PHYLOGENY OF DINOFLAGELLATES BASED ON MITOCHONDRIAL CYTOCHROME b AND NUCLEAR SMALL SUBUNIT rDNA SEQUENCE COMPARISONS

Huan Zhang
Department of Marine Sciences, University of Connecticut, Groton, Connecticut 06340 USA

Debashish Bhattacharya
Department of Biological Sciences and Roy J. Carver Center for Comparative Genomics, University of Iowa, 312 Biology Building, Iowa City, Iowa 52242-1324 USA

and

Senjie Lin
Department of Marine Sciences, University of Connecticut, Groton, Connecticut 06340 USA

Despite their evolutionary and ecological importance, dinoflagellate phylogeny remains poorly resolved. Here we explored the utility of mitochondrial cytochrome b (cob) in inferring a dinoflagellate tree and focused on resolving the relationship between fucoxanthin-and peridinin-containing taxa. Trees were inferred using cob and small subunit rDNA alone or in combination as concatenated data and including members of the six major dinoflagellate orders. Many regions of the cob DNA or protein and rDNA trees were congruent with support for the monophyly of Symbiodinium spp. Freudenthal and of the Prorocentrales and the early divergence of Crypthecodinium cohnii Seligo in Grasse. However, these markers provided differing support for the monophyly of Pfiesteria spp. Steidinger et Burkholder (only supported strongly by rDNA) and of the fucoxanthin dinoflagellates with Akashiwo sp. (Hirasaka) Hansen et Moestrup (Gymnodiniales, only supported strongly by the cob data). The approximately unbiased (AU) test was used to assess these results using 13- and 11-taxon (excluding apicomplexans) backbone maximum likelihood trees inferred from the combined cob + rDNA data. The AU test suggested that our data were insufficient to resolve the phylogenetic position of Symbiodinium spp. and that the ancestral position of C. cohnii might have resulted from long-branch attraction to the apicomplexan outgroup. We found significant support, however, for the association of fucoxanthin dinoflagellates with Akashiwo sp. The monophyly and relatively derived position of the Gymnodiniales in our cob DNA and protein trees and in the cob + rDNA tree is consistent with the tertiary endosymbiotic origin of the plastid in fucoxanthin dinoflagellates.

Key index words: cob; cytochrome b; dinoflagellates; phylogeny; rDNA; tertiary endosymbiosis

Abbreviations: AU, approximately unbiased; COB, cytochrome b; cob, gene coding for COB; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; ML, maximum likelihood; MP, maximum parsimony; mt, mitochondrial; SSU, small subunit

Dinoflagellates (subphylum Dinoflagellata, phylum Dinozoa) are an evolutionarily diverse group of protists that are closely related to the apicomplexan parasites and together with the ciliates form the crown group Alveolata (Gajadhar et al. 1991, Cavalier-Smith 1998). Dinoflagellates play pivotal roles in the marine ecosystem and are of tremendous economic significance. They are important primary producers, with their abundance second only to diatoms; they support coral reef ecosystems through symbiotic associations; and they include species that cause harmful algal blooms. Dinoflagellates are considered unusual eukaryotes because, among other well-recognized unique cytological features (Hackett et al. 2004a), they harbor a highly reduced plastid genome (approximately 14 genes, Hackett et al. 2004b) of 1–3 gene minicircles (Zhang et al. 1999, Barbrook and Howe 2000, Koumandou et al. 2004), a nuclear-encoded form II RUBISCO of alpha-proteobacterial origin (Morse et al. 1995. Rowan et al. 1996, Zhang and Lin 2003), and a complex RNA editing machinery (Lin et al. 2002). The ecological and economic importance of dinoflagellates, combined with their unusual genetic/genomic traits, renders them an evolutionary group of high biological interest.

Despite the increasing research effort in recent years to study the phylogeny of dinoflagellates and their plastids, many questions remain open regarding the evolution of the "host" cell and the photosynthetic organelle (Hackett et al. 2004a). Resolution of these
issues would be greatly aided by the availability of a resolved dinoflagellate host tree. The small subunit (SSU) rRNA gene (rDNA) has been instrumental in advancing our understanding of dinoflagellate evolution; however, the low resolving power of this gene regarding lineage relationships has resulted in an incomplete understanding of the major splits among dinoflagellates despite a taxonomically broad sampling of data (Saunders et al. 1997, Saldañarriaga et al. 2001).

