J Mol Evol (2002) 55:414–430
DOI: 10.1007/s00239-002-2338-9

JOURNAL OF MOLECULAR EVOLUTION

Differences in Pyrenoid Morphology Are Correlated with Differences in the *rbcL* Genes of Members of the *Chloromonas* Lineage (Volvocales, Chlorophyceae)

Hisayoshi Nozaki,[1] Keisuke Onishi,[1] Eiko Morita[2]

[1] Department of Biological Sciences, Graduate School of Science, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[2] National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan

Received: 1 February 2002 / Accepted: 8 April 2002

Abstract. *Chloromonas* is distinguished from *Chlamydomonas* primarily by the absence of pyrenoids, which are structures that are present in the chloroplasts of most algae and are composed primarily of the CO_2-fixing enzyme Rubisco. In this study we compared sequences of the *rbcL* (Rubisco large subunit-encoding) genes of pyrenoid-less *Chloromonas* species with those of closely related pyrenoid-containing *Chlamydomonas* species in the "*Chloromonas* lineage" and with those of 45 other green algae. We found that the proteins encoded by the *rbcL* genes had a much higher level of amino acid substitution in members of the *Chloromonas* lineage than they did in other algae. This kind of elevated substitution rate was not observed, however, in the deduced proteins encoded by two other chloroplast genes that we analyzed: *atpB* and *psaB*. The rates of synonymous and nonsynonymous nucleotide substitutions in the *rbcL* genes indicate that the rapid evolution of these genes in members of the *Chloromonas* lineage is not due to relaxed selection (as it preasumably is in parasitic land plants). A phylogenetic tree based on *rbcL* nucleotide sequences nested two *Chlamydomonas* species as a "pyrenoid-regained" clade within a monophyletic *Chloromonas* "pyrenoid-lost" clade. Character-state optimization with this tree suggested that the loss and the regain of pyrenoids were accompanied by eight synapomorphic amino acid replacements in the Rubisco large subunit, four of which are positioned in the region involved in its dimerization. However, both the *atpB* and the *psaB* sequence data gave robust support for a rather different set of phylogenetic relationships in which neither the "pyrenoid-lost" nor the "pyrenoid-regained" clade was resolved. The appearance of such clades in the *rbcL*-based tree may be an artifact of convergent evolutionary changes that have occurred in a region of the large subunit that determines whether Rubisco molecules will aggregate to form a visible pyrenoid.

Key words: ATP synthase β-subunit (*atpB*) gene — *Chloromonas* — Gene evolution — Large subunit of Rubisco (*rbcL*) gene — Morphological evolution — P700 chlorophyll a-apoprotein A2 (*psaB*) gene — Pyrenoid — Volvocales

Introduction

It is a basic tenet of modern biology that the morphological diversity of organisms must ultimately be explicable in terms of differences in DNA sequences. Recent comparative studies of homeobox-containing genes and other regulatory loci have provided many important new insights regarding phylogenetic relationships among major groups of organisms, evolu-

Correspondence to: Hisayoshi Nozaki; *email*: nozaki@biol.s.u-tokyo.ac.jp

tionary duplication and divergence of regulatory genes and their products, and the evolution of morphology in a very general sense (e.g., Bürglin 1995; Hasebe and Ito 2000). However, there are still very few cases in which direct relationships have been established between specific morphological characters and nucleotide substitutions in particular genes. The present study was initiated to determine whether such a relationship could be established with respect to a characteristic feature of algal morphology: the pyrenoid.

Pyrenoids are electron-dense bodies found in the chloroplast stroma of most eukaryotic algae and certain species of hornworts (Antherocerotopsida) but not in the chloroplasts of other bryophytes or in those of the vascular plants (e.g., Bold et al. 1987). The matrix of the pyrenoid is composed predominantly of the CO_2-fixing enzyme, ribulose bisphosphate carboxylase–oxygenase (Rubisco), which may constitute as much as 85% of the total polypeptide mass of a pyrenoid (Griffiths l970; Satoh et al. 1984; Okada 1992; Morita et al. 1996). It has been suggested that the pyrenoid may be an important component of the so-called CO_2-concentrating mechanism (CCM) that accounts for the high affinity for CO_2 that characterizes algal photosynthesis (Kuchitsu et al. 1988; Tsuzuki and Miyachi 1991; Ramazanov et al. 1994). The presence/absence, numbers, locations and appearance of pyrenoids have all been used as important morphological attributes to delineate green algal taxa at the generic and/or specific level (e.g., Ettl 1983; Nozaki et al. 1998a). Meanwhile, sequences of the *rbcL* genes that encode the Rubisco large subunit have proven to be extremely useful for evaluating green algal phylogenies (e.g., Manhart 1994; McCourt et al. 1995; Nozaki et al. 1995, 1997a, b, 1998a; Hepperle et al. 1998). However, attempts to explore the possible evolutionary relationships between these two sets of characters have not been reported before now.

While we were performing physiological studies recently with a group of green algae known as the *Chloromonas* lineage (Morita et al. 1999), we realized that they were ideally suited for a detailed study of the possible evolutionary relationship between changes in *rbcL* sequence and changes in pyrenoid morphology. The group we studied included six representatives of the genus *Chloromonas*—which is distinguished from the more widely known genus, *Chlamydomonas*, primarily by the absence of pyrenoid—plus the five species of *Chlamydomonas* that had previously been inferred to be close relatives of *Chloromonas* on the basis of their 18S *r*RNA sequence similarities (Buchheim et al. 1997). Our comparative analysis of *rbcL* sequences confirmed that these 11 algal isolates are closely related, indicated that clades exhibiting both loss and regain of pyrenoids were present within the group (Morita et al. 1999), and suggested that further studies were warranted, to determine whether these molecular and morphological features might be both causally and phylogenetically linked. To that end, we compared members of the *Chloromonas* lineage to 45 other green algae with respect to the sequences of their *rbcL* genes and two other chloroplast genes: *atpB* (which encodes the β subunit of ATP synthase) and *psaB* (which encodes the P700 chlorophyll *a*-apoprotein A2).

We found that certain amino acid replacements in the Rubisco large subunit appeared to represent synapomorphies for the "pyrenoid-lost" and the "pyrenoid-regained" clades that we had resolved with an *rbcL*-based phylogeny of the *Chloromonas* lineage (Morita et al. 1999). However, phylogenies based on *atpB* and *psaB* sequence data provided no support for either of these clades and raised the possibility that the regular association of particular amino acids with the presence or absence of pyrenoids in this lineage may reflect convergent evolution rather than common pathways of descent.

Materials and Methods

Taxa Studied. Species names, strain designations, and other information regarding the 56 operational taxonomic units (OTUs) that were examined in this study are provided in Table 1. All but two of these OTUs are members of the "CW" group of green algae, which includes the *Chloromonas* lineage (Buchheim et al. 1997) and all others that have a clockwise configuration of basal bodies (see Friedl 1997). Representatives of nine families of CW algae were studied. The two remaining OTUs are representatives of the "DO" group of green algae, in which the basal bodies are in a "directly opposed" orientation (Friedl 1997); they served as the outgroup for this study.

DNA Sequence Data. DNA extraction methods were as described previously (Nozaki et al. 2000). Partial *rbcL*, *atpB*, and *psaB* genes were amplified by the polymerase chain reaction (PCR) as described previously (Nozaki et al. 1995, 1999, 2000, 2002), except that three additional forward primers and two additional reverse primers were used in the analysis of *atpB* genes. The forward primers were atpB-F30 (5′-GGTGATAAT/CTGTGTAA/CGT/AGC-3′), atpB-F40 (5′-ATGGTCAAATGAACGAGCCG-3′), and atpB-F11(5′-GAACCA/GCCA/GGGA/GGCT/ACGTATGC G-3′). The reverse primers were atpB-R12 (5′-GAAACTTCA/TGCA/TCCG/AGCTTGAAC-3′) and atpB-R41 (5′-AA/GA/GTAA/GAAI GCT/CTGT/CTCIGG-3′). The PCR products were purified and sequenced as described previously (Nozaki et al. 1995, 2002). In some of the OTUs studied, one or more of the chloroplast genes analyzed contained group I or group II introns, but only the presumptive coding regions were used in this study. Coding regions and splice sites of the presumptive introns were identified according to Cech (1988), Lambowitz and Belfort (1993), and Nozaki et al. (1998b), using the GENETYX program (Software Development Co., Tokyo). The *rbcL* exons analyzed in this study corresponded to positions 31–1158 of the spinach *rbcL* gene, constitute ca. 80% of the total *rbcL* coding region, and encode 90% of the active-site residues of the Rubisco large subunit (RbcL) (Knight et al. 1990). The *atpB* and *psaB* exons we analyzed corresponded to positions 229–1356 and 274–1665 of the *Chlorella vulgaris atpB* and *psaB* genes, respectively (Wakasugi et al. 1997).

The only adjustment that was required for unambiguous alignment of the coding sequences of all three chloroplast genes from all 56 OTUs was to delete a six-nucleotide insertion that was present in the *psaB* genes of *Carteria crucifera* and *Carteria cerasiformis* between nucleotide 1464 and nucleotide 1465 [numbered with reference to the *psaB* gene of *Chlorella vulgaris* (Wakasugi et al. 1997)].

The chi-square test of homogeneity of base frequency across the 56 OTUs was carried out by PAUP 4.0b8 (Swofford 2001) for each of these three chloroplast genes, and no significant P values ($< 5\%$) were obtained.

Calculation of the Rates of Synonymous and Nonsynonymous Nucleotide Substitutions and of Amino Acid Interchanges. Nucleotide and amino acid substitutions within the chloroplast genes and gene products of the 54 CW OTUs were calculated with respect to the corresponding sequences of the outgroup—the two DO algae, *Scenedesmus quadricauda* and *Pediastrum duplex* (Table 1). The rates of synonymous and nonsynonymous nucleotide substitutions (K_s and K_a, respectively) with Poisson correction (Nei and Gojobori 1986) were calculated using DnaSP 3.5 (Rozas and Rozas 1999) and the numbers of amino acid substitutions were calculated with PAUP 4.0b8. The variance associated with each of the individual values made it difficult to calculate accurate standard deviations for such values within each of the algal groups of interest. Therefore, approximate standard deviations (SD*; Table 2) were calculated without taking into account the variances of the individual values. Differences in the rates of amino acid substitutions in all three chloroplast genes were examined in all possible pairwise combinations of the 54 CW OTUs. The pairwise relative rate test of HYPHY 0.71 beta (Muse and Pond 2000) was used to test for statistical significance of the differences, using the equal-input model for amino acids and no Γ variation, with the two DO algae designated as the outgroup.

To represent differences in evolutionary rates of amino acid substitutions within the whole CW group, a maximum likelihood (ML) tree based on the Dayhoff model was constructed for each of the three chloroplast genes by ProML of PHYLIP 3.6a2 (Felsenstein 2001), based on the deduced amino acid sequences from the 54 CW OTUs and the two DO algae (Table 1).

