *Synechococcus*-like sequences from DNA were recovered from surface waters while sequences recovered from 50 m were more like a divergent *Prochlorococcus* sequence. This *Prochlorococcus*-like sequence from the water column and the sequences from the *Prochlorococcus* isolates (ours and others) demonstrate the high genetic diversity within this genus and the need to carefully select gene probes and conditions for studying expression of *rbcL* in this organism in the oceans. Interestingly, the *rbcL* mRNA sequence recovered from deep water (80 m), while expected to be chromophytic in nature, was more like an *rbcL* from a recently discovered bacterial manganese oxidizer (5). Such organisms are known to be present in coastal environments but were unexpected at this open-ocean site. This suggests either that such bacterial autotrophs are active and may play a previously unrecognized role in the upper ocean carbon dynamics or that the chromophytic branch of the *rbcL* tree may be more diverse than is currently recognized. The only active, definitively chromophytic *rbcL* sequence recovered was from the continental shelf station GOMst8. These waters are recognized as often containing abundant populations of various chromophytic-type algae, including prymnesiophyte algae (48).

The marine *Synechococcus* and *Prochlorococcus* genera are recognized as exhibiting a great amount of diversity. The guanine-plus-cytosine content of from 39 to 71 mol% in the *Synechococcus* group spans almost the entire range of those of all procaryotes (49). Among the three marine clusters (A, B, and C) the range in G+C is 47 to 69 mol%. In this study we sequenced *rbcL* from two *Prochlorococcus* strains and from a member of marine cluster B, *Synechococcus* WH8007. All three sequences clustered among the form IB *rbcL*-containing cyanobacteria. When considered together with other divergent *rbcL* genes recovered from *Prochlorococcus* sp. GP2 (41) and *Synechococcus* sp. WH7803 (51), these sequences represent a broad range of diversity of *rbcL* types and show that the marine A, B, and C *Synechococcus* clusters (49) are widely dispersed across the *rbcL* phylogenetic tree. It should be recognized that phylogenies based on the *rbcL* locus may be different from those based upon other gene loci (16s rRNA [3, 22] and *tufA* [11]) because of either gene duplication and differential retention of the *rbcL* locus or past horizontal gene transfer events (12, 27). For example, the diversity between chromophyte algae (form ID), the cyanobacteria-to-higher-plant clade (form IB), the appearance of *Prochlorococcus* sp. GP2 and *Synechococcus* sp. WH7803 among form IA sequences, and the recent observations of form II-like RuBisCOs in peridinin-containing dinoflagellates (20, 39) indicate multiple origins of the *rbcL* gene and that past lateral gene transfer events have occurred among the phytoplankton divisions (or plastid precursor organisms) and autotrophic bacteria. Therefore, *rbcL* may be better suited for studying the relationship of organisms within a clade once the gene has become firmly established within that section of the phylogenetic tree. We intend to take advantage of this greater level of genetic resolution to identify members of the marine autotrophic picoplankton community by direct methods in situ.

TABLE 2. Percent identities of PCR-length, full-length, and nearly full-length form I *rbcL* sequences shown in Fig. 1 and 2[a]

