Phycologia (2001) Volume 40 (2), 153–161

Published 10 July 2001

Phylogenetic analyses of Bolidophyceae (Heterokontophyta) using *rbc*L gene sequences support their sister group relationship to diatoms

N. DAUGBJERG[1]* AND L. GUILLOU[2]

[1]*Department of Phycology, Botanical Institute, University of Copenhagen, Øster Farimagsgade 2D, DK-1353 Copenhagen K, Denmark*

[2]*Institut de Ciencies del Mar, Passeig Joan de Borbó, s/n-08039 Barcelona, Spain*

N. DAUGBJERG & L. GUILLOU. 2001. Phylogenetic analyses of Bolidophyceae (Heterokontophyta) using *rbc*L gene sequences support their sister group relationship to diatoms. *Phycologia* 40: 153–161

Nearly complete (> 97%) chloroplast-encoded *rbc*L gene sequences were determined for *Bolidomonas mediterranea* from the eastern Mediterranean Sea and two strains of *B. pacifica* var. *eleuthera* from the equatorial Pacific Ocean. The three sequences were added to an *rbc*L data matrix comprising 45 taxa representing 11 of the 12 recognized classes of Heterokontophyta. Phylogenetic analyses that used maximum parsimony (MP), maximum likelihood (ML) and neighbour-joining (NJ) algorithms supported their sister group relationship to the Bacillariophyceae. Centric and pennate diatoms and the three novel *Bolidomonas rbc*L sequences share a two-amino-acid insertion at position 1336, relative to the homologous position in the complete sequence for the centric diatom *Odontella sinensis*. This synapomorphic character state further supports a common evolutionary history of bolidophyceans and diatoms. Chloroplast- and nuclear-encoded molecular markers thus favour the idea that the ancestor of the diatoms possessed two flagella. The Rubisco spacer region was also determined. Sequences showed no differences among the strains of *B. pacifica* examined but these differed by 25% from *B. mediterranea*. New observations are presented on the distribution of *Bolidomonas* and the phylogenetic relationships of bolidomonads and other heterokont algae are discussed.

INTRODUCTION

The evolutionary history of the heterokont algae has resulted in the formation of a morphologically diverse assemblage of organisms. Some species are single-celled flagellates or coccoids, some form colonies, and others are multicellular with a complex organization of the thallus. The heterokont algae inhabit most ecosystems and dominate in many freshwater and marine environments. Molecular techniques (e.g. nucleotide sequence determination and the use of species-specific DNA probes) are currently being applied to describe the biodiversity of protists at a multitude of taxonomic levels. This effort has also resulted in a substantial contribution to our understanding of the diversity of the heterokont algae (= chromophyte algae, in part). In less than 10 years, four new classes of heterokont algae (Pelago-, Phaeothamnio-, Bolido- and Pinguiophyceae) have been circumscribed, by use of light and electron microscopy in combination with nucleotide gene sequences (Andersen *et al.* 1993; Bailey *et al.* 1998; Guillou *et al.* 1999a; Kawachi *et al.* 2000). Hence, at present we recognize *c.* 12 classes within the Heterokontophyta.

The presence of two opposite rows of tubular flagellar hairs constitutes the synapomorphic character state uniting all auto- and heterotrophic heterokonts, whereas the morphological features used to characterize classes primarily relate to the flagellate state (diatoms being one exception, owing to the near or complete absence of zoids). For autotrophic groups, the composition of photosynthetic pigments (particularly carotenoids) has also been invoked successfully at the class level, but these chemosystematic markers appear to have a complicated evolutionary history (Daugbjerg & Andersen 1997a). Ultrastructural characters regarded as conservative and therefore useful in delimiting the groups of heterokonts include (1) the flagellar apparatus configuration; (2) the number and position of transitional plates in the flagellar transition region; and (3) the presence/absence of a helix (gyres) in the transition region (*sensu* Andersen *et al.* 1991).

A few studies have attempted to determine the phylogeny of the heterokont algae using morphological characters (e.g. Andersen 1991; Saunders *et al.* 1995; Potter *et al.* 1997) and/or carotenoids (Daugbjerg & Andersen 1997a). Such cladistic analyses have generally produced tree topologies with low resolution, and this may in part explain why many attempts have been made to use gene sequence data to reconstruct relationships among the heterokonts. Most of these studies are based on nuclear-encoded small-subunit (SSU) rDNA or chloroplast-encoded *rbc*L gene sequences (e.g. Cavalier-Smith *et al.* 1995; Cavalier-Smith & Chao 1996; van de Peer *et al.* 1996; Daugbjerg & Andersen 1997a; Medlin *et al.* 1997; Bailey *et al.* 1998). Although the branching pattern differs slightly when comparing tree topologies obtained from SSU rDNA and *rbc*L analyses, the genealogies suggest the same sister group relationships for some of the classes. For example, most molecular analyses ally the Xanthophyceae with the Phaeophyceae, although such a relationship is not indicated from comparative studies on morphology and biochemical features (Bold & Wynne 1985). The Phaeothamniophyceae are related to the clade containing the Xanthophyceae, and the Phaeophyceae and the Raphidophyceae form a sister group to the Xantho-/Phaeo-/Phaeothamniophyceae clade. On the other hand, the phylogeny of the chrysophyceans, eustigmatophyceans, dictyochophyceans and pelagophyceans is still uncer-

*Corresponding author (nielsd@bot.ku.dk).

tain, with low (< 50%) bootstrap support for any particular branching pattern (e.g. Saunders *et al.* 1995; Daugbjerg & Andersen 1997a; Andersen *et al.* 1998a, b; Bailey *et al.* 1998).

In order to clarify the phylogeny suggested by the *rbc*L gene for the heterokont algae, we have sequenced two species (three strains) of Bolidophyceae. In particular, we wished to assess the support for the sister group relationship between the Bolidophyceae and the diatoms that was suggested recently by analysis of nuclear-encoded SSU rDNA (Guillou *et al.* 1999a). The Rubisco spacer region and a 156-bp fragment of the flanking *rbc*S gene were also determined.