Phylogenetic Analyses of the Chloromonas *Lineage*. Sequence data derived in each of the three chloroplast gene analyses were used to deduce the apparent phylogenetic relationships among the 11 OTUs constituting the *Chloromonas* lineage (Table 1), using two more distantly related *Chlamydomonas* (*Cd.*) species (namely, *Cd. moewusii* and *Cd. kuwadae*) plus *Carteria obtusa* as the outgroup, and PAUP 4.0b8 to generate several types of trees from each set of sequences. First, the unambiguously aligned 1128-bp *rbcL* or *atpB* sequences, or the 1392-bp *psaB* sequences, were subjected to likelihood ML analysis based on the F81 model with empirical base frequencies: ML trees were constructed based on the heuristic search using stepwise addition of 10 random replications [with the tree bisection–reconnection (TBR) branch-swapping algorithm], and the robustness of the resulting lineages was tested by estimation of quartet puzzling support (QPS) values (which have the same practical meaning as the bootstrap values) with 1000 puzzling steps (comparable to the number of the bootstrap replicates) (Strimmer and von Haeseler 1996). Next, distance matrices for these same sets of aligned sequences were calculated using the Jukes–Cantor method (Jukes and Cantor 1969; Tajima and Takezaki 1994), each of these distance matrices was used to construct a neighbor-joining (NJ) phylogenetic tree with the algorithm of Saitou and Nei (1987), and the robustness of the resulting lineages was tested by a bootstrap analysis (Felsenstein 1985) with 1000 replications. Then an equal maximal parsimony (MP) analysis was carried out, including a bootstrap analysis based on 1000 replications of the branch-and-bound search. The deduced amino acid sequences encoded by the *rbcL*, atpB, and *psaB* genes were then used for constructing additional gene trees, as follows: MP trees based on the branch-and-bound search were calculated by PAUP 4.0b8, NJ trees based on Kimura (1983) distances were calculated using CLUSTAL X (Thompson et al. 1997), and ML analyses based on the Dayhoff model were carried out using PUZZL 4.0 (Strimmer and von Haeseler 1996). In each case bootstrap/QPS values were determined, based on 1000 replications.

NJ tree based on K_s and K_a [with Poisson correction (Nei and Gojobori 1986)] in the *rbcL* genes were constructed by MEGA 2.0 (Kumar et al. 2001) with a bootstrap analysis based on 1000 replications. The ratios of synonymous-to-nonsynonymous nucleotide substitutions (K_s/K_a ratios) in the *rbcL* genes at branches were calculated based on the topology and branch lengths of the K_s- and K_a-based NJ trees constructed.

Results

Amino Acid Substitutions and Synonymous and Nonsynonymous Nucleotide Substitutions. In Table 2 we compare the average abundance of nucleotide and deduced-amino acid substitutions that we observed in three chloroplast genes within three groups of CW algae, namely, the 11 members of the *Chloromonas* lineage, the 40 other photoautotrophic members of the CW group, and the photoheterotrophic colonial volvocalean genus *Astrephomene* (all with reference to the two species in the DO outgroup). No significant differences among the three lineages were observed in the case of the *psaB* genes, where the average amino acid substitution levels varied only between 36.7 and 38.1 per 464 residues. In the case of the *atpB* genes the results were rather similar: in the *Chloromonas* lineage and *Astrephomene* the differences in amino acid substitution levels [29.4 ± 7.2 (SD*) versus 30.3 ± 5.0 per 376 residues] clearly were not significantly different, and the level of substitution in the remaining members of the CW group (24.7 ± 4.7) was only slightly lower. In marked contrast, the level of amino acid substitutions encoded by the *rbcL* genes was substantially higher in the members of the *Chloromonas* lineage than it was in either of the other two lineages: 33.2 ± 5.3 per 376 residues in the *Chloromonas* group *versus* 12.5 ± 1.1 in *Astrephomene* and an average of 17.1 ± 4.3 in the 40 other OTUs in the CW group.

The phylogram of ML tree based on the *rbcL* amino acid sequences showed a significant difference in evolutionary rates between the *Chloromonas* lineage and other members of the CW group (Fig. 1). The evolutionary rates in the *rbcL* amino acid substitutions seemed to be accelerated, especially in the common ancestral branch of the *Chloromonas* lineage as well as in several branches within the lineage (Fig. 1). However, such a significant difference in amino acid substitutions was not recognized in either the *atpB* or the *psaB* amino acid-based tree (Figs. 2 and 3).

Table 1. List of the chlorophycean taxa/strains included in the phylogenetic analyses and DDBJ/EMBL/GenBank accession numbers of the three chloroplast genes

Taxon	Strain designation	Accession No.		
		atpB	*rbcL*	*psaB*
DO group				
Scenedesmus quadricauda	NIES[a]-96	AB084305[b]	AB084332[b]	AB084339[b]
Pediastrum duplex	NIES-213	AB084306[b]	AB084333[b]	AB084340[b]
CW group				
Chlamydomonadaceae				
Chloromonas lineage				
Chlamydomonas augustae	UTEX[c] 1969	AB084307[b]	AB022227-8[d]	AB084341[b]
Chlamydomonas bipapillata	SAG[e] 11-47	AB084308[b]	AB022225[d]	AB084342[b]
Chlamydomonas macrostellata	SAG 72.81	AB084309[b]	AB022229-30[d]	AB084343[b]
Chlamydomonas mutabilis	UTEX 578	AB084310[b]	AB022224[d]	AB084344[b]
Chlamydomonas radiata	UTEX 966	AB084311[b]	AJ001878[f]	AB084345[b]
Chloromonas clathrata	UTEX 1970	AB084312[b]	AB022533-4[d]	AB084346-7[b]
Chloromonas insignis	NIES-447	AB084313[b]	AB022226[d]	AB084348[b]
Chloromonas palmelloides	SAG 32.86	AB084314[b]	AB022530[d]	AB084349[b]
Chloromonas rosae	UTEX 1337	AB084315[b]	AB022535-6[d]	AB084350-1[b]
Chloromonas rosae	SAG 26.90	AB084316[b]	AB022531-2[d]	AB084352-3[b]
Chloromonas serbinowii	UTEX 492	AB084317[b]	AJ001879[f]	AB084354[b]
Excluding *Chloromonas* lineage				
Chlamydomonas reinhardtii	137C	M13704[g]	J01399[h]	AB044470[i]
Chlamydomonas debaryana	UTEX 1344	AB014034[j]	D86838[k]	AB044469[i]
Chlamydomonas moewusii	UTEX 9	AB014035[j]	M15842[l]	AB084355[b]
Chlamydomonas kuwadae	92-514-H-13	AB084318[b]	AB084334[b]	AB084356[b]
Chlamydomonas tetragama	NIES-446	AB084319[b]	AJ001880[f]	AB084357[b]
Carteria crucifera	NIES-421	AB084320[b]	D63431[m]	AB084358[b]
Carteria cerasiformis	NIES-425	AB084321[b]	D89768[n]	AB084359[b]
Carteria radiosa	NIES-432	AB084322[b]	D89770[n]	AB084360[b]
Carteria obtusa	NIES-428	AB084323[b]	D89769[n]	AB084361-3[b]
Pseudocarteria mucosa	NIES-522	AB084324[b]	AB084335[b]	AB084364[b]
Lobomonas monstruosa	NIES-474	AB044533[i]	AB044171[i]	AB044472[i]
Vitreochlamys ordinata	Nozaki S-4	AB014036[j]	AB014041[j]	AB044529[i]
Vitreochlamys pinguis	NIES-1148	AB076120[o]	AB050490-1[p]	AB076157[o]
Vitreochlamys aulata	NIES-1140	AB076121[o]	AB050486-7[p]	AB076158[o]
Haematococcaceae				
Haematococcus lacustris	NIES-144	AB084325[b]	AB084336-7[b]	AB084365[b]
Chlorogonium neglectum	NIES-439	AB084326[b]	AB010243[q]	AB084366[b]
Chlorogonium kasakii	CCAP[r] 12/8	AB014037[j]	AB010244[q]	AB084367[b]
Chlorogonium elongatum	IAM[S] C 293	AB084327[b]	AJ001881[f]	AB084368[b]
Chlorogonium euchlorum	CCAP 12/3	AB084328[b]	AJ001882[f]	AB084369[b]
Chlorogonium fusiforme	NIES-123	AB084329[b]	AB010242[q]	AB084370[b]
Phacotaceae				
Pteromonas angulosa	KR/91/2	AB014038[j]	AJ001887[f]	AB084371-2[b]
Phacotus lenticularis	KR/91/1	AB014039[j]	AJ001883[f]	AB084373-4[b]
Tetrabaenaceae				
Tetrabaena socialis	NIES-571	AB014014[j]	D63443[m]	AB044466[i]
Basichlamys sacculifera	N1ES-566	AB014015[j]	D63430[m]	AB044467-8[i]
Goniaceae				
Gonium pectorale	NIES-569	AB014016-7[j]	D63437[m]	AB044463[i]
Gonium octonarium	GO-LC-1 +	AB014018[j]	D63436[m]	AB044462[i]
Gonium quadratum	NIES-653	AB014019[j]	D63438[m]	AB044464[i]
Gonium multicoccum	UTEX 2580	AB014020[j]	D63435[m]	AB044461[i]
Gonium viridistellatttm	UTEX 2519	AB014021[j]	D86831[k]	AB044465[i]
Astrephomene gubernaculifera	NIES-418	AB014022-3[j]	D63428[m]	AB044458[i]
Astrephomene gubernaculifera	UTEX 1394	AB044181[i]	AB044169-70[i]	AB044459[i]
Astrephomene perforata	N1ES-564	AB014024[j]	D63429[m]	AB044460[i]
Volvocaceae				
Pandorina monrum	NIES-574	AB014025-6[j]	D63442[m]	AB044452[i]
Volvulina compacta	NIES-582	AB014029[j]	D86832[k]	AB044446[i]
Yamagishiella unicocca	UTEX 2428	AB014030[j]	D86823[k]	AB044443[i]
Platydorina caudata	UTEX 1658	AB014032[j]	D86828[k]	AB044442[i]
Eudorina elegans	NIES-456	AB014009[j]	D63432[m]	AB044435[i]
Pleodorina californica	UTEX 809	AB014004[j]	D63439[m]	AB044430[i]
Volvox carteri	NIES-732	AB013999[j]	D63446[m]	AB044425[i]

Table 1. Continued

Taxon	Strain designation	Accession No.		
		atpB	*rbcL*	*psaB*
Volvox rousseletii	UTEX 1862	AB014003[j]	D63448[m]	AB044429[i]
Dunaliellaceae				
Dunaliella parva	UTEX 1983	AB084330[b]	AJ001877[f]	AB084375[b]
Tetrasporaceae				
Paulschulzia pseudovolvox	UTEJX 167	AB014040[j]	D86837[k]	AB044473[i]
Characiochloridaceae				
Characiochloris sasae	NIES-567	AB084331[b]	AB084338[b]	AB084376[b]

[a] Microbial Culture Collection at the National Institute for Environmental Studies (Watanabe et al. 2000).
[b] Data from the present study.
[c] Culture Collection of Algae at the University of Texas at Austin (Starr and Zeikus 1993).
[d] Data from Morita et al. (1999).
[e] Sammlung von Algenkulturen at the University of Göttingen (Schlösser 1994).
[f] Data from Hepperle et al. (1998).
[g] Data from Woessner et al. (1986).
[h] Data from Dron et al. (1982).
[i] Data from Nozaki et al. (2000).
[j] Data from Nozaki et al. (1999).
[k] Data from Nozaki et al. (1997a).
[l] Data from Yang et al. (1986).
[m] Data from Nozaki et al. (1995).
[n] Data from Nozaki et al. (1997b).
[o] Data from Nozaki et al. (2002).
[p] Data from Nakazawa et al. (2001).
[q] Data from Nozaki et al. (1998a).
[r] Culture Collection of Algae and Protozoa (Thompson et al. 1988).
[s] IAM Culture Collection at the University of Tokyo (Sugiyama et al. 1988).