Organism or sequence	% Identity with[b]:																															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1. *Rhodobacter*																																
2. *Xanthobacter*	91 (86)																															
3. Mn oxidizer	77 (74)	77 (73)																														
4. *Odontella*	75 (72)	75 (70)	75 (73)																													
5. *Cylindrotheca*	74 (71)	74 (68)	77 (73)	92 (93)																												
6. *Heterosigma*	77 (74)	74 (71)	79 (72)	91 (87)	90 (86)																											
7. *Ectocarpus*	74 (72)	71 (69)	77 (73)	87 (84)	85 (83)	90 (86)																										
8. *Cryptomonas*	74 (71)	72 (70)	79 (73)	82 (82)	82 (82)	86 (83)	86 (83)																									
9. *Porphyridium*	77 (74)	74 (70)	81 (73)	84 (84)	88 (86)	90 (85)	85 (84)	91 (86)																								
10. *Pleurochrysis*	76 (72)	73 (69)	79 (73)	83 (82)	82 (82)	87 (85)	86 (83)	87 (86)	89 (86)																							
11. *Chrysochromulina*	76 (74)	74 (71)	78 (74)	81 (81)	82 (81)	85 (82)	82 (82)	85 (84)	90 (84)	92 (93)																						
12. *Emiliania*	75 (74)	72 (71)	78 (74)	81 (81)	80 (81)	82 (83)	83 (83)	82 (85)	86 (85)	91 (93)	94 (94)																					
13. *Hydrogenovibrio*	66 (58)	69 (59)	57 (57)	59 (55)	62 (54)	59 (54)	57 (54)	60 (55)	61 (55)	59 (54)	57 (53)	57 (53)																				
14. *Chromatium*	64 (58)	66 (58)	62 (56)	60 (54)	62 (55)	59 (55)	57 (53)	60 (54)	61 (55)	58 (52)	57 (52)	57 (52)	88 (83)																			
15. *Nitrobacter*	64 (57)	66 (57)	63 (57)	60 (56)	62 (56)	60 (56)	59 (54)	59 (54)	61 (55)	58 (54)	57 (52)	57 (53)	83 (81)	84 (79)																		
16. *Prochlorococcus* GP2	66 (60)	67 (60)	63 (59)	66 (56)	62 (55)	60 (54)	59 (54)	61 (54)	62 (54)	58 (53)	59 (52)	59 (53)	85 (84)	83 (81)	79 (79)																	
17. *Spinacia*	65 (58)	68 (58)	65 (57)	58 (53)	61 (54)	59 (54)	59 (54)	60 (54)	62 (54)	61 (54)	61 (55)	62 (56)	81 (74)	79 (70)	74 (70)	75 (70)																
18. *Micromonas*	64 (58)	66 (60)	60 (57)	62 (59)	61 (57)	61 (57)	59 (55)	57 (54)	59 (54)	59 (54)	59 (54)	58 (54)	80 (75)	80 (72)	76 (74)	77 (68)	88 (86)															
19. *Tetraselmis*	61 (58)	62 (59)	55 (54)	59 (56)	60 (57)	57 (55)	56 (54)	53 (52)	57 (52)	57 (53)	56 (54)	57 (54)	77 (71)	79 (70)	74 (72)	73 (60)	87 (87)	87 (85)														