MATERIAL AND METHODS

Information on growth conditions and extraction of total genomic DNA is outlined in Guillou *et al.* (1999a). The *rbc*L gene in the three *Bolidomonas* strains (Table 1) was PCR amplified as two fragments, by use of primer combinations DPrbcL1-NDrbcL8R and NDrbcL5F-NDrbcS (see Daugbjerg & Andersen 1997b for primer sequences). PCR conditions were denaturation at 94° C for 1 min (initial denaturation 3 min), annealing at 50° C for 1 min and extension at 72° C for 1 min (final extension 6 min). The cycle was repeated 32 times. The PCR reaction volumes, chemicals and their concentrations are provided in Hansen *et al.* (2000). Cleaned PCR products were sequenced with the Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer, Foster City, CA), following the manufacturer's recommendations. Cycle sequencing was performed using an ABI PRISM 377 DNA sequencer (Perkin Elmer). Fragments obtained from respective PCR primers were assembled by use of Sequencher version 3.1.1 (Gene Codes Corporation, Ann Arbor, MI). Both strands were determined for all *Bolidomonas* strains, and the sequences have been deposited in GenBank (Table 1). By use of BioEdit version 4.8.9 (Hall 1999), the *rbc*L nucleotides from *Bolidomonas* species were added to a data matrix comprising 45 heterokont algal taxa. The *rbc*L sequences, representing 3–7 taxa from each of 11 out of 12 recognized classes of Heterokontophyta, were unambiguously aligned and included positions 100–1473, relative to the nucleotide sequence of *Odontella sinensis* (Kowallik *et al.* 1995). The likelihood-mapping method included in Puzzle version 5.0 (Strimmer & von Haeseler 1997) was used to evaluate *a priori* the phylogenetic signal of the data matrix, comprising either all codon positions or first and second codon positions only. For likelihood calculations, we used the HKY85 model that allows unequal base frequencies. Quartets (subsamples of four sequences) were resampled 10,000 times. The alignment including all codon positions was analysed by use of maximum parsimony (MP), maximum likelihood (ML) and neighbour-joining (NJ) methods. In MP analysis, characters were unordered, and all codon positions were weighted equally. Gaps were treated as missing data. The two amino acids inserted at position 1336–1341 were coded as present (1) in diatoms and bolidophytes and as absent (0) in all other taxa included. Cladograms were sought by random addition of sequences (1000 replicates) by use of the heuristic search option with branch swapping (tree bisection reconnection) in PAUP* version 4.0b4a (Swofford 1998). ML analyses were used with fastDNAml version 1.0.8 (Olsen *et al.* 1994). Optimal trees were sought by varying the transition to transversion (Ts/Tv) ratio until the best ML score was reached. ML analyses were repeated five times by use of the Ts/Tv ratio (= 0.7) that gave the best ML score. PAUP* was also used to compute dissimilarity values, which were converted to evolutionary distances by correction for multiple substitutions according to the Kimura two-parameter model. The distance matrix was used to build a tree with the NJ method. Bootstrap analyses (Felsenstein 1985) were performed to assess the robustness of clades (1000 replicates in MP and NJ and 100 in ML). Earlier studies based on *rbc*L sequences agree with the current view that the chloroplast in heterokonts is likely to have originated from a single secondary endosymbiosis, when the ancestor of all photosynthetic heterokonts engulfed and retained a red alga (e.g. Daugbjerg & Andersen 1997b; Medlin *et al.* 1997). Hence, three red algae were chosen as outgroup taxa. A single cryptomonad (*Guillardia theta*) was also included, since its plastid genome has been shown to originate from the red algae (e.g. Douglas & Penny 1999). Table 1 provides strain numbers and GenBank accession numbers for all taxa included in the present study.

RESULTS AND DISCUSSION

The phylogenetic analyses of the Heterokontophyta are based on nearly complete *rbc*L sequences – *c.* 100 bp at the 5′ end of the *rbc*L gene are lacking for most taxa (the first forward primer lies at the very beginning of the gene). Likelihood-mapping analysis showed that resolution was slightly higher when all codon positions were included, compared to the inclusion of the first and second codon positions alone (data not shown). Hence, all codon positions were included in the phylogenetic inferences. Table 2 lists the number of variable and parsimony-informative sites in the first, second and third codon positions among the ingroup taxa (i.e. heterokonts and a cryptomonad) and between ingroup and outgroup (i.e. heterokonts, a cryptomonad and red algae). The number of variable and parsimony-informative sites was only slightly higher when the three outgroup taxa (red algae) were included. Hence, only a few of the total number of parsimony-informative sites are due to the inclusion of the red algae. Third codon positions possess the highest number of variable and parsimony informative sites, owing to mutational saturation from multiple hits.

Maximum parsimony analysis of the unweighted data matrix resulted in a single most parsimonious tree (Fig. 1). The tree topology obtained from ML and NJ analyses is illustrated in Figs 2 and 3, respectively. All methods used to infer phylogeny strongly support monophyly for most of the classes of autotrophic heterokonts included. Only the Phaeothamniophyceae *sensu* Bailey *et al.* (1998) and the Chrysophyceae are not monophyletic in all analyses. The cryptomonad *G. theta* is the sister group to the heterokonts (Figs 1–3). In terms of bootstrap values, the relationships between the classes are generally not well resolved. However, there is one exception: all analyses strongly support a sister-group relationship between the Bolidophyceae and diatoms (bootstrap value = 100%, Figs 1–3). This relationship, first indicated by a nuclear-encoded gene (SSU rDNA; Guillou *et al.* 1999a), supports the idea that the ancestor of the diatoms possessed two flagella (one with tubular hairs and the other naked). We speculate

Table 1. List of species included in the present study. Strain numbers[1–6] and GenBank accession numbers for the *rbc*L sequences are also provided.