Table 2. Rates of synonymous (K_s) and nonsynonymous (K_a) substitutions and amino acid substitutions (P_a) of three chloroplast genes in various members of the CW group (Table 1)[a]

Gene/alga	K_s[b]	K_a[b]	P_a[c]
RbcL (376 codons)			
Chloromonas lineage (11 OTUs)	0.724 ± 0.044*,[d]	0.070 ± 0.007*	33.2 ± 5.3*
Astrephomene (3 OTUs)	0.835 ± 0.035*	0.036 ± 0.004*	12.5 ± 1.1*
Others (40 OTUs)	0.653 ± 0.101*	0.041 ± 0.007*	17.1 ± 4.3*
AtpB (376 codons)			
Chloromonas lineage	0.977 ± 0.276*	0.077 ± 0.016*	29.4 ± 7.2*
Astrephomene	0.838 ± 0.116*	0.070 ± 0.012*	30.3 ± 5.0*
Others	0.758 ± 0.098*	0.064 ± 0.013*	24.7 ± 4.7*
PsaB (464 codons)			
Chloromonas lineage	0.932 ± 0.138*	0.064 ± 0.004*	38.1 ± 2.6*
Astrephomene	0.842 ± 0.070*	0.064 ± 0.004*	36.7 ± 3.7*
Others	0.915 ± 0.144*	0.065 ± 0.007*	37.9 ± 5.8*

[a] Calculated with two species of the DO group (Table 1).
[b] Calculated by DnaSP (Rozas and Rozas 1999).
[c] Difference in number of amino acids calculated by PAUP 4.0b8 (Swofford 2001).
[d] Approximate standard deviation (SD*) calculated without taking into account variances of individual values.

The significance of this difference in amino acid substitution level was reinforced when pairwise relative rate tests were performed for the *rbcL* genes: 73% of the pairs in which a member of the *Chloromonas* lineage was compared to some other CW alga exhibited differences in the amino acid substitutions that were significant at the 1% level (Table 3), whereas only 12–14% of the pairwise within-group comparisons revealed amino acid substitution differences that were significant at this level. Furthermore, the number of significant differences observed in pairwise comparisons of the *atpB* genes between members of the *Chloromonas* lineage and the non-*Chloromonas* group (14%) was not substantially higher than the number seen (10%) when parallel comparisons were made for pairs within the non-*Chloromonas* group (Table 3), and very few significant differences were observed when either within-

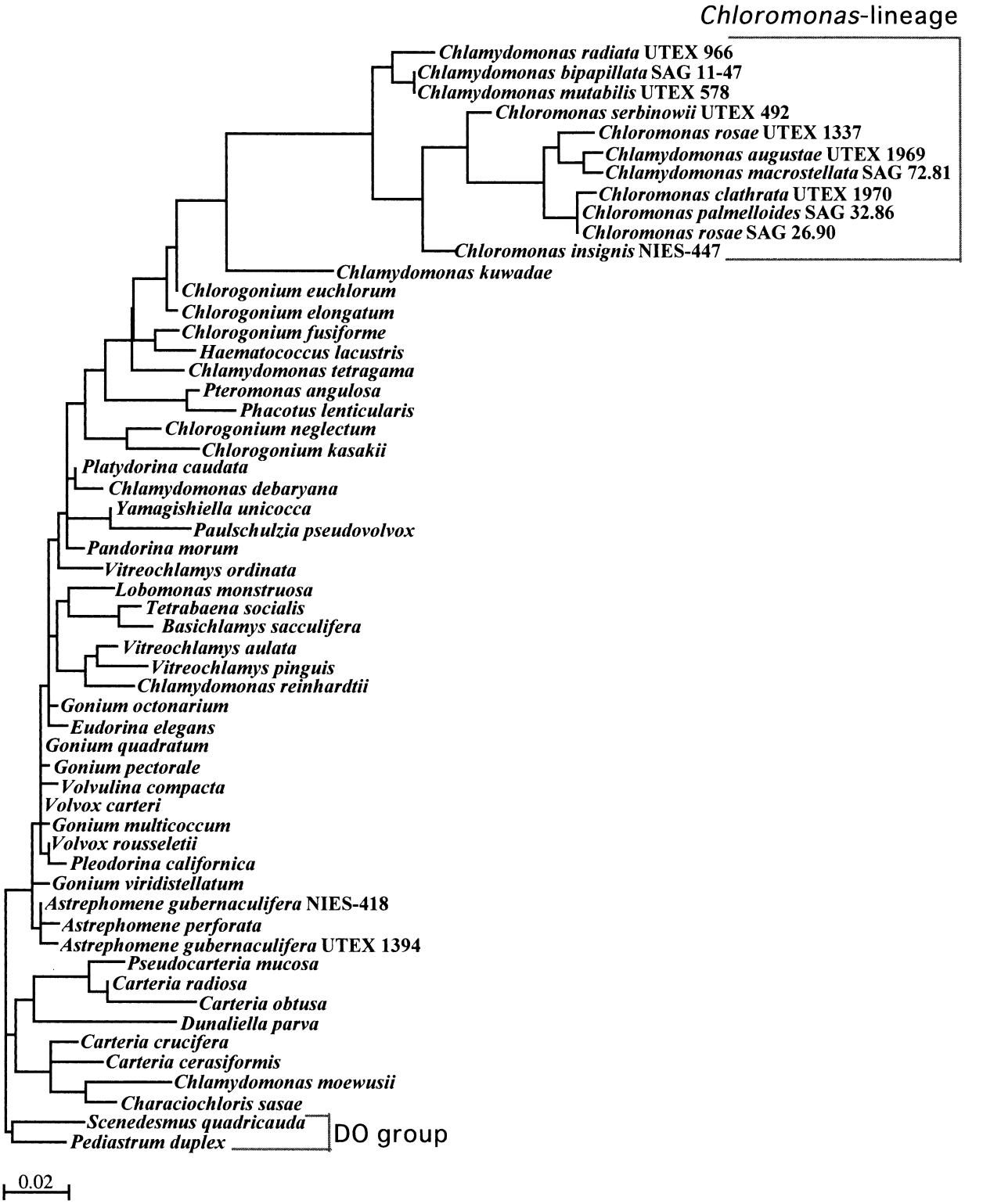


Fig. 1. Maximum likelihood (ML) tree based on the deduced amino acid sequences of *rbcL* genes from 54 strains of the CW group and two DO algae (Table 1). The tree was constructed by the ML method using the Dayhoff model by ProML of PHYLIP 3.6a2 (Felsenstein 2001). Branch lengths are proportional to the Dayhoff distances, which are indicated by the scale bar below the tree.

group or between-group comparisons were made for the *psaB* genes (Table 3).

When the rates of synonymous and nonsynonymous nucleotide substitutions in the *rbcL* genes were examined (Table 2), it became clear that the elevated level of amino acid substitutions observed in the *Chloromonas* lineage could be attributed to the substantially higher average rate of nonsynonymous nucleotide substitutions in the *Chloromonas* lineage (K_a = 0.070) than in the non-*Chloromonas* algae

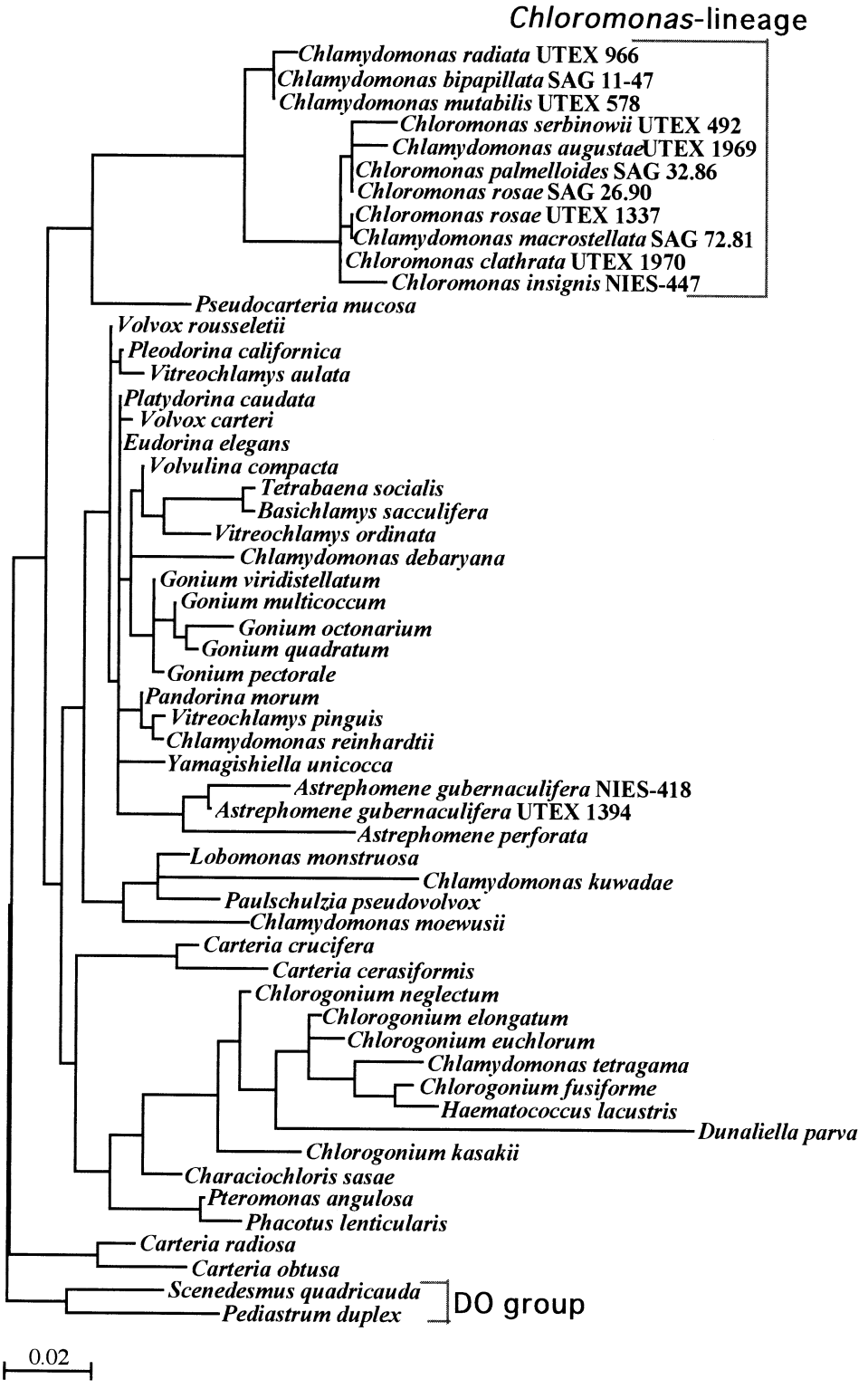


Fig. 2. Maximum likelihood (ML) tree based on the deduced amino acid of sequences of *atpB* genes from 54 strains of the CW group and two DO algae (Table 1). The tree was constructed by the ML method using the Dayhoff model by ProML of PHYLIP 3.6a2 (Felsenstein 2001). Branch lengths are proportional to the Dayhoff distances, which are indicated by the scale bar below the tree.