20. *Chlorella*	66 (58)	66 (58)	63 (57)	60 (54)	61 (54)	59 (54)	58 (55)	57 (54)	59 (54)	59 (54)	58 (54)	59 (54)	81 (74)	80 (71)	76 (72)	76 (71)	89 (88)	89 (89)	90 (90)													
21. *Prochloron*	64 (59)	65 (60)	65 (59)	59 (58)	62 (56)	60 (58)	60 (56)	62 (56)	63 (56)	59 (55)	58 (54)	58 (55)	81 (78)	79 (74)	77 (75)	77 (74)	81 (85)	81 (80)	78 (73)	81 (86)												
22. *Prochlorothrix*	66 (60)	67 (61)	64 (58)	63 (58)	65 (57)	62 (58)	60 (55)	61 (56)	62 (57)	60 (56)	59 (55)	59 (56)	86 (78)	86 (75)	81 (75)	80 (75)	86 (84)	86 (82)	84 (81)	88 (86)	88 (89)											
23. *Synechococcus* 7002	67 (61)	68 (59)	64 (57)	63 (58)	65 (57)	62 (56)	61 (56)	62 (56)	63 (57)	59 (57)	60 (57)	61 (57)	86 (79)	85 (74)	80 (74)	79 (75)	85 (81)	84 (81)	83 (81)	85 (83)	90 (89)	95 (90)										
24. *Anabaena* 7120	64 (59)	66 (59)	63 (57)	63 (57)	63 (57)	65 (56)	61 (56)	59 (54)	59 (54)	61 (55)	58 (54)	58 (54)	58 (55)	86 (78)	81 (74)	80 (74)	79 (75)	85 (83)	84 (82)	81 (80)	84 (83)	83 (86)	87 (85)	87 (84)								
25. *Synechococcus* 6301	66 (58)	66 (59)	64 (57)	61 (56)	63 (56)	60 (55)	58 (55)	59 (56)	61 (56)	58 (54)	59 (56)	59 (56)	89 (80)	84 (74)	79 (73)	81 (75)	85 (81)	85 (80)	83 (79)	87 (82)	86 (85)	91 (87)	92 (87)	91 (83)								
26. GOMst8 mRNA 5m[c]	74	72	76	80	81	83	84	84	86	92	91	92	59	57	57	60	62	58	56	58	61	61	61	60	59							
27. GOMst4 mRNA 80m	82	79	87	80	82	82	79	80	82	81	78	77	66	63	64	62	64	63	61	64	63	65	65	65	66	77						
28. GOMst4 DNA2 5m	66	66	63	61	63	61	59	60	62	59	60	61	87	82	75	87	79	78	75	79	76	85	85	83	86	60	64					
29. GOMst4 mRNA 50m	68	67	67	63	66	64	62	65	66	64	62	61	84	81	79	92	78	77	74	77	76	82	82	81	83	64	66	86				
30. GOMst4 DNA1 5m	61	62	60	58	60	57	56	56	58	56	56	56	84	79	74	76	80	81	80	82	83	86	88	86	95	57	62	86	79			
31. *Prochlorococcus* Pacific	65	66	64	61	63	60	58	58	61	57	58	56	87	83	77	79	83	81	80	84	86	87	91	89	97	59	64	86	81	92		
32. *Prochlorococcus* MED	64	65	63	60	62	59	57	58	61	57	58	59	86	81	77	78	82	81	80	84	82	88	89	88	97	59	64	86	80	94	94	
33. *Synechococcus* 8007	64	65	63	60	62	59	57	58	61	57	58	59	87	83	77	79	83	82	81	84	85	88	90	88	97	59	64	86	81	94	95	97