Taxon	Strain no.	GenBank accession no.
Bacillariophyceae		
Odontella sinensis (Greville) Grunow	?	Z67753
Rhizosolenia setigera Brightwell	CCMP 1330[1]	AF015568
Skeletonema costatum (Greville) Cleve	CCMP 1332[1]	AF015569
Detonula confervacea (Cleve) Gran	?	AB018006
Thalassiosira nordenskioeldii Cleve	?	AB018007
Cylindrotheca sp.	N1	M59080.1
Chrysophyceae		
Chromulina nebulosa Cienkowski	CCMP 263[1]	AF155876
Chrysocapsa vernalis Starmach	CCMP 277[1]	AF155877
Epipyxis pulchra Asmund & Hilliard	CCMP 382[1]	AF015571
Hibberdia magna (Belcher) R.A. Andersen	CCMP 453[1]	AF015572
Bolidophyceae		
Bolidomonas mediterranea Guillou & Chrétiennot-Dinet	MINB 11E5[2]	AF333977
Bolidomonas pacifica var. *eleuthera* Guillou & Chrétiennot-Dinet	OLI 120SD[2]	AF333978
Bolidomonas pacifica var. *eleuthera* Guillou & Chrétiennot-Dinet	OLI 41SA[2]	AF333979
Dictyochophyceae		
Apedinella radians (Lohman) Campbell	K-0077[3]	AF015573
Pseudopedinella elastica Skuja	CCMP 716[1]	U89899
Rhizochromulina sp.	CCMP 237[1]	AF015574
Eustigmatophyceae		
Eustigmatos magna Hibberd	CCMP 387[1]	AF015575
Nannochloropsis sp.	CCMP 531[1]	AF015577
Nannochloropsis salina (Droop) Hibberd	CCMP 369[1]	AF015576
Vischeria helvetica (Vischer & Pascher) Hibberd	UTEX 49[4]	AF015579
Pelagophyceae		
Aureococcus anophagefferens Hargraves & Sieburth	CCMP 1784[1]	AF117906
Aureoumbra lagunensis Stockwell, De-Yoe, Hargraves & Johnson	CCMP 1681[1]	AF117786
Pelagococcus subviridis Norris	CCMP 1429[1]	AF015580
Pelagomonas calceolata Andersen & Saunders	CCMP 1214[1]	U89898
Pulvinaria sp.	CCMP 292[1]	AF015583
Sarcinochrysis marina Geitler	CCMP 770[1]	AF015584
Phaeophyceae		
Elachista fucicola (Velley) Areschoug	G 115[5]	AF055398
Petalonia fascia (O.F. Müller) Kuntze	?	AB022243
Punctaria plantaginea (Roth) Greville	G 170[5]	AF055410
Scytosiphon lomentaria (Lyngbye) Link	?	AB022238
Stictyosiphon soriferus (Reinke) Rosenvinge	42[5] (= K-0359)	AF055413
Striaria attenuata (Greville) Greville	Sat49	AF055415
Phaeothamniophyceae		
Phaeoschizochlamys mucosa Lemmermann	CCMP 635[1]	AF064747
Phaeothamnion confervicola Lagerheim	CCMP 637[1]	AF064746
Pleurochloridella botrydiopsis Pascher	CCMP 1665[1]	AF069499
Stichogloea globosa Starmach	ACO 1315	AF155584
Tetrasporopsis fuscescens (De Toni) Lemmermann	ACO 12	AF155585
Raphidophyceae		
Chattonella subsalsa Biecheler	CCMP 217[1]	AF015581
Heterosigma akashiwo (Hada) Hada *ex* Sournia	?	X61918
Vacuolaria virescens Cienkowski	SAG1 195[6]	AF015582
Synurophyceae		
Mallomonas asmundae (Wujek & van der Veer) K.H. Nichols	CCMP 1658[1]	AF015585
Synura uvella Ehrenberg *em.* Korshikov	CCMP 870[1]	AF015586
Xanthophyceae		
Botrydiopsis intercedens Vischer & Pascher	UTEX 296[4]	AF015587
Botrydium stoloniferum Mitra	UTEX 156[4]	AF064743
Bumilleriopsis filiformis Vischer	UTEX 309[4]	U89900
Mischococcus sphaerocephalus Vischer	UTEX 150[4]	AF064744
Tribonema aequale Pascher	CCMP 1275[1]	AF084611
Xanthonema debile (Vischer) P.C. Silva	UTEX 155[4]	AF084612
Cryptophyceae		
Guillardia theta D.R.A. Hill & Wetherbee	?	AF041468
Rhodophyceae (=outgroup)		
Mastocarpus papillatus (C. Agardh) Kützing	?	U04028
Porphyra purpurea (Roth) C. Agardh	Avonport	U38804
Porphyridium aerugineum Geitler	?	X17597

Cultures originally obtained from [1] Provasoli-Guillard National Center for Culture of Marine Phytoplankton; [2] Roscoff Culture Collection of Marine Phytoplankton; [3] Scandinavian Culture Collection of Algae and Protozoa; [4] Culture Collection of Algae at the University of Texas; [5] P.M. Pedersen's culture collection; and [6] Sammlung von Algenkulturen der Universität Göttingen.

Table 2. Number of variable and parsimony informative sites (first, second and third codon positions, respectively) for ingroup taxa (= heterokont algae and a cryptomonad) and ingroup and outgroup taxa (= heterokont algae, a cryptomonad and red algae). Based on 1373 positions of the *rbc*L gene.

	First (variable/ informative)	Second (variable/ informative)	Third (variable/ informative)
Ingroup taxa	224/188	148/115	428/408
Ingroup and outgroup taxa	231/193	153/119	433/412

that early in the evolutionary history of the diatoms the ancestor has undergone a tremendous secondary reduction in flagellar apparatus features. This reduction has resulted in the retention of flagella only in the male gametes in the centric diatoms. All traces of flagella have been lost in the pennate diatom lineages.