($K_a < 0.04$), with no substantial difference in the rates of synonymous substitutions (K_s) between the groups. However, *Astrephomene* exhibited higher K_s rates (0.835) than those of the other members (0.653, 0.724) of the CW group (Table 2). Contrary distinction to the *rbcL* genes, there was no substantial dif-

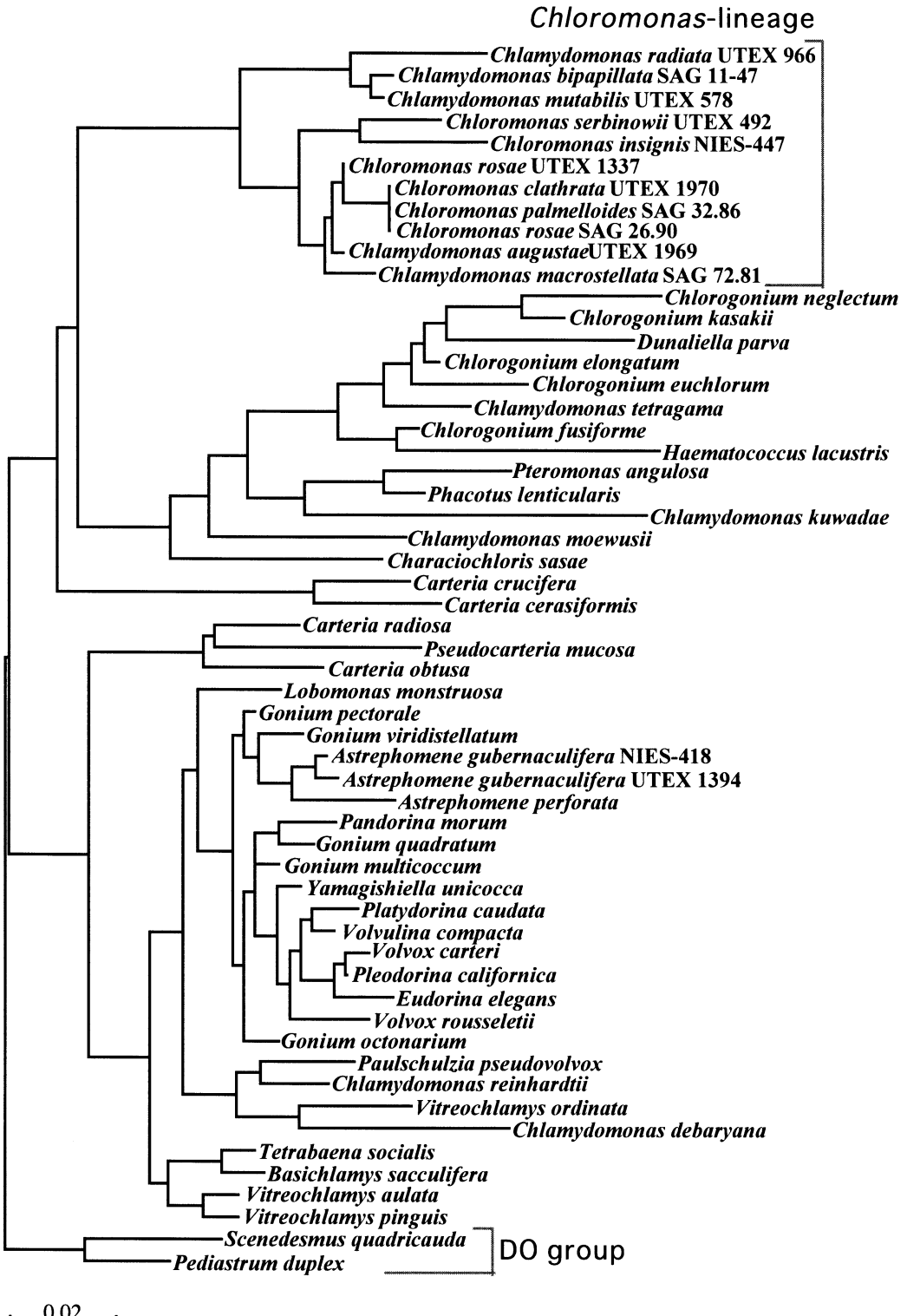


Fig. 3. Maximum likelihood (ML) tree based on the deduced amino acid sequences of *psaB* genes from 54 strains of the CW group and two DO algae (Table 1). The tree was constructed by the ML method using the Dayhoff model by ProML of PHYLIP 3.6a2 (Felsenstein 2001). Branch lengths are proportional to the Dayhoff distances, which are indicated by the scale bar below the tree.

ference in K_a between the *Chloromonas* lineage and the other algae with respect to either the *atpB* or the *psaB* genes (Table 2).

Phylogenetic Relationships Within the Chloromonas *Lineage Based on Nucleotide Sequences.* The results of our present phylogenetic analysis of the

Table 3. Numbers of pairs with significant differences in rates of amino acid substitutions of three chloroplast genes based on pairwise relative rate tests between various members of the CW group (Table 1)[a]

Gene/*p*-test	Between *Chloromonas* lineage and others (946)[b]	Within *Chloromonas* lineage (110)[b]	Within others (1806)[b]
RbcL			
$P < 5\%$	788 (83%)	32 (29%)	430 (24%)
$P < 1\%$	649 (73%)	15 (14%)	211 (12%)
AtpB			
$P < 5\%$	284 (30%)	0 (0%)	348 (19%)
$P < 1\%$	133 (14%)	0 (0%)	174 (10%)
PsaB			
$P < 5\%$	36 (4%)	3 (3%)	361 (20%)
$P < 1\%$	2 (<1%)	0 (0%)	117 (6%)

[a] Calculated by HYPHY 0.71 beta (Muse and Pond 2001) with two species of the DO group (Table 1) designated as the outgroup, based on *p* distances without the Γ distribution model.
[b] Total numbers of pairs of OTUs.

Chloromonas lineage based on nucleotide sequences of the *rbcL* genes (Fig. 4) are essentially the same as those reported previously (Morita et al. 1999). In both of these *rbcL*-based studies, the *Chloromonas* lineage was divided into two monophyletic groups, one (Group A) composed of three *Chlamydomonas* species (namely, *Cd. bipapillata*, *Cd. mutabillis*, and *Cd. radiata*) and the other (Group B) containing the remaining eight OTUs in the *Chloromonas* lineage (Fig. 4). Within Group B, the one species of *Chloromonas* (*Cr.*) that has an atypical pyrenoid [*Cr. insignis* (see Morita et al. 1999)] was positioned basally relative to a monophyletic (''pyrenoid-lost'') clade (*rbcL* clade C) that contains five pyrenoid-less strains (*Cr. serbinowii*, *Cr. rosae* UTEX 1337, *Cr. clathrata*, *Cr. palmelloides*, and *Cr. rosae* SAG 26.90) and within which is embedded a derived, monophyletic ''pyrenoid-regained'' clade (*rbcL* clade D) of two *Chlamydomonas* species (*Cd. augustae* and *Cd. macrostellata*) (Fig. 4).

However, although the phylogenetic relationships within the *Chloromonas* lineage that were inferred from the nucleotide sequences of the *atpB* and *psaB* genes were essentially identical (Figs. 5 and 6), they both differed in a very important way from the relationships that were inferred from analysis of the *rbcL* nucleotide sequences. All three gene trees robustly resolved the same two monophyletic groups, A and B, but phylogenetic analyses of the *atpB* and *psaB* sequences led to very different conclusions regarding the relationships within Group B than our analysis of the *rbcL* sequences had—particularly with respect to the positions of *Cr. insignis* and *Cd. augustae*: instead of providing any support for the validity of either a pyrenoid-lost *rbcL* clade C or a derived pyrenoid-regained *rbcL* clade D (Fig. 4), the *atpB* and *psaB* sequences indicated that Group B should be divided into two sister monophyletic groups, one composed of *Cr. serbinowii* and *Cr. insignis* and the other containing the six remaining OTUs of the *Chloromonas* group (Figs. 5 and 6). Within the latter group three pyrenoid-less strains of *Chloromonas* (*Cr. clathrata*, *Cr. palmelloides*, and *Cr. rosae* SAG 26.90) clearly constitute a robust monophyletic group (Figs. 5 and 6). However, instead of having a distal position within Group B as a derived pyrenoid-regained clade in the *rbcL* gene tree (Fig. 4), *Cd. augustae* is positioned basally in the *atpB*- and *psaB*-based trees, with the relative phylogenetic positions of *Cd. macrostellata* and *Cr. rosae* UTEX 1337 being very poorly defined in both the *atpB*- and the *psaB*-trees (Figs. 5 and 6).

When the NJ trees based on K_a and K_s in the *rbcL* genes were constructed, they differed in tree topology with regard to the *rbcL* clades C and D (Fig. 7). In the K_a-based tree, both of the clades were resolved robustly (Fig. 7A). In contrast, they were not resolved in the K_s-based analysis (Fig. 7B). K_s/K_a ratios at branches within the *Chloromonas* lineages were >3 except for the distal branch bearing *Cr. clathrata*, which showed K_a higher than K_s (Fig. 7). On the other hand the K_s/K_a ratio at the basal ancestral branch of the *Chloromonas* lineage was relatively low (1.86) (Fig. 7).

Phylogenetic Relationships Within the Chloromonas *Lineage Based on Amino Acid Sequences.* Phylogenetic relationships within the *Chloromonas* lineage were also analyzed using the deduced amino acid sequences encoded by the *rbcL*, *atpB*, and *psaB* genes from the same 14 OTUs as in the nuceotide-based analyses. Two major sublineages (Groups A and B) that were resolved in the nucleotide-based trees were identically resolved in each of the amino acid-based trees. Once again, however, the phylogenetic relationships within Group B that were inferred from the *rbcL* data were different from those inferred from the *atpB* and *psaB* data.