[a] PCR sequence size is a 480- to 483-bp amplified sequence which is translated into amino acids (160 to 161 residues). Nearly full-length sequences were used in some cases where the entire *rbcL* gene sequence has not yet been reported.

[b] Percent identities were determined using the PALIGN program contained in PCGENE. Full-length and nearly full-length sequence identities are shown in parentheses.

[c] Sequences 26 to 33 were determined in this study.

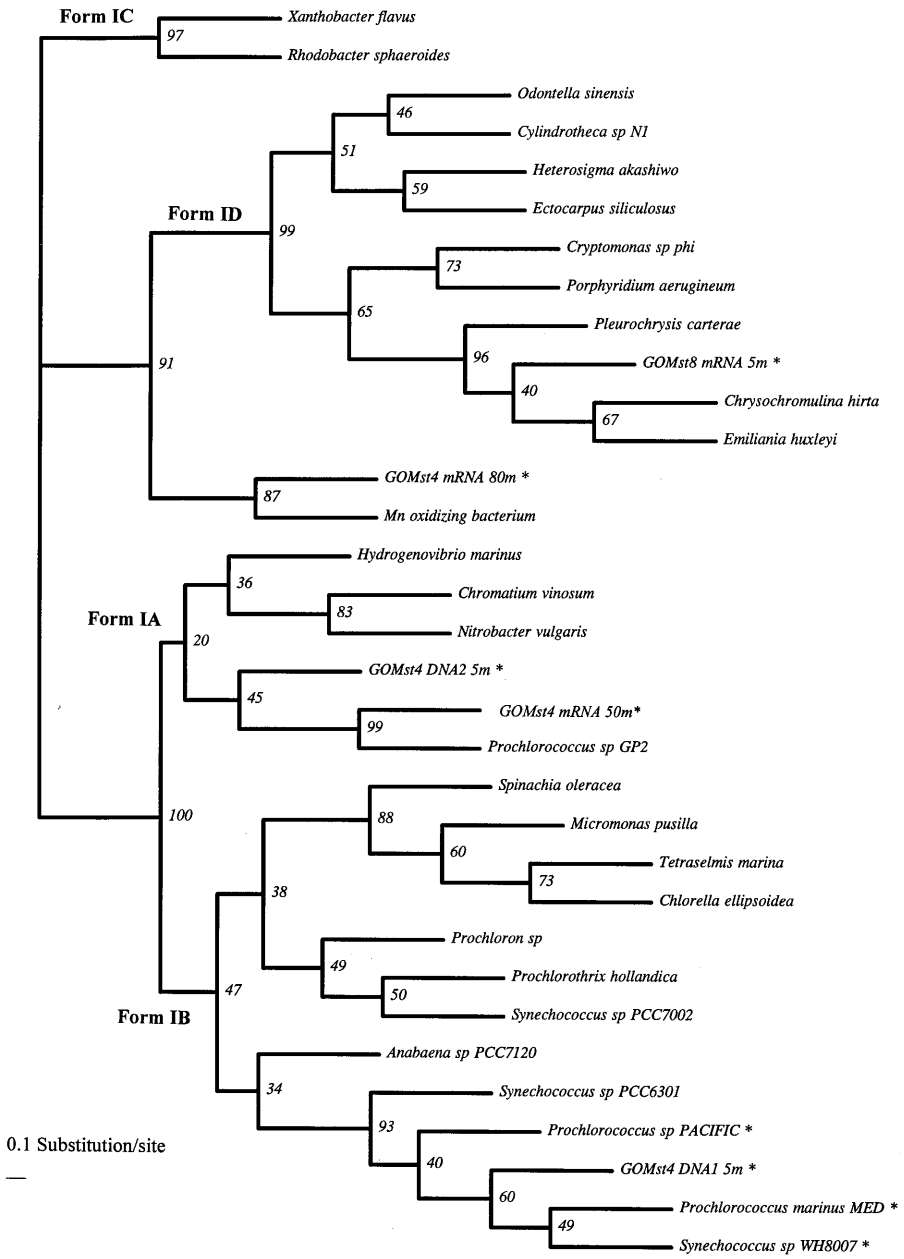

FIG. 2. Consensus neighbor-joining phylogenetic tree of amino acid sequences deduced from PCR-length (480 to 483 bp) fragments of *rbcL* genes from a variety of autotrophic taxa and transcriptionally active environmental organisms. Sequences determined in this study are indicated by an asterisk. Bootstrap values are shown inside each node (percentages of 500 bootstraps).

ACKNOWLEDGMENTS

This research was supported by National Science Foundation grants OCE9216937 and OCE9696198 to J.H.P. and L.C., respectively, and Department of Energy grant DE-FG05-93ER6199 to J.H.P. S.L.P. was supported by a U.S. Geological Survey Coastal Science Fellowship and a Knight Oceanographic Research Fellowship from the USF Department of Marine Science.

The Mediterranean strain of *Prochlorococcus marinus* was kindly provided by Brian Palenik, Scripps Institute of Oceanography, University of California-San Diego, La Jolla.