There seems to be moderate support for the idea that the Raphidophyceae is sister group to the clade containing Xantho-/Phaeo-/Phaeothamniophyceae (67% in MP, 71% ML and 63% in NJ). Unfortunately, increasing the taxon sampling by adding novel *rbc*L sequences from a new lineage of heterokonts did not further unravel the evolutionary history of the deep nodes. The bolidophyceans and diatoms form a sister group to a clade containing raphido-, phaeo-, phaeothamnio- and xantho-phyceans: the early divergence of diatoms found by use of nuclear-encoded SSU rDNA sequences (e.g. Andersen *et al.* 1998a, b; Guillou *et al.* 1999a) is not supported by *rbc*L sequence analyses. In the NJ analysis the Pelagophyceae form a sister group to the clade containing the Bolidophyceae and diatoms (supported by a bootstrap value of 78%, Fig. 3). MP and ML analyses suggest an early divergence of the Eustigmatophyceae and Chrysophyceae/Synurophyceae clades (Figs 1–2), but phylogenetic inference with use of NJ favours an early divergence of the Dictyochophyceae (Fig. 3). A previous study, which also used *rbc*L gene sequences but included only the first and second codon positions, revealed an early divergence for the eustigmatophyceans (Daugbjerg & Andersen 1997a). The relatively early divergence of the chrysophycean/synurophycean clade seen in Figs 1–2 differs from the results of Daugbjerg & Andersen (1997a), where this clade formed a sister group to the diatoms; the inclusion of bolidophyceans could explain the difference. MP and ML analyses indicate that the Chrysophyceae are not monophyletic, since *Chrysocapsa vernalis, Hibberdia magna* and *Epipyxis pulchra* are more closely related to the Synurophyceae (*Synura uvella* and *Mallomonas asmundae*) than to *Chromulina nebulosa* (Figs 1–2). In contrast, NJ analysis implies monophyly of the Chrysophyceae (Fig. 3). A recent study by Andersen *et al.* (1999), based on SSU rDNA, also showed that the Synurophyceae are embedded within the Chrysophyceae. They concluded nevertheless that it is premature to reduce the Synurophyceae to the order Synurales within the Chrysophyceae. Alhough based on fewer taxa, parsimony and ML analyses of the *rbc*L gene support the transfer of the Synurales back into the Chrysophyceae.

In the present study, we have added the *rbc*L sequence of *Tetrasporopsis fuscescens*, available in GenBank, so that our analyses include a total of five taxa assigned to the Phaeothamniophyceae (*sensu* Bailey *et al.* 1998). Our analyses (Figs 1–3) reveal maximum bootstrap support (100%) for a relationship between *Phaeothamnion confervicola, Stichogloea globosa* and *Phaeoschizochlamys mucosa*. The relationship between *T. fuscescens* and *Pleurochloridella botrydiopsis* is weakly supported by bootstrap analyses (≤ 54 %). All taxa possess electron-opaque vesicles just beneath the plasma membrane and a unique combination of carotenoids (fucoxanthin and heteroxanthin). Overall, however, the Phaeothamniophyceae is shown to be para- or polyphyletic by *rbc*L data (Figs 1–3). In accordance with this, close inspection of high-performance liquid chromatography analyses likewise suggest a close relationship between *Phaeothamnion confervicola, Phaeoschizochlamys mucosa* and another species of *Stichogloea* (*S. doederleinii*), because pigment ratios were observed to be nearly identical, whereas they differed from the pigment ratios obtained for *Pleurochloridella botrydiopsis* (Bailey *et al.* 1998). Hence, pigment ratios and *rbc*L gene sequence analyses (Figs 1–3 and Bailey *et al.* 1998) imply that the classification of the Phaeothamniophyceae needs further consideration.

Ultrastructural features of the flagellar apparatus are considered to be of major evolutionary significance (e.g. Moestrup 1982; Preisig 1989) and have been used in various schemes of protist classifications. Guillou *et al.* (1999a) regarded the location of the basal bodies above the nucleus, lack of a helix in the transition region, the reduced flagellar apparatus (no flagellar roots), the ring-shaped chloroplast nucleoid and the presence of diatoxanthin as significant features, all suggesting that the bolidophytes are more closely related to the diatoms than to any other group of heterokonts. But none of these are unique to bolidophyceans and diatoms. However, in addition to the results of phylogenetic analyses based on SSU rDNA sequences (Guillou *et al.* 1999a), the present study has added two new distinctive traits in favour of a close relationship between the Bolidophyceae and Bacillariophyceae: (1) maximum bootstrap support (100% in MP, ML and NJ) from chloroplast-encoded *rbc*L sequence data; and (2) the shared two-amino-acid insertion at position 1336 in the *rbc*L gene. A previous combined analysis of SSU rDNA sequences and morphological characters revealed a lineage characterized by a very reduced flagellar apparatus (RFA) (Saunders *et al.* 1995). This comprised diatoms, dictyochophyceans and pelagophyceans. In the present study, which differs from that of Saunders *et al.* (1995) by including the Bolidophyceae (which also have RFA) and in being based on nearly complete sequences of *rbc*L, MP, ML and NJ analyses all failed to reveal the RFA group as a monophyletic assemblage (Figs 1–3).

Complete *rbc*L sequences are available from most plant lineages and a search in GenBank revealed that the total length of the gene in many vascular plants is either 1428 or 1434 bp. In Chlorophyta, Glaucophyta and Euglenophyta, *rbc*L is 1428 bp long, whereas it is 1467 bp in Rhodophyta, Cryptophyta and Haptophyta. In heterokonts, the complete gene sequence is also 1467 bp long, except in diatoms, which possess a two-amino-acid insertion at position 1336 (relative to the homologous position of *O. sinensis*). The *rbc*L sequences determined in the three strains of *Bolidomonas* share this insertion, which comprises glycine and proline. The most parsimonious explanation is that this is a synapomorphy for diatoms and bolidophyceans, supporting a close relationship be-