The phylogenetic analysis based on RbcL amino acid sequences, like the analysis based on *rbcL* nucleotide sequences, indicated that the five pyrenoid-

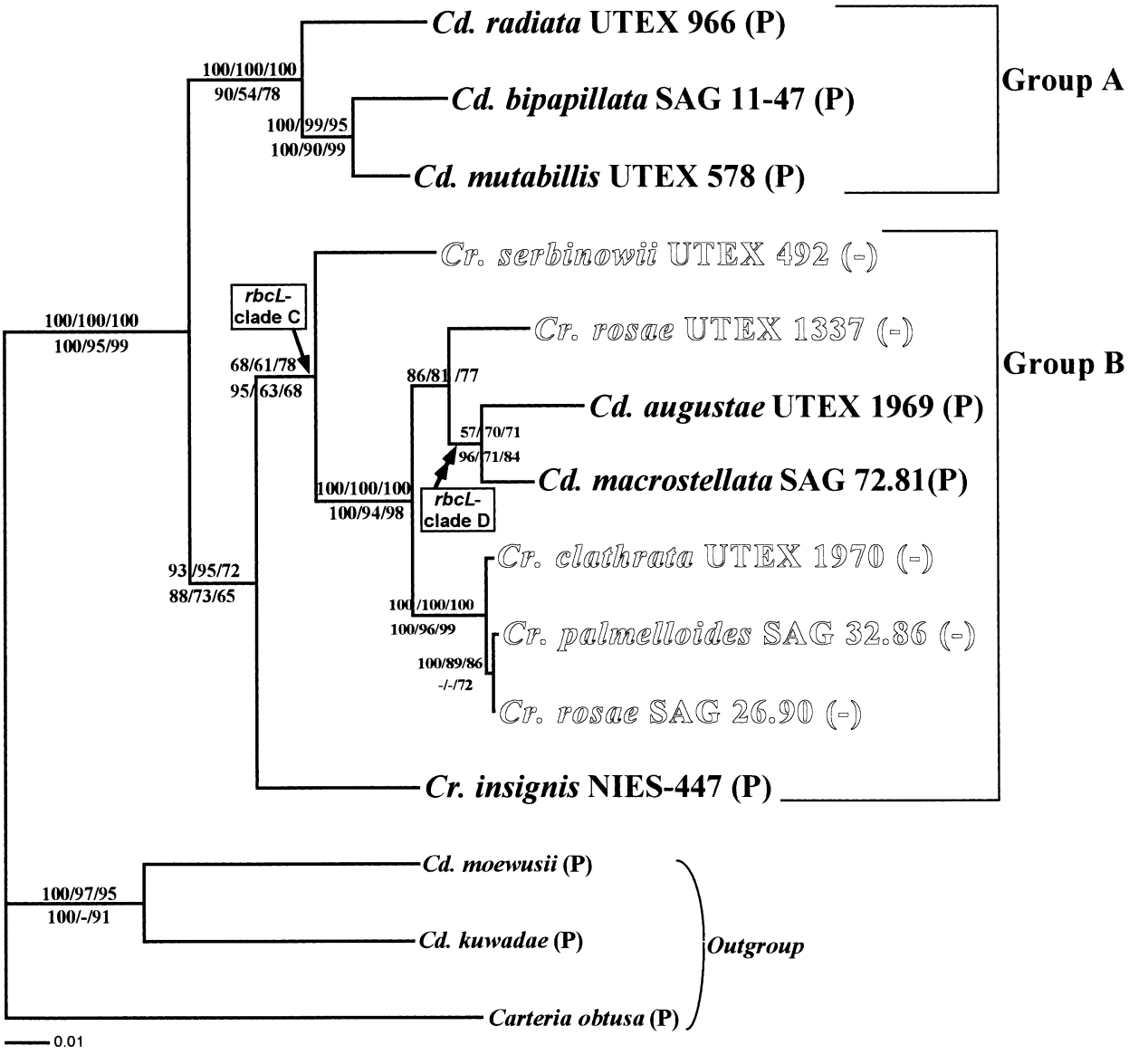


Fig. 4. Maximum likelihood (ML) tree of 11 strains of the *Chloromonas* lineage (*Cd.*, *Chlamydomonas*; *Cr.*, *Chloromonas*) and two species of *Chlamydomonas* plus *Carteria obtusa* designated as the outgroup (Table 1), based on the aligned nucleotide sequences for 1128 base pairs in coding regions of *rbcL* genes. The single tree was obtained using the F81 model, based on the heuristic search using stepwise addition of 10 random replications (with the TBR branch-swapping algorithm) by PAUP 4.0b8 (Swofford 2001). Branch lengths are proportional to the F81 distances, which are indicated by the scale bar below the tree. Numbers at the left, in the middle, and at the right above branches represent 50% or higher bootstrap/QPS values (based on 1000 replications) of the ML, NJ (based on the Jukes–Cantor distances), and MP analyses, respectively, of the *rbcL* nucleotide sequences. Branches also resolved in the consensus bootstrap tree of NJ (based on Kimura distance), MP, and ML (based on the Dayhoff model) analyses of the deduced amino acid sequences (363 amino acids) of the *rbcL* genes from the same 14 strains are presented by bootstrap/QPS values (50% or higher) based on 1000 replications shown below the branches (ML/NJ/MP). The *arrow* and *double arrow* at the branches indicate loss and regain of pyrenoids, respectively, deduced from the character optimization by PAUP 4.0b8 (Swofford 2001). The P or – in parentheses after the species/strain designation represents the presence or absence of pyrenoids, respectively (based on Morita et al. 1998, 1999).

less strains of *Chloromonas* plus *Cd. augustae* and *Cd. macrostellata* constituted a robust monophyletic group (the pyrenoid-lost *rbcL* clade C), within which the two *Chlamydomonas* species constituted a derived pyrenoid-regained clade (*rbcL* clade D) (Fig. 4). Bootstrap and QPS value supporting these two clades in ML and NJ methods were even higher for the amino acid-based analysis than for the nucleotide-based analysis (Fig. 4). Character-state optimization based on accelerated transformation for each of the 12 most parsimonious trees derived from the RbcL amino acid sequences by the branch-and-bound search of PAUP 4.0b8 indicated that *rbcL* clade C was supported by five synapomorphic amino acid residues [at positions 145, 210, 225, 276, and 309, numbered according to the deduced *Chlorella vulgaris*

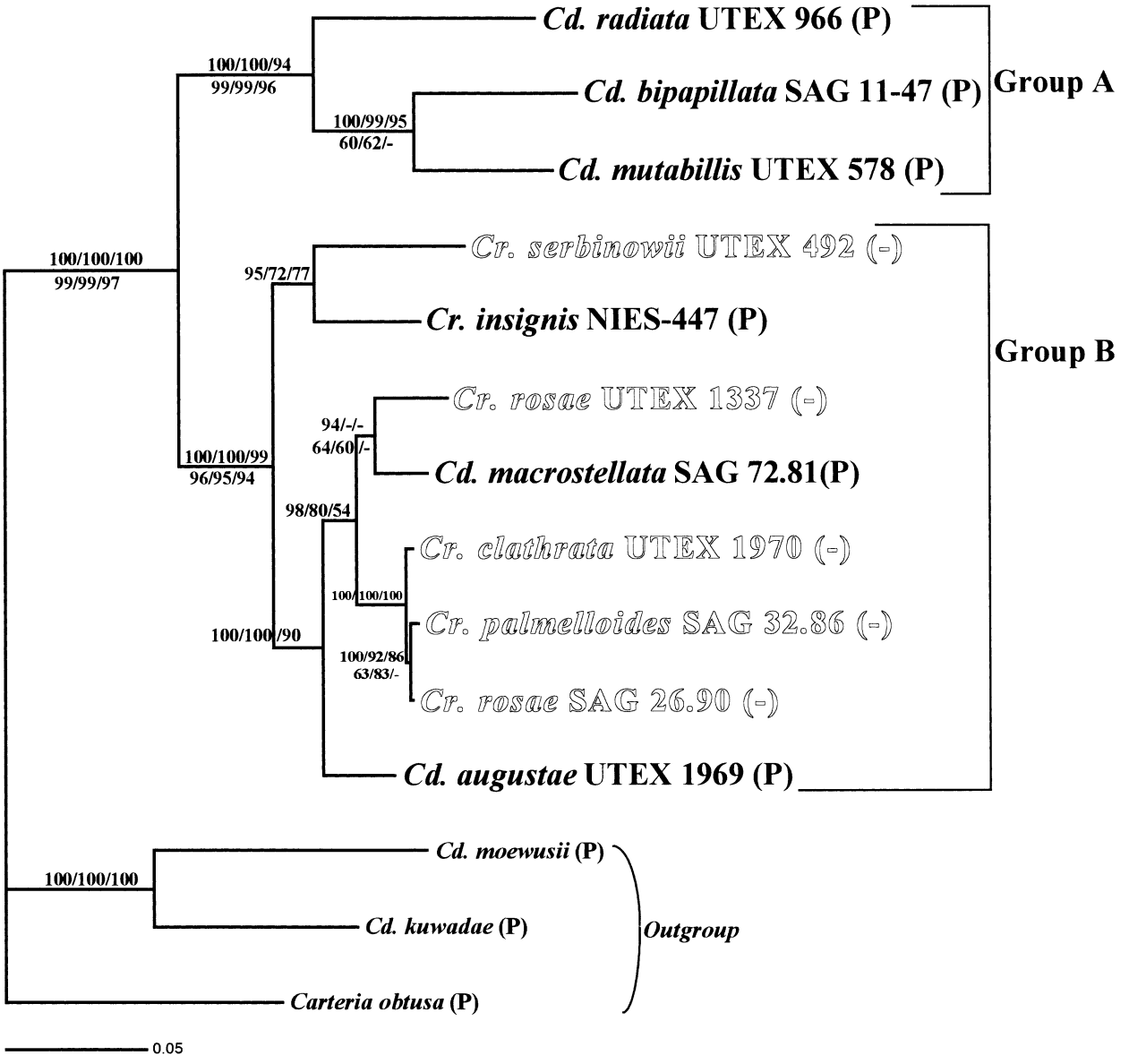


Fig. 5. Maximum likelihood (ML) tree of 11 strains of the *Chloromonas* lineage (*Cd.*, *Chlamydomonas; Cr.*, *Chloromonas*) and two species of *Chlamydomonas* plus *Carteria obtusa* designated as the outgroup (Table 1), based on the aligned nucleotide sequences for 1128 base pairs in coding regions of *atpB* genes. The single tree was obtained using the F81 model, based on the heuristic search using stepwise addition of 10 random replications (with the TBR branch-swapping algorithm) by PAUP 4.0b8 (Swofford 2001). Branch lengths are proportional to the F81 distances, which are indicated by the scale bar below the tree. Numbers at the left, in the middle, and at the right above branches represent 50% or higher bootstrap/QPS values (based on 1000 replications) of the ML, NJ (based on the Jukes–Cantor distances), and MP analyses, respectively, of the *rbcL* nucleotide sequences. Branches also resolved in the consensus bootstrap tree of NJ (based on Kimura distance), MP, and ML (based on the Dayhoff model) analyses of the deduced amino acid sequences (363 amino acids) of the *rbcL* genes from the same 14 strains are presented by bootstrap/QPS values (50% or higher) based on 1000 replications shown below the branches (ML/NJ/MP). The P or – in parentheses after the species/strain designation represents the presence or absence of pyrenoids, respectively (based on Morita et al. 1998, 1999).

sequence (Wakasugi et al. 1997)] and that *rbcL* clade D was supported by either two or three synapomorphic amino acid residues [at positions 32, (222), and 247]. Five of these specific residues in the *Chloromonas* lineage (those at positions 145, 210, 222, 225, and 309) represent unique characters for all 56 OTUs analyzed (Table 1). In addition, of course, character optimization for the 12 most parsimonious RbcL amino acid-based trees indicated that "loss of pyrenoid" constituted a synapomorphy for *rbcL* clade C, whereas "regain of pyrenoid" constituted a synapomorphy for *rbcL* clade D (Fig. 4). Furthermore, two autopomorphic amino acid residues (at positions 142 and 252) were resolved in the distal branch bearing *Cd. clathrata*.