REFERENCES

1. **Bainbridge, G., P. Madgwick, S. Parmar, R. Mitchell, M. Paul, J. Pitts, A. J. Keys, and M. A. J. Parry.** 1995. Engineering rubisco to change its catalytic properties. J. Exp. Bot. **46:**1269–1276.
2. **Becker-André, M.** 1993. Absolute levels of mRNA by polymerase chain reaction-aided transcript titration assay. Methods Enzymol. **218:**420–445.
3. **Bhattacharya, D., and L. Medlin.** 1995. The phylogeny of plastids: a review based on comparisons of small-subunit ribosomal RNA coding regions. J. Phycol. **31:**489–498.
4. **Campbell, L., L. P. Shapiro, and E. Haugen.** 1994. Immunochemical characterization of eukaryotic ultraplankton from the Atlantic and Pacific Oceans. J. Plankton Res. **16:**35–51.
5. **Caspi, R., M. G. Haygood, and B. M. Tebo.** 1996. Unusual ribulose-1,5-bisphosphate carboxylase/oxygenase genes from a marine manganese-oxidizing bacterium. Microbiology **142:**2549–2559.
6. **Chernykh, N. A.** 1995. Detection of the ribulose bisphosphate carboxylase gene in natural samples by polymerase chain reaction. Mikrobiologiya **64:** 792–796.
7. **Chisholm, S. W., R. J. Olson, E. R. Zettler, R. Goericke, J. B. Waterbury, and N. A. Welschmeyer.** 1988. A novel free-living prochlorophyte abundant in the