Recent studies clearly show the advantage of using multiple genes for phylogenetic analysis (Pryer et al. 2001, Mattern 2004, Yoon et al. 2004). Mitochondrial (mt) DNA is a useful candidate for phylogeny reconstruction for several reasons. First, mt genes are generally conserved although are more variable than nuclear genes such as SSU rDNA (Avisé 1994, Saccone et al. 2000, Garesse and Vallejo 2001). Second, the mitochondrion likely emerged at nearly the same time as the nucleus (Gray et al. 1999), rendering mt genes despite their organellar location appropriate for studying eukaryotic host cell phylogeny (Lang et al. 1998, 1999). Third, some mt genes are closer to a molecular clock than SSU rDNA (i.e. they have a relatively constant mutation rate across taxa) and thus are ideal for cross-taxon phylogenetic studies (Saccone et al. 2000). Cytochrome b (cob) is one of the mt genes most widely used for phylogenetic and population genetic analyses (Conway et al. 2000, Taylor and Hellberg 2003). A combination of cob and nuclear genes has provided robust phylogenetic trees for alveolates and other organisms (Séritawa et al. 2000, Rathore et al. 2001). To date, cob has been studied only for three species of dinoflagellates (Lin et al. 2002), and its utility for inferring the dinoflagellate phylogeny has yet to be explored. Here we cloned and sequenced cob for 14 species from six major dinoflagellate orders: Gonyaulacales, Gymnodiniales, Prorocentrales, Peridiniinales, Suessiales, and Dinamoebiales (Pfiesteria and its relatives). We then used the existing and newly obtained cob gene sequences to reconstruct the dinoflagellate tree.

Based on cob and SSU rDNA analyses, we also attempted to address the phylogeny of the Gymnodiniales. This clade includes taxa (e.g. Gymnodinium Stein sensu stricto, Akashiwo spp.) that, like most photosynthetic dinoflagellates, possess peridinin as the main accessory pigment as well as taxa (Karenia spp. Hansen et Moestrup, Karlodinium micrum [Leadbeater et Dodge] Larsen, and Tohayana spp. de Salas, Bolck, Botes and Halegraeff) that lack peridinin but contain chl c2 + c3 and 19′-hexanoyloxyfucoxanthin and/or 19′-butanoyloxyfucoxanthin, similar to the haptophyte algae (Daughbjerg et al. 2000, Tengs et al. 2000, de Salas et al. 2003). It has been postulated that a tertiary plastid endosymbiosis (involving a haptophyte alga) in the common ancestor of fucoxanthin-containing dinoflagellates explains this distribution of pigment characteristics (Tengs et al. 2000). This hypothesis was reinforced by the finding of oxygen evolving enhancer 1 (psbO) (Ishida and Green 2002) and plastid targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (unpublished data) of haptophyte origin in the K. brevis nuclear genome. Despite these data, the monophyly of fucoxanthin-containing taxa has had varying levels of support from analyses of nuclear and chloroplast SSU rDNA (Tengs et al. 2000, Saldañarriaga et al. 2001, de Salas et al. 2003) and remains to be substantiated. The cob analyses presented here allowed us to address the position of fucoxanthin-containing taxa in the dinoflagellate tree.

**MATERIALS AND METHODS**

**Algal cultures.** Cultures of dinoflagellates and other algae used in this study were obtained from several sources (Table 1). The photosynthetic species were grown in f/2 medium, whereas heterotrophic dinoflagellates (Pfiesteria shumwayae [Glasgow et Burkholder, CCMP1827, CCMP1828, and CCMP1835]) were grown with an algal prey (Rhodomonas sp. [Karsten CCMP768]). Seawater was adjusted to 28 psu for most species and to 15 psu for Rhodomonas sp. (Rhodomonas sp., K. micrum, Akashiwo sp., Ceratium sp. Schrank, and the heterotrophic dinoflagellates; and 25 ± 1°C for K. brevis and the two Symbiodinium species (CCMP830, CCMP832). Illumination was provided in a 12:12-h light:dark cycle with a photon flux of around 100 μmol photons · m⁻² · s⁻¹. The growth rate was monitored by microscopic cell counts using a Sedgewick-Rafter chamber (Phycotech, St. Joseph, MI, USA).

**Sample collection and DNA extraction.** Samples were collected when cultures were in the exponential growth phase. Samples of heterotrophic dinoflagellates were collected after feeding was discontinued for 2 days and few cells of the prey alga Rhodomonas sp. were present. Previous studies had shown that Rhodomonas cob and SSU rDNA sequences were distinct from those of dinoflagellates. The PCR primers used to amplify the dinoflagellates genes did not therefore target the Rhodomonas homologs, even if contaminating DNA was present in the samples (Zhang and Lin 2002). The cells were harvested by