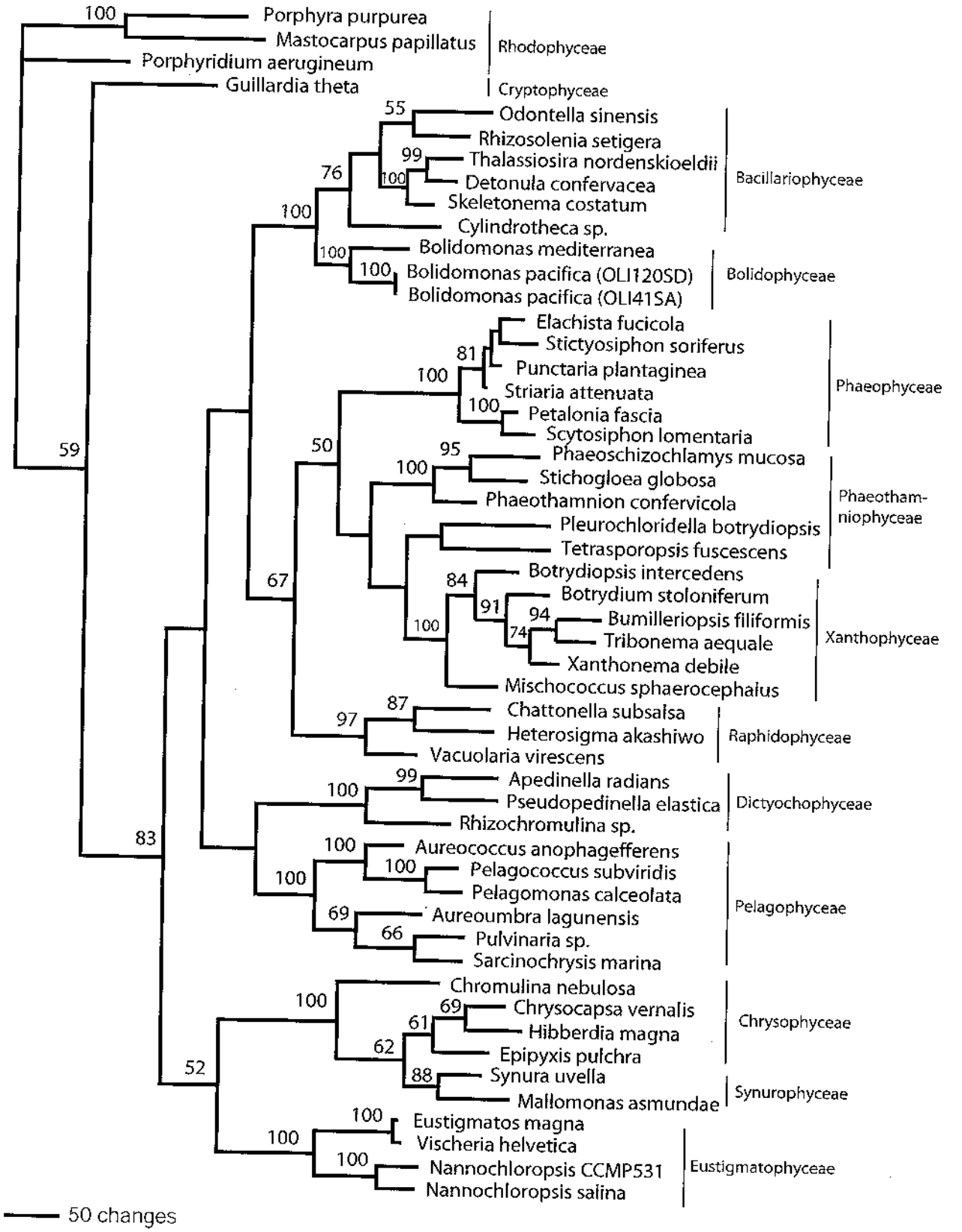

Fig. 1. Phylogeny of 48 autotrophic heterokonts and a cryptomonad inferred from a 1374-bp fragment of the chloroplast-encoded *rbc*L gene, by use of the MP method. The tree was rooted by use of the three red algae. The analysis produced a single most-parsimonious tree (length: 4803 steps; CI = 0.284; RI = 0.569). Branch lengths are proportional to the number of character changes. Bootstrap values (> 50%) from 1000 replicates are shown to the left of internal nodes.