oceanic euphotic zone. Nature (London) **334:**340–343.
8. **Chisholm, S. W., S. L. Frankel, R. Goericke, R. J. Olsen, B. Palenik, J. B. Waterbury, L. West-Johnsrud, and E. R. Zettler.** 1992. *Prochlorococcus marinus* nov. gen. sp.: an oxyphototrophic marine prokaryote containing divinyl chlorophyll *a* and *b*. Arch. Microbiol. **157:**297–300.
9. **Clegg, M. T.** 1993. Chloroplast gene sequences and the study of plant evolution. Proc. Natl. Acad. Sci. USA **90:**363–367.
10. **DeLong, E. F.** 1992. Archaea in coastal marine environments. Proc. Natl. Acad. Sci. USA **89:**5685–5689.
11. **Delwiche, C. F., M. Kuhsel, and J. D. Palmer.** 1995. Phylogenetic analysis of *tufA* sequences indicates a cyanobacterial origin of all plastids. Mol. Phylogenet. Evol. **4:**110–128.
12. **Delwiche, C. F., and J. D. Palmer.** 1996. Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids. Mol. Biol. Evol. **13:**873–882.
13. **Felsenstein, J.** 1989. PHYLIP—phylogeny inference package. Cladistics **5:** 164–166.
14. **Felsenstein, J.** 1993. PHYLIP—phylogeny inference package (version 3.5c). Department of Genetics, University of Washington, Seattle.
15. **Fuhrman, J. A., K. McCallum, and A. A. Davis.** 1993. Phylogenetic diversity of subsurface marine microbial communities from Atlantic and Pacific Oceans. Appl. Environ. Microbiol. **59:**1294–1302.
16. **Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field.** 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature (London) **345:** 60–63.
17. **Golenberg, E. M., D. E. Giannasi, M. T. Clegg, C. J. Smiley, M. Durbin, D. Henderson, and G. Zurawski.** 1990. Chloroplast DNA sequence from a Miocene magnolia species. Nature (London) **344:**656–658.
18. **Jung, V., S. B. Pestka, and S. Pestka.** 1993. Cloning of polymerase chain reaction-generated DNA containing terminal restriction endonuclease recognition sites. Methods Enzymol. **218:**357–362.
19. **Manen, J. F., V. Savolainen, S. De Marchi, and B. Rion.** 1995. Chloroplast DNA sequences from a Miocene diatomite deposit in Ardèche (France). C. R. Acad. Sci. Ser. III Sci. Vie **318:**971–975.
20. **Morse, D., P. Salois, P. Markovic, and J. W. Hastings.** 1995. A nuclear-encoded form II rubisco in dinoflagellates. Science **268:**1622–1624.
21. **Moyer, C. L., F. C. Dobbs, and D. M. Karl.** 1994. Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. Appl. Environ. Microbiol. **60:**871–879.
22. **Nelissen, B., Y. Van der Peer, A. Wilmotte, and R. De Wachter.** 1995. An early origin of plastids within the cyanobacterial divergence is suggested by evolutionary trees based on complete 16S rRNA sequences. Mol. Biol. Evol. **12:**1166–1173.
23. **Page, R. D. M.** 1996. Treeview: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. **12:**357–358.
24. **Palenik, B.** 1994. Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. Appl. Environ. Microbiol. **60:**3212–3219.
25. **Palenik, B., and R. Haselkorn.** 1992. Multiple evolutionary origins of prochlorophytes, the chlorophyll *b*-containing prokaryotes. Nature (London) **355:**265–267.
26. **Palenik, B., and H. Swift.** 1996. Cyanobacterial evolution and prochlorophyte diversity as seen in DNA-dependent RNA polymerase gene sequences. J. Phycol. **32:**638–646.
27. **Palmer, J. D.** 1995. Rubisco rules fall; gene transfer triumphs. Bioessays **17:**1005–1008.
28. **Paul, J. H.** 1996. Carbon cycling: molecular regulation of photosynthetic carbon fixation. Microb. Ecol. **32:**231–245.
29. **Paul, J. H., and S. L. Pichard.** 1995. Extraction of DNA and RNA from aquatic environments, p. 153–177. *In* J. D. Trevors and J. D. Van Elsas (ed.), Nucleic acids in the environments: methods and applications. Springer-Verlag, Berlin, Germany.
30. **Paul, J. H., and S. L. Pichard.** 1996. Molecular approaches to studying natural communities of autotrophs, p. 301–309. *In* M. E. Lindstrom and F. R. Tabita (ed.), Microbial growth on C_1 compounds. Kluwer Academic Publishers, Dordrecht, The Netherlands.
31. **Paul, J. H., L. Cesares, and J. Thurmond.** 1990. Amplification of the *rbcL* gene from dissolved and particulate DNA from aquatic environments. Appl. Environ. Microbiol. **56:**1963–1966.
32. **Pichard, S. L., and J. H. Paul.** 1991. Detection of gene expression in genetically engineered microorganisms and natural phytoplankton populations in the marine environment by mRNA analysis. Appl. Environ. Microbiol. **57:** 1721–1727.
33. **Pichard, S. L., M. E. Frischer, and J. H. Paul.** 1993. Ribulose bisphosphate carboxylase gene expression in subtropical marine phytoplankton populations. Mar. Ecol. Prog. Ser. **101:**55–65.