In contrast, phylogenetic relationships based on the amino acid sequences encoded by the *atpB* or *psaB* genes did not resolve either *rbcL* clade C or *rbcL*

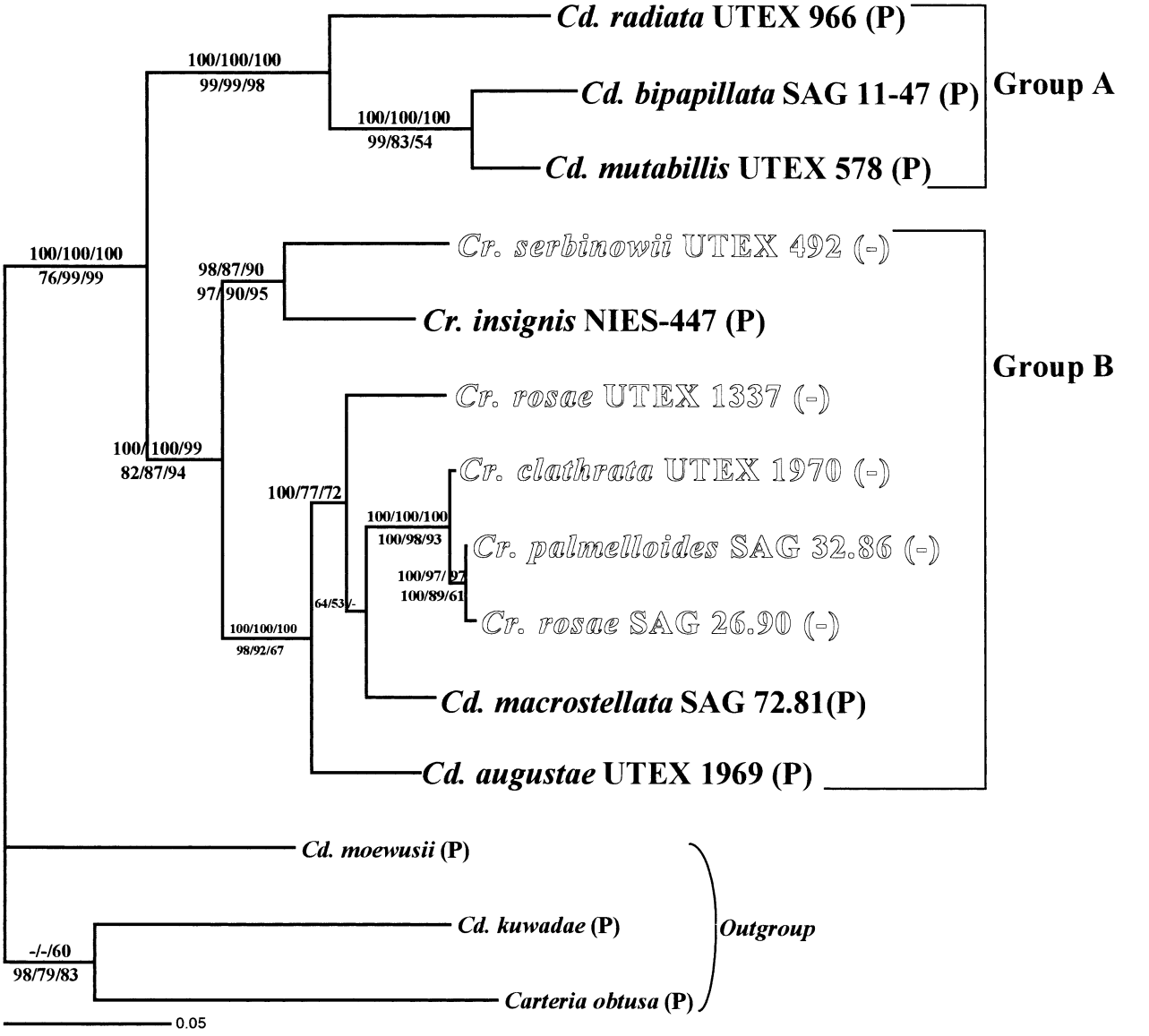
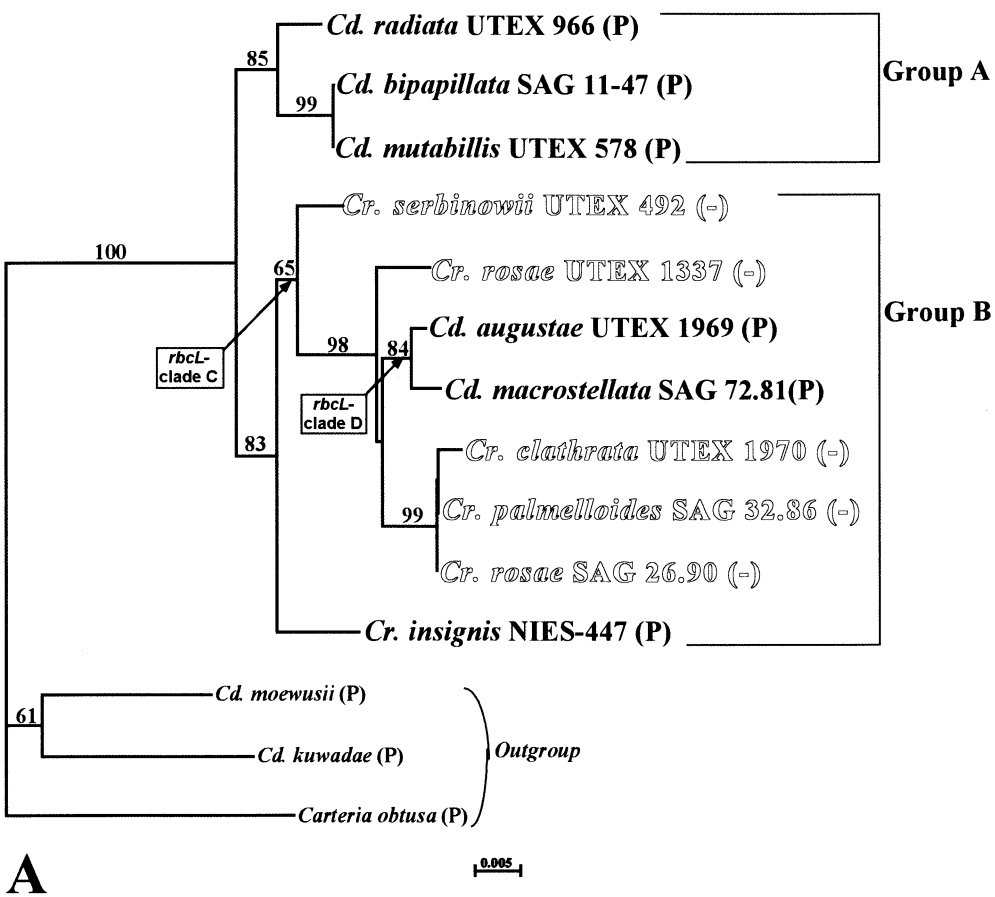


Fig. 6. Maximum likelihood (ML) tree of 11 strains of the *Chloromonas* lineage (*Cd*., *Chlamydomonas*; *Cr*., *Chloromonas*) and two species of *Chlamydomonas* plus *Carteria obtusa* designated as the outgroup (Table 1), based on the aligned nucleotide sequences for 1392 base pairs in coding regions of *psaB* genes. The single tree was obtained using the F81 model, based on the heuristic search using stepwise addition of 10 random replications (with the TBR branch-swapping algorithm) by PAUP 4.0b8 (Swofford 2001). Branch lengths are proportional to the F81 distances, which are indicated by the scale bar below the tree. Numbers at the left, in the middle, and at the right above branches represent 50% or higher bootstrap/QPS values (based on 1000 replications) of the ML, NJ (based on the Jukes–Cantor distances), and MP analyses, respectively, of the *rbcL* nucleotide sequences. Branches also resolved in the consensus bootstrap tree of NJ (based on Kimura distance), MP, and ML (based on the Dayhoff model) analyses of the deduced amino acid sequences (464 amino acids) of the *rbcL* genes from the same 14 strains are presented by bootstrap/QPS values (50% or higher) based on 1000 replications shown below the branches (ML/NJ/MP). The P or – in parentheses after the species/strain designation represents the presence or absence of pyrenoids, respectively (based on Morita et al. 1998, 1999).

clade D. The *psaB* amino acid-based tree indicated with high bootstrap/QPS values (90–98% in ML and NJ analyses) that Group B was composed of two sister monophyletic groups, one of which contained *Cr. insignis* and *Cr. serbinowii*, in agreement with a conclusion derived from analysis of the corresponding nt sequences (Fig. 6). The *atpB* amino acid sequences did not resolve such groups within Group B robustly (Fig. 5).

Discussion

Nonsynonymous nucleotide substitutions and amino acid replacements were much more abundant in the *rbcL* genes of members of the *Chloromonas* lineage than they were in the 43 other taxa of CW algae that we analyzed (Table 2). In addition, relative rate tests indicated a significant difference ($p < 1\%$) in the abundance of amino acid substitutions in 73% of



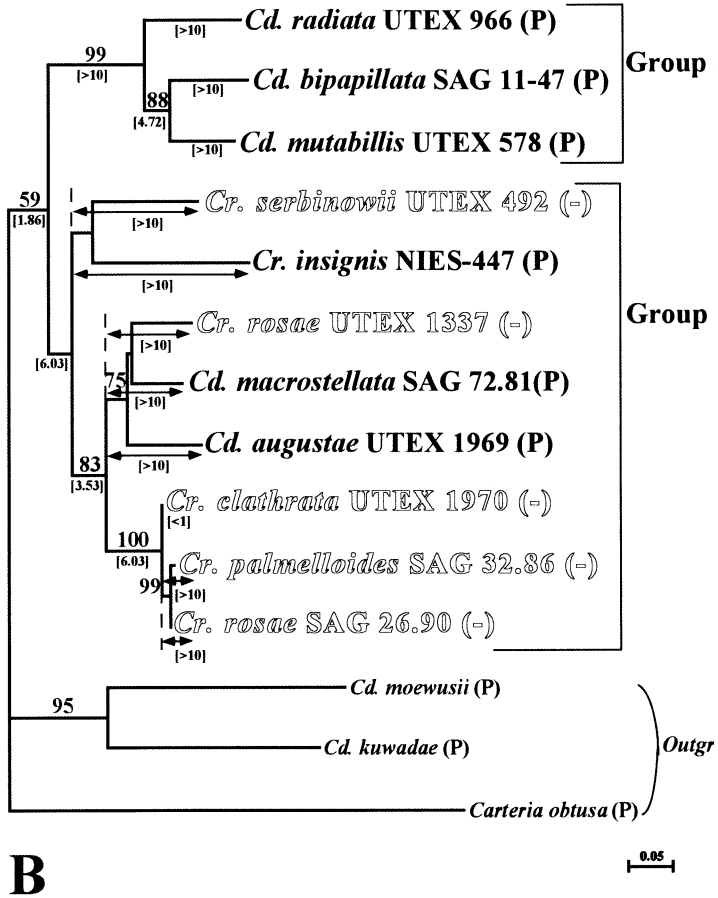


Fig. 7. Comparison of neighbor-joining (NJ) trees based on the nonsynonymous (**A**) and synonymous (**B**) substitutions (K_a and K_s, respectively) for 1128 base pairs of the *rbcL* genes from 11 strains of the *Chloromonas* lineage (*Cd.*, *Chlamydomonas*; *Cr.*, *Chloromonas*) and two species of *Chlamydomonas* plus *Carteria obtusa* designated as the outgroup (Table 1). Numbers above branches represent 50% or higher bootstrap values (based on 1000 replications) based on the NJ analyses of the substitutions. Branch lengths are proportional to K_a (**A**) and K_s (**B**), which are indicated below the trees. K_s/K_a ratios at branches calculated based on comparison between these two trees are indicated by numbers with brackets below branches (**B**). P or - in parentheses after the species/strain designation represents presence or absence of pyrenoids, respectively (based on Morita et al. 1998, 1999).