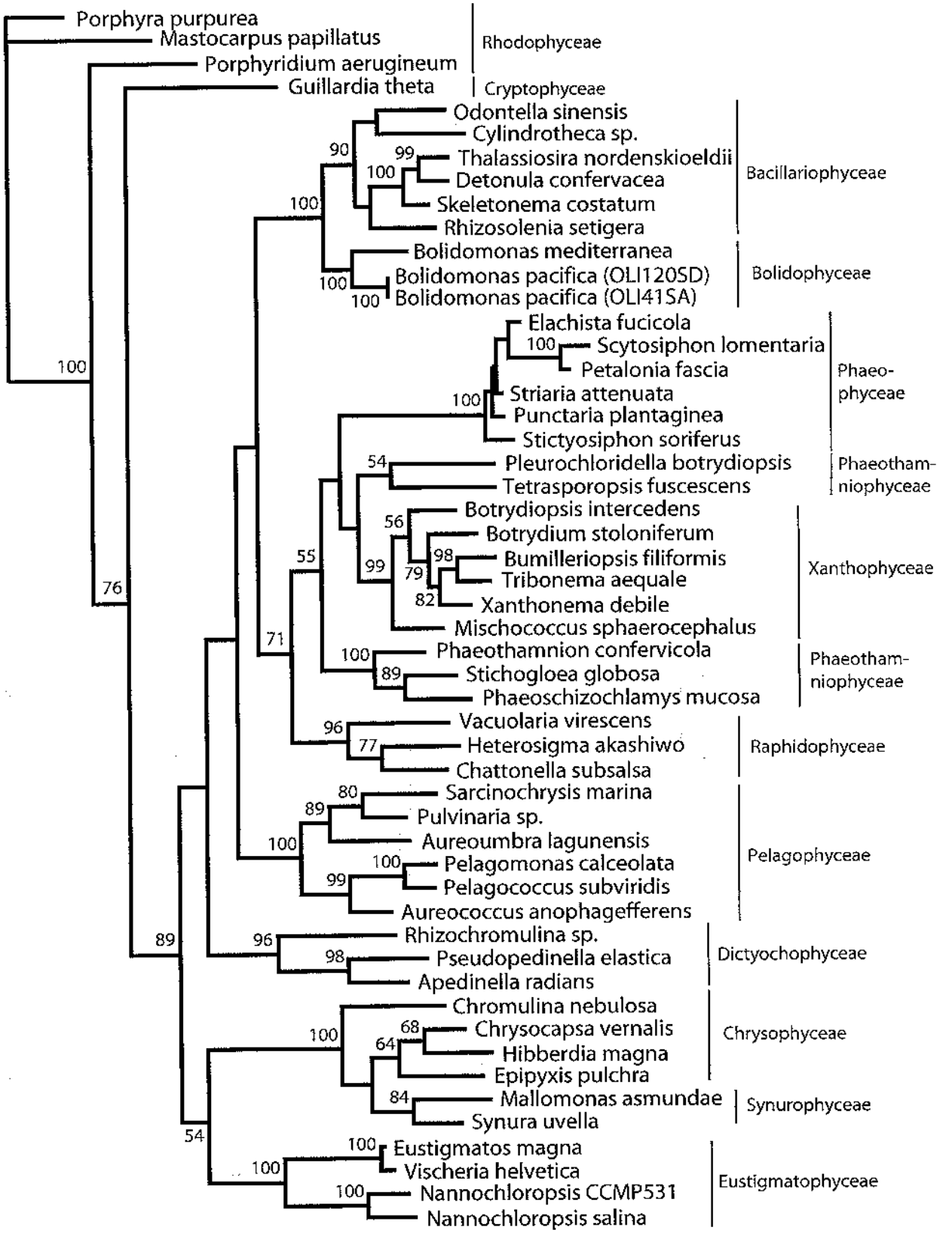

Fig. 2. Phylogeny of 48 autotrophic heterokonts and a cryptomonad inferred from a 1374-bp fragment of the chloroplast-encoded *rbc*L gene, by use of the ML method. The tree was rooted by use of the three red algae. The best −ln likelihood score was 25800.31, obtained with a Ts/Tv ratio equal to 0.7. Branch lengths are proportional to the mean number of nucleotide substitutions per site. Bootstrap numbers (> 50%) from 100 replicates are given to the left of internal nodes.

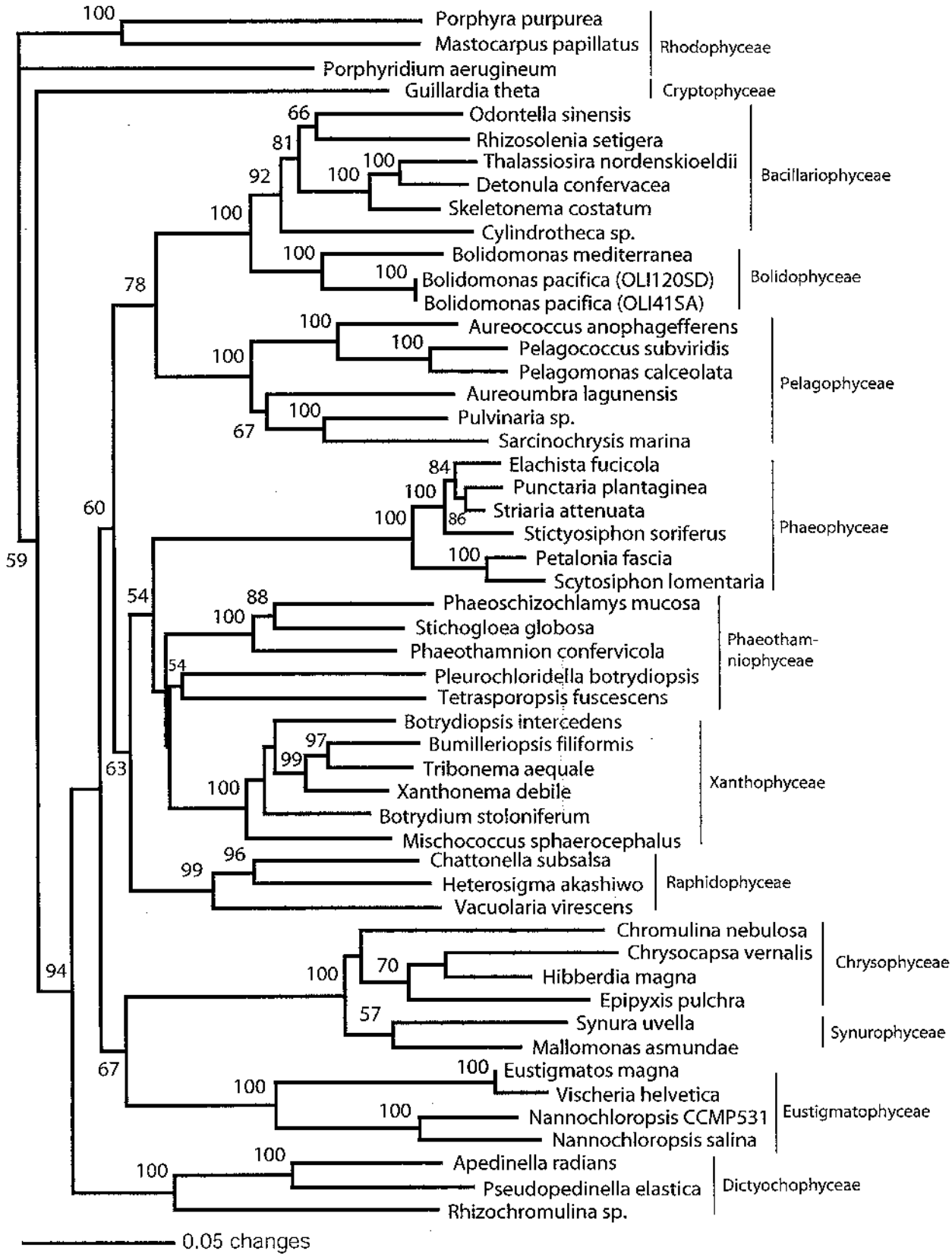

Fig. 3. Phylogeny of 48 autotrophic heterokonts and a cryptomonad inferred from a 1374-bp fragment of the chloroplast-encoded *rbc*L gene by use of the NJ method. Evolutionary distances were corrected for multiple substitutions according to the Kimura two-parameter model. The ingroup was polarized by use of the three red algae. Bootstrap numbers (> 50%) from 1000 replicates are given to the left of internal nodes.

```
Bmed (90.0%)   ----AATTATAATTAAAAATTAATAATTAAGGAGTATTTGAATA
Bop1 (79.0%)   ----CG.-....AA.CT.....C.T.....-.....
Cyli (86.1%)   TAAA..AAG.T.C.TTT..-A..A.T.A.....
Thal (81.6%)   TAAA-----.T.C.TTT.-.AC.C.T.....
Deto (86.8%)   TTAA-----.T.C.TTT.-.AT...T.....
Odon (84.6%)   TAA-----G.T.C.TTT...A..A.TAA.....
```