34. **Pichard, S. L., L. Campbell, J. B. Kang, F. R. Tabita, and J. H. Paul.** 1996. Regulation of ribulose bisphosphate carboxylase gene expression in natural phytoplankton communities. I. Diel rhythms. Mar. Ecol. Prog. Ser. **139:**257–265.
35. **Pichard, S. L., L. Campbell, K. Carder, J. B. Kang, J. Patch, F. R. Tabita, and J. H. Paul.** 1997. Analysis of ribulose bisphosphate carboxylase gene expression in natural phytoplankton communities by group-specific gene probing. Mar. Ecol. Prog. Ser.**149:**239–253.
36. **Platt, T., and W. K. W. Li (ed.).** 1986. Photosynthetic picoplankton. Can. Bull. Fish. Aquat. Sci. **214:**1–583.
37. **Preston, G. M.** 1996. Polymerase chain reaction with degenerate oligonucleotide primers to clone gene family members. Methods Mol. Biol. **58:**303–309.
38. **Read, B. A., and F. R. Tabita.** 1994. High substrate specificity factor ribulose bisphosphate carboxylase/oxygenase from eukaryotic marine algae and properties of recombinant cyanobacterial rubisco containing "algal" residue modifications. Arch. Biochem. Biophys. **312:**210–218.
39. **Rowan, R., S. M. Whitney, A. Fowler, and D. Yellowlees.** 1996. Rubisco in marine symbiotic dinoflagellates: form II enzymes in eukaryotic oxygenic phototrophs encoded by a nuclear multigene family. Plant Cell **8:**539–553.
40. **Saunders, G. W., D. Potter, M. P. Paskind, and R. A. Andersen.** 1995. Cladistic analyses of combined traditional and molecular data sets reveal an algal lineage. Proc. Natl. Acad. Sci. USA **92:**244–248.
41. **Shimada, A., S. Kanai, and T. Maruyama.** 1995. Partial sequence of ribulose-1,5-bisphosphate carboxylase/oxygenase and the phylogeny of *Prochloron* and *Prochlorococcus* (Prochlorales). J. Mol. Evol. **40:**671–677.
42. **Shimada, A., T. Maruyama, and S. Miyachi.** 1996. Vertical distributions and photosynthetic action spectra of two oceanic picophytoplankters, *Prochlorococcus marinus* and *Synechococcus* sp. Mar. Biol. **127:**15–23.
43. **Steel, M. A., P. J. Lockart, and D. Penny.** 1993. Confidence in evolutionary trees from biological sequence data. Nature (London) **364:**440–442.
44. **Tabita, F. R.** 1995. The biochemistry and metabolic regulation of carbon metabolism and CO_2 fixation in purple bacteria, p. 885–914. *In* R. E. Blankenship, M. T. Madigan, and C. E. Bauer (ed.), Anoxygenic photosynthetic bacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands.
45. **Uemura, K., H. Tokai, T. Higuchi, H. Murayama, H. Yamamoto, Y. Enomoto, S. Fujiwara, J. Hamada, and A. Yokota.** 1995. Distribution of fallover in the carboxylase reaction among ribulose 1,5-bisphosphate carboxylase/oxygenases of photosynthetic organisms. Energy Convers. Manage. **36:**771–774.
46. **Urbach, E., D. L. Robertson, and S. W. Chisholm.** 1992. Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. Nature (London) **355:**267–270.
47. **Valentin, K., and K. Zetsche.** 1989. The genes of both subunits of ribulose-1,5-bisphosphate carboxylase constitute an operon on the plastome of a red alga. Curr. Genet. **16:**203–209.
48. **Vargo, G.** Personal communication.
49. **Waterbury, J. B., and R. Rippka.** 1989. Order *Chroococcales* Wettstein 1924, emend. Rippka et al., 1979, p. 1728–1739. *In* J. T. Staley, M. P. Bryant, N. Pfennig, and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 3. Williams and Wilkins, Baltimore, Md.
50. **Waterbury, J. B., and J. M. Willey.** 1988. Isolation and growth of marine planktonic cyanobacteria. Methods Enzymol. **167:**100–105.
51. **Watson, G. M. F., and F. R. Tabita.** 1996. Regulation, unique gene organization, and unusual primary structure of carbon fixation genes from a marine phycoerythrin-containing cyanobacterium. Plant Mol. Biol. **32:**1103–1116.
52. **Watson, G. M. F., and F. R. Tabita.** 1997. Microbial ribulose-1,5-bisphosphate carboxylase/oxygenase: a molecule for phylogenetic and enzymological investigations. FEMS Microbiol. Lett. **146:**13–22.
53. **Wolfe, K. H., W. H. Li, and P. M. Sharp.** 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. USA **84:**9054–9058.
54. **Wood, A. M., and D. Townsend.** 1990. DNA polymorphism within the WH7803 serogroup of marine *Synechococcus* spp. (Cyanobacteria). J. Phycol. **26:**576–585.
55. **Xu, H. H., and F. R. Tabita.** 1996. Ribulose-1,5-bisphosphate carboxylase/oxygenase gene expression and diversity of Lake Erie planktonic microorganisms. Appl. Environ. Microbiol. **62:**1913–1921.
56. **Zehr, J. P., and L. A. McReynolds.** 1989. Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine cyanobacterium *Trichodesmium thiebautii*. Appl. Environ. Microbiol. **55:**2522–2526.