all cases in which *rbcL* gene of a member of the *Chloromonas* lineage was compared directly to the *rbcL* gene of some other CW alga (Table 3). However, our data also indicate that this high rate of amino acid replacement in the *rbcL* genes is not a general feature of chloroplast genes in the *Chloromonas* lineage, because it is not observed in either their *atpB* or their *psaB* genes (Tables 2 and 3, Figs. 1–3). Thus, accelerated evolution appears to have occurred rather specifically in the *rbcL* genes of the *Chloromonas* lineage.

Wolfe and dePamphilis (1998) demonstrated that certain taxa of parasitic higher plants have high K_s rates in their *rbcL* genes, probably due to the fact that parasitic or heterotrophic growth results in relaxation of functional constraints on the RbcL encoded by *rbcL*. They speculated that relaxed functional constraints on the chloroplast in parasitic higher plants led at first to an increase in K_s due to increased error rates for plastid DNA replication and/or repair (Wolfe and dePamphilis 1998). One of the genera included in our present study is photoheterotrophic: members of the genus *Astrephomene* require an exogenous source of reduced carbon for growth (Brooks 1972). We found that the K_s rates of the *rbcL* genes in *Astrephomene* were somewhat elevated, whereas the K_a rates were not elevated relative to the average values for the other members of the CW group that we examined (Table 2), comparable to the heterotrophic condition of the higher-plant *rbcL* genes. Thus, the heterotrophy in *Astrephomene* may have resulted in the relaxed functional constraints on *rbcL* genes. However, this consideration does not appear to be relevant to the *Chloromonas* lineage, because the K_s rates for *Chloromonas*-lineage *rbcL* genes were lower than those for *Astrephomene*, and the K_a rates of *Chloromonas*-lineage *rbcL* genes were much higher than those of either *Astrephomene* or the group of 40 other photoautotrophic CW taxa that were analyzed (Table 2). These findings are not unexpected, because all members of the *Chloromonas* lineage grow well without exogenous reduced carbon sources and exhibit levels of Rubisco activity that are consistent with their maximal rates of photosynthesis (Morita et al. 1998, 1999). Therefore, the rapid evolution of the *rbcL* genes in the *Chloromonas* lineage cannot be attributed to the relaxation of functional constraints on RbcL that accompanies heterotrophy, as it apparently can in the parasitic higher plants.

The rapid evolution of *rbcL* genes in the *Chloromonas* lineage is associated with a significant elevation in K_a: the rate of nonsynonymous nucleotide substitutions (Table 2). Positive natural selection can result in enhanced rates of nonsynonymous nucleotide substitutions, of course. But typically positive natural selection results in a K_s/K_a ratio that is lower than 1.0 [as seen, for example, in the major histocompatibility complex class I and class II loci (Hughes and Nei 1988, 1989)]. However, the K_s/K_a ratios for the *rbcL* genes in the *Chloromonas* lineage were always >3 when calculated with reference to other members of the *Chloromonas* lineage or to other algae in the CW group (data not shown). In addition, K_s/K_a ratios calculated at branches of the *Chloromonas* lineage were always >1 than one except for the distal branch bearing *Cr. clatharata* (Fig. 7), which exhibited two autopomorphic RbcL amino acid replacements. Thus, we found no evidence that rapid evolution of the RbcL amino acid sequences in the *Chloromonas* lineage could be attributed totally to typically positive selection. However, the K_s/K_a ratio at the basal branch bearing all the OTUs of the *Chloromonas* lineage was comparatively low (1.86) (Fig. 7B). Therefore, some natural selection might have occurred in the *rbcL* genes of the common ancestor of the *Chloromonas* lineage.

On the other hand, there probably has been a decrease in the strength of stabilizing selection in this lineage. Phylogenetic analyses of the RbcL amino acid sequences gave strong support for the existence of the "pyrenoid-lost" *rbcL* clade C and the "pyrenoid-regained" *rbcL* clade D that had initially been identified on the basis of phylogenetic analysis of the *rbcL* nucleotide sequences (Morita et al. 1999; the present study). Furthermore, character optimization suggested that seven or eight amino acid replacements in the RbcL molecule (together with the pyrenoid loss/regain character) constituted synapomorphies for these two clades. However, neither of these two clades was supported by phylogenetic analyses of the *atpB* and *psaB* gene sequences. Indeed, the *atpB* and *psaB* sequences provided robust support for a very different phylogenetic conclusion: namely, that *Cr. serbinowii* (a member of *rbcL* clade C) and *Cr. insignis* (an *rbcL* clade C outsider) constitute a monophyletic group that is the most basal branch of Group B and that *Cd. augustae* (a member of *rbcL* clade D) represents the next most basal branch of this Group (Figs. 1–3). Thus, *rbcL* clades C and D appear to be artifacts, possibly resulting from a convergent association of certain *rbcL* amino acid residues with the presence or absence of pyrenoids within the *Chloromonas* lineage. We believe that it is significant that although none of the amino acid residues that had been thought to be synapomorphic for *rbcL* clades C and D reside in the active site of Rubisco, four of them (residues 210, 247, 276, and 309) lie in the region where two large subunits make contact with one another in a Rubisco dimer (Knight et al. 1990). Therefore these residues very likely are involved in quaternary interactions of RbcL molecules that help determine whether Rubisco molecules will associate to the extent that a visible pyrenoid will be formed. In fact, RbcL amino acid substitutions in 79

amino acids corresponding to the intradimer interface of RbcL (Knight et al. 1990) were extremely high, especially in Group B of the *Chloromonas* lineage (data not shown), in which evolution of pyrenoid loss can be deduced (Fig. 4). However, even when excluding all of the sequences of these 79 amino acids, phylogenetic analyses of the *rbcL* nucleotide sequences resolved the *rbcL* clades C and D, and a high level of RbcL amino acid substitutions was recognized in the *Chloromonas* lineage (data not shown). Therefore, the rapid evolution of RbcL amino acids of the *Chloromonas* lineage cannot be attributed largely to the amino acid substitutions in such intradimer *rbcL* residues.

Thus, our working hypothesis is that the rapid evolution of the *rbcL* genes in the *Chloromonas* lineage has occurred as a consequence of the following two steps. In the common ancestor of the *Chloromonas* lineage, RbcL amino acid substitutions may have accumulated as a consequence of relaxation of functional constraints and/or natural selection on the region of the RbcL molecule that is not directly involved in subunit association. After the differentiation of Groups A and B of the *Chloromonas* lineage, further RbcL amino acid substitutions may have occurred in Group B as a result of relaxation of functional constrains mainly on the region of the RbcL molecule that is directly involved in subunit association and pyrenoid formation. Such relaxation may have been caused by the acquisition of novel physiological characteristics related to the CCM in this lineage, such as elevated carbonic anhydrase (CA) activity. Palmqvist et al. (1995) demonstrated that the pyrenoid-less green alga *Coccomyxa* has a high level of intracellular CA activity, and Morita et al. (1999) have shown that the pyrenoid-containing *Chlamydomonas* members of the *Chloromonas* lineage have higher CA activity than *Cd. reinhardtii* (Morita et al. 1998).

Disappearance of the pyrenoid during the evolution of the land plants from green algae may also have involved specific amino acid substitutions in RbcL. This possibility would not be easy to test by comparative studies of the *rbcL* genes of land plants parallel to our studies of the *Chloromonas* lineage. The hornworts are the only land plants including closely related lineages that differ with respect to the presence or absence of pyrenoids (e.g., Vaughn et al. 1990), but they exhibit RNA editing in the chloroplast genes (Yoshinaga et al. 1996), which would make deducing amino acid sequences from DNA sequence data difficult.

Since the pyrenoid is believed to be one of the factors contributing to the CCM of most algae (see Morita et al. 1999), loss of pyrenoids during land plant evolution might have occurred when the CCM was no longer necessary for efficient photosynthesis. As plants moved to the land, they encountered a comparatively rich CO_2 supply in the air (Badger 1987) and the CO_2 affinity of Rubisco may have been adequate (Coleman and Espie 1985) to have caused a weakening of the selective forces that had been acting to retain pyrenoids in algal chloroplasts. In algae such as those in the *Chloromonas* lineage, however, a CCM is still required, because of the low abundance of CO_2 in water (Badger 1987). Thus, the loss of pyrenoids during *Chloromonas* evolution must have had a different basis—probably the acquisition of novel physiological characteristics that contributed to the CCM in a new and different manner, which then relaxed the functional constraints on the *rbcL* genes, as discussed above.

It would be interesting to know whether the small subunit of Rubisco (RbcS) may also be involved in the association of Rubisco molecules that is required for pyrenoid formation and whether the *rbcS* genes may also have undergone rapid evolution in the *Chloromonas* lineage. However, no information is available yet regarding *rbcS* gene sequences in these algae. Comparative studies of the *rbcS* genes within the *Chloromonas* lineage and other green algae may be fruitful in the future, as a way of improving our understanding of the physiological and evolutionary relationships between Rubisco sequences and pyrenoid structure and function.