Fig. 4. Alignment of the Rubisco spacer region in two species of *Bolidomonas* and four diatoms. Numbers in parenthesis refer to A + T content, in %. Possible ribosomal binding sites are marked in bold. Species abbreviations: Bmed = *Bolidomonas mediterranea*; Bop1 = *B. pacifica* (OLI120SD); Cyli = *Cylindrotheca* sp.; Thal = *Thalassiosira nordenskioeldii*; Deto = *Detonula confervacea*; Odon = *Odontella sinensis*.

tween them. Although the first 32 bp in *B. mediterranea* and 46 bp in two strains of *B. pacifica* were not determined, we anticipate that the total length of the *rbc*L gene is identical to that of diatoms (= 1473 bp).

The two strains (OLI120SD and OLI41SA) of *B. pacifica* var. *eleuthera* possessed identical *rbc*L sequences. The SSU rDNA sequences are also identical (Guillou *et al.* 1999b). The *rbc*L sequence divergence between *B. pacifica* and *B. mediterranea* is 7%, which is only slightly higher than that estimated for divergences between three species of *Nannochloropsis* (*N. salina, N. oculata* (Droop) Hibberd and *N.* sp. CCMP 531), which ranged between 3 and 6% (Daugbjerg & Andersen 1997a).

The Rubisco spacer region is identical in the two *B. pacifica* strains examined and had a length of 38 bp (Fig. 4). The spacer region in *B. mediterranea* was slightly longer (40 bp). The spacer region is characterized by a high A + T content (90% in *B. mediterranea* and 79% in *B. pacifica*), and the sequence divergence between the two species is 25% (Fig. 4). In diatoms (*O. sinensis, Detonula confervacea, Thalassiosira nordenskioeldii* and *Cylindrotheca* sp.), the length of the Rubisco spacer was 38–43 bp and the sequence divergence 8–25%. The sequence divergence between diatoms and bolidophyceans was 39–50% (Fig. 4). In diatoms, A + T constitute more than 80% of the nucleotides in the Rubisco spacer. Although the spacer region from a relative low number of diatoms has been determined, we conclude that it is not suitable as a molecular marker in phylogenetic studies. It is too short and the sequence divergence too high for unambiguous alignment. The Rubisco spacer region is again short in the two pelagophyceans *Aureococcus* (39 bp) and *Aureoumbra* (43 bp), with a 30% sequence divergence (Bailey & Andersen 1999). By contrast, the spacer region in brown algae is much longer (151–204 bp) and has been used successfully to study four orders of Phaeophyceae (Siemer *et al.* 1998). A 156-bp fragment at the 5′ end of the *rbc*S gene did not differ in the two *B. pacifica* strains, but 18 substitutions (= 11.5%) were recorded when homologous positions in *B. pacifica* and *B. mediterranea* were compared (data not shown).

Since the original description of the two *Bolidomonas* species as new components of the picoeukaryote plankton in the Mediterranean Sea and the Pacific Ocean, the species have been encountered in coastal waters at Roscoff, France (F. Not, unpublished observations), and in the Weddell Sea, Antarctica (B. Díez, unpublished observations). Although their biogeographical status remains to be fully explored, it is evident that species of *Bolidomonas* are widely distributed. A study on the diversity and abundance of bolidophyceans in oligotrophic areas (Pacific and Mediterranean waters), which used molecular probes and pigment analyses, has revealed that these organisms only constitute a minor proportion of the eukaryote phytoplankton (Guillou *et al.* 1999b). Projects have been initiated to examine the ecological importance of Bolidophyceae in coastal ecosystems.

ACKNOWLEDGEMENTS

We thank Kim B. Pedersen for providing valuable technical assistance. Øjvind Moestrup, two anonymous reviewers and the editor are thanked for useful comments on earlier versions of the manuscript. This study was supported by the Danish Science Foundation to N.D. (96701499) and PICODIV (EVK3-CT1999-00021).

REFERENCES

ANDERSEN R.A. 1991. The cytoskeleton of chromophyte algae. *Protoplasma* 164: 143–159.

ANDERSEN R.A., BARR D.J.S., LYNN D.H., MELKONIAN M., MOESTRUP Ø. & SLEIGH M.A. 1991. Terminology and nomenclature of the cytoskeletal elements associated with the flagellar ciliary apparatus in protists. *Protoplasma* 164: 1–8.

ANDERSEN R.A., SAUNDERS G.W., PASKIND M.P. & SEXTON J.P. 1993. Ultrastructure and 18S rRNA gene sequence for *Pelagomonas calceolata* gen. et sp. nov. and the description of a new algal class, the Pelagophyceae classis nov. *Journal of Phycology* 29: 701–715.

ANDERSEN R.A., BRETT R.W., POTTER D. & SEXTON J.P. 1998a. Phylogeny of the Eustigmatophyceae based upon 18S rDNA, with emphasis on *Nannochloropsis*. *Protist* 149: 61–74.

ANDERSEN R.A., POTTER D., BIDIGARE R.R., LATASA M., ROWAN K. & O'KELLY C.J. 1998b. Characterization and phylogenetic position of the enigmatic golden alga *Phaeothamnion confervicola:* ultrastructure, pigment composition and partial SSU rDNA sequence. *Journal of Phycology* 34: 286–298.

ANDERSEN R.A., VAN DE PEER Y., POTTER D., SEXTON J.P., KAWACHI M. & LAJEUNESSE T. 1999. Phylogenetic analysis of the SSU rRNA from members of the Chrysophyceae. *Protist* 150: 71–84.