Acknowledgments. We are grateful to Prof. D.L. Kirk (Washington University, St. Louis) for critically reading and correcting the manuscript. Thanks are also due to Prof. F. Tajima (University of Tokyo) for his valuable discussion on the evolution of the genes. The study was supported in part by a Grant-in-Aid for Scientific Research (No. 11440241 to H.N.) and by a Grant-in-aid for scientific Research on Priority Areas (C) "Genome Biology" (No. 1320611) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

Badger MR (1987) CO_2-concentrating mechanism in aquatic phototrophs. In: Hatch MD, Boardman NK (eds) The biochemistry of plants: A comprehensive treatise, Vol 10. Photosynthesis. Academic Press, New York, pp 219–274

Bold HC, Alexopoulos CJ, Delevoryas T (1987) Morphology of plants and fungi, 5th ed. Harper & Row, New York

Brooks AE (1972) The physiology of *Astrephomene gubernaculifera*. J Protozool 19:195–199

Buchheim MA, Buchheim JA, Chapman RL (1997) Phylogeny of *Chloromonas*: A study of 18S ribosomal RNA gene sequences. J Phycol 33:286–293

Bürglin TR (1995) The evolution of homeobox genes. In: Arai R, Kato M (eds) Biodiversity and evolution. National Science Museum Foundation, Tokyo, pp 291–336

Cech TR (1988) Conserved sequences and structures of group I introns: Building an active site for RNA catalysis—a review. Gene 73:259–271

Coleman B, Espie GS (1985) CO_2 uptake and transport in leaf mesophyll cells. Plant Cell Environ 8:449–457

Dron M, Rahire M, Rochaix J-D (1982) Sequence of the chloroplast DNA region of *Chlamydomonas reinhardii* containing the gene of the large subunit of ribulose bisphosphate carboxylase and parts of its flanking genes. J Mol Biol 162:775–793

Ettl H (1983) Chlorophyta I. Phytomonadia. In: Ettl H, Gerloff J, Mollenhauer D (eds) Süßwasserflora von Mitteleuropa, Bd 9. Gustav Fischer Verlag, Stuttgart, pp i–xiv, 1–807

Felsenstein J (1985) Confidence limits on phylogenies: an approach using bootstrap. Evolution 38:16–24

Felsenstein J (2001) PHYLIP (phylogeny inference package) version 3.6a2. Distributed by the author. Department of Genetics, University of Washington, Seattle

Friedl TA (1997) The evolution of the green algae. In: Bhattacharya D (ed) Origin of algae and their plastids. Springer-Verlag, Wien, pp 87–101

Griffiths DJ (1970) The pyrenoid. Bot Rev 36:29–58

Hasebe M, Ito M (2000) Evolution of reproductive organs in vascular plants. In: Kato M (ed) The biology and biodiversity. Springer-Verlag, Tokyo, pp 243–255

Hepperle D, Nozaki H, Hohenberger S, Huss VAR, Morita E, Krienitz L (1998) Phylogenetic position of the Phacotaceae within the Chlamydophyceae as revealed by analysis of 18S *r*DNA and *rbcL* sequences. J Mol Evol 47:420–430

Hughes AL, Nei, M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335:167–170

Hughes AL, Nei M (1989) Nucleotide substitution at major histocompatibility complex class II loci: Evidence for overdominant selection. Proc Natl Acad Sci USA 86:958–962

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic Press, New York, pp 21–132

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge

Knight S, Andersson I, Brändén C-I (1990) Crystallographic analysis of ribulose 1,5-bisphosphate carboxylase from spinach at 2.4 A resolution. Subunit interactions and active site. J Mol Biol 215:113–160

Kuchitsu K, Tsuzuki M, Miyachi S (1988) Characterization of the pyrenoid isolated from unicellular green alga *Chlamydomonas reinhardtii*: Particulate from Rubisco protein. Protoplasma 144:17–24

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: Molecular evolutionary genetics analysis software. Bioinformatics 17:1244–1245

Lambowitz AM, Belfort M (1993) Introns as mobile genetic elements. Annu Rev Biochem 62:587–622

Manhart J (1994) Phylogenetic analysis of green plant rbcL sequences. Mol Phylogenet Evol 3:114–127

McCourt RM, Karol KG, Kaplan S, Hoshaw RW (1995) Using *rbcL* sequences to test hypotheses of chloroplast and thallus evolution in conjugating green algae (Zygnematales, Charophyceae). J Phycol 31:989–995

Morita E, Kuroiwa H, Kuroiwa T, Nozaki H (1997) High localization of ribulose-1,5-bisphosphate carboxylase/oxygenase in the pyrenoids of *Chlamydomonas reinhardtii* (Chlorophyta), as revealed by cryofixation and immunogold electron microscopy. J Phycol 33:68–72

Morita E, Abe T, Tsuzuki M, Fujiwara S, Sato N, Hirata A, Sonoike K, Nozaki H (1998) Presence of the CO_2-concentrating mechanism in some species of the pyrenoid-less free-living algal genus *Chloromonas* (Volvocales, Chlorophyta). Planta 204:269–276

Morita E, Abe T, Tsuzuki M, Fujiwara S, Sato N, Hirata A, Sonoike K, Nozaki H (1999) Role of pyrenoids in the CO_2-concentrating mechanism: comparative morphology, physiology and molecular phylogenetic analysis of closely related strains of *Chlamydomonas* and *Chloromonas* (Volvocales). Planta 208:365–372

Muse SV, Pond SK (2000) HYPHY (Hypothesis Testing Using Phylogenies) version 0.71 beta (http://peppercat.statgen.ncsu.edu/~hyphy/). Department of Statistics, North Carolina State University, Raleigh

Nakazawa A, Krienitz L, Nozaki H (2001) Taxonomy of the unicellular green algal genus *Vitreochlamys* (Volvocales), based on comparative morphology of cultured material. Eur J Phycol 36:113–128

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Nozaki H, Ito M, Sano R, Uchida H, Watanabe MM, Kuroiwa T (1995) Phylogenetic relationships within the colonial Volvocales (Chlorophyta) inferred from *rbcL* gene sequence data. J Phycol 31:970–979

Nozaki H, Ito M, Sano R, Uchida H, Watanabe MM, Takahashi H, Kuroiwa T (1997a) Phylogenetic analysis of *Yamagishiella* and *Platydorina* (Volvocaceae, Chlorophyta) based on *rbcL* gene sequences. J Phycol 33:272–278

Nozaki H, Ito M, Watanabe MM, Takano H, Kuroiwa T (1997b) Phylogenetic analysis of morphological species of *Carteria* (Volvocales, Chlorophyta) based on *rbcL* gene sequences. J Phycol 33:864–867

Nozaki H, Ohta N, Morita E, Watanabe MM (1998a) Toward a natural system of species in *Chlorogonium* (Volvocales, Chlorophyta): A combined analysis morphological and *rbcL* gene sequence data. J Phycol 34:1024–1037

Nozaki H, Ohta N, Yamada T, Takano H (1998b) Characterization of *rbcL* group IA introns from two colonial volvocalean species (Chlorophyceae). Plant Mol Biol 37:77–85

Nozaki H, Ohta N, Takano H, Watanabe MM (1999) Reexamination of phylogenetic relationships within the colonial Volvocales (Chlorophyta): An analysis of *atpB* and *rbcL* gene sequences. J Phycol 35:104–112

Nozaki H, Misawa K, Kajita T, Kato M, Nohara S, Watanabe MM (2000) Origin and evolution of the colonial Volvocales (Chlorophyceae) as inferred from multiple, chloroplast gene sequences. Mol Phylogenet Evol 17:256–268

Nozaki H, Takahara M, Nakazawa A, Kita Y, Yamada T, Takano H, Kawano S, Kato M (2002) Evolution of *rbcL* group IA introns and intron open reading frames within the colonial Volvocales (Chlorophyceae). Mol Phylogenet Evol 23: 326–338

Okada M (1992) Recent studies on the composition and the activity of algal pyrenoids. In: Round FE, Chapman DJ (eds) Progress in phycological research, vol 8. Biopress, Bristol, pp 117–138

Palmqvist K, Sültemeyer D, Baldet R Andrews TJ, Badger MR (1995) Characterization of inorganic carbon fluxes, carbonic anhydrase(s) and ribulose-1,5-biphosphate carboxylase-oxygenase in the green unicellular alga *Coccomyxa*. Planta 197:352–361

Ramazanov Z, Rewat M, Henk MC, Mason CB, Matthews SW, Moroney JV (1994) The induction of the CO_2-concentrating mechanism is correlated with the formation of starch sheath around the pyrenoid of *Chlamydomonas reinhardtii*. Planta 195:210–216

Rozas J, Rozas R (1999) DnaSP version 3: An integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15:174–175

Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Satoh H, Okada M, Nakayama K, Miyaji K (1984) Purification and further characterization and pyrenoid proteins and ribulose 1,5-bisphosphate carboxylase-oxygenase from green alga *Bryopsis maxima*. Plant Cell Physiol 25:1205–1214

Schlösser UG (1994) SAG-Sammlung von Algenkulturen at the University of Göttingen. Catalogue of Strains 1994. Bot Acta 107:111–186

Starr RC, Zeikus JA (1993) UTEX—The Culture Collection of Algae at the University of Texas at Austin. J Phycol 29(2) (Suppl):1–106

Strimmer K, von Haeseler A (1996) Quartet puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. Mol Biol Evol 13:964–969

Sugiyama J, Yokota A, Unami E, Onodera Y, Nishida H, Miura Y, Isobe M, Kasahara Y, Sakamoto K (1998) IAM. Catalogue of strains, 2nd ed. Microbiology Research Foundation, Tokyo

Swofford DL (2001) PAUP* 4.0: Phylogenetic analysis using parsimony, version 4.0b8. Computer program. Distributed by Sinauer Associates, Fitchburg, MA

Tajima F, Takezaki N (1994) Estimation of evolutionary distance for reconstructing molecular phylogenetic trees. Mol Biol Evol 11:278–286

Thompson AS, Rhodes JC, Pettman I (1988) Culture collection of algae and protozoa, catalogue of strains. Natural Environment Research Council, Freshwater Biological Association, Ambleside

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X Windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Tsuzuki M, Miyachi S (1991) CO_2 syndrome in *Chlorella*. Can J Bot 69:1003–1007

Vaughn KC, Campbell EO, Hasegawa H, Owen HA, Renzaglia KS (1990) The pyrenoid is the site of ribulose 1,5-bisphosphate carboxylase/oxygenase accumulation in the hornwort (Bryophyta: Anthocerotae) chloroplast. Protoplasma 156:117–129

Wakasugi T, Nagai T, Kapoor M, Sugita M, Ito M, Ito S, Tsudzuki J, Nakashima K, Tsudzuki T, Suzuki Y, Hamada A, Ohta T, Inamura A, Yoshinaga K, Sugiura M (1997) Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: The existence of genes possibly involved in chloroplast division. Proc Natl Acad Sci USA 94:5967–5972

Watanabe MM, Kawachi M, Hiroki M, Kasai F (2000) NIES-Collection. List of strains. Algae and protozoa, 6th ed. National Institute for Environmental Studies, Tsukuba

Woessner JP, Gillham NW, Boynton JE (1986) The sequence of the chloroplast *atpB* gene and its flanking regions in *Chlamydomonas reinhardtii*. Gene 44:17–28

Wolfe AD, dePamphilis CW (1998) The effect of relaxed functional constraints on the photosynthetic gene *rbcL* in photosynthetic and nonphotosynthetic parasitic plants. Mol Biol Evol 15:1243–1258

Yang RCA, Dove M, Seligy VL, Lemieux C, Turmel M, Narang SA (1986) Complete nucleotide sequence and mRNA-mapping of the large subunit gene of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) from *Chlamydomonas moewusii*. Gene 50:259–270

Yoshinaga K, Iinuma H, Masuzawa T, Ueda K (1996) Extensive RNA editing of U to C in addition to C to U substitution in the *rbcL* transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. Nucleic Acids Res 24:1008–1014