BAILEY J.C. & ANDERSEN R.A. 1999. Analysis of clonal cultures of the brown tide algae *Aureococcus* and *Aureoumbra* (Pelagophyceae) using 18S rRNA, *rbc*L and RUBISCO spacer sequences. *Journal of Phycology* 35: 570–574.

BAILEY J.C., BIDIGARE R.R., CHRISTENSEN S.J & ANDERSEN R.A. 1998. Phaeothamniophyceae classis nova: a new lineage of chromophytes based upon photosynthetic pigments, *rbc*L sequence analysis and ultrastructure. *Protist* 149: 245–263.

BOLD H. & WYNNE M.J. 1985. *Introduction to the algae, structure and reproduction.* Prentice Hall, Englewood Cliffs, New Jersey. 720 pp.

CAVALIER-SMITH T. & CHAO E.E. 1996. 18S rRNA sequence of *Heterosigma carterae* (Raphidophyceae), and the phylogeny of heterokont algae (Ochrophyta). *Phycologia* 35: 500–510.

CAVALIER-SMITH T., CHAO E.E. & ALLSOPP T.E.P. 1995. Ribosomal RNA evidence for chloroplast loss within Heterokonta: pedinellid relationships and a revised classification of ochristan algae. *Archiv für Protistenkunde* 145: 209–220.

DAUGBJERG N. & ANDERSEN R.A. 1997a. A molecular phylogeny of the heterokont algae based on analyses of chloroplast-encoded *rbc*L sequence data. *Journal of Phycology* 33: 1031–1041.

DAUGBJERG N. & ANDERSEN R.A. 1997b. Phylogenetic analyses of *rbc*L sequences from haptophytes and heterokont algae suggest their chloroplasts are unrelated. *Molecular Biology and Evolution* 14: 1242–1251.

DOUGLAS, S.E. & PENNY, S.L. 1999. The plastid genome of the cryptophyte alga, *Guillardia theta:* complete sequence and conserved synteny groups confirm its common ancestry with red alga. *Journal of Molecular Evolution* 48: 236–244.

FELSENSTEIN J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

GUILLOU L., CHRÉTIENNOT-DINET M.-J., MEDLIN L.K., CLAUSTRE H., LOISEAUX-DE GÖER S. & VAULOT D. 1999a. *Bolidomonas:* a new genus with two species belonging to a new algal class, the Bolidophyceae (Heterokonta). *Journal of Phycology* 35: 368–381.

GUILLOU L., MOON-VAN DER STAAY S-Y., CLAUSTRE H., FATENSKY F. & VAULOT D. 1999b. Diversity and abundance of Bolidophyceae (Heterokonta) in two oceanic regions. *Applied Environmental Microbiology* 65: 4528–4536.

HALL T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

HANSEN G., DAUGBJERG N. & HENRIKSEN P. 2000. Comparative study of *Gymnodinium mikimotoi* and *Gymnodinium aureolum,* comb. nov. (= *Gyrodinium aureolum*) based on morphology, pigment composition, and molecular data. *Journal of Phycology* 36: 394–406.

KAWACHI M., INOUYE I., HONDA D., O'KELLY C.J., BAILEY J.C., BIDIGARE R.R. & ANDERSEN R.A. 2000. The Pinguiophyceae classis nova, chromophytes algae producing large amounts of omega-3 fatty acids. *Journal of Phycology* 36, supplement: 35.

KOWALLIK K.V., STOEBE B., SCHAFFRAN I., KROTH-PANCIC P. & FREIER U. 1995. The chloroplast genome of a chlorophyll a+c-containing alga, *Odontella sinensis. Plant Molecular Biology Reports* 13: 336–342.

MEDLIN L.K., KOOISTRA W.H.C.F., POTTER D., SAUNDERS G.W. & ANDERSEN R.A. 1997. Phylogenetic relationships of the 'golden algae' (haptophytes, heterokont chromophytes) and their plastids. *Plant Systematics and Evolution* 11, supplement: 187–219.

MOESTRUP Ø. 1982. Flagellar structure in algae: a review, with new observations particularly on the Chrysophyceae, Phaeophyceae (Fucophyceae), Euglenophyceae, and *Reckertia. Phycologia* 21: 427–528.

OLSEN G., MATSUDA H., HAGSTROM R. & OVERBEEK R. 1994. fastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *CABIOS* 10: 41–48.

POTTER D., SAUNDERS G.W. & ANDERSEN R.A. 1997. Phylogenetic relationships of the Raphidophyceae and Xanthophyceae as inferred from nucleotide sequences of the 18S ribosomal RNA gene. *American Journal of Botany* 84: 966–972.

PREISIG H.R. 1989. The flagellar base ultrastructure and phylogeny of chromophytes. In: *The chromophyte algae: problems and perspectives* (ed. by J.C. Green, B.S.C. Leadbeater & W.L. Diver), pp. 167–187. Clarendon Press, Oxford.

SAUNDERS G.W., POTTER D., PASKIND M.P. & ANDERSEN R.A. 1995. Cladistic analyses of combined traditional and molecular data sets reveal an algal lineage. *Proceedings of the National Academy of Sciences of the United States of America* 92: 244–248.

SIEMER B.L., STAM W.T., OLSEN J.L. & PEDERSEN P.M. 1998. Phylogenetic relationships of the brown algal orders Ectocarpales, Chordariales, Dictyosiphonales, and Tilopteridales (Phaeophyceae) based on Rubisco large subunit and spacer sequences. *Journal of Phycology* 34: 1038–1048.

STRIMMER K. & VON HAESELER A. 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proceedings of the National Academy of Sciences of the United States of America* 94: 6815–6819.

SWOFFORD D.L. 1998. *PAUP*. Phylogenetic Analysis Using Parsimony (* and other methods).* Version 4. Sinauer Associates, Sunderland, Massachusetts.

VAN DE PEER Y., VAN DER AUWERA G. & DE WACHTER R. 1996. The evolution of Straminopiles and Alveolates as derived by 'substitution rate calibration' of small ribosomal subunit RNA. *Journal of Molecular Evolution* 42: 201–210.

Accepted 12 February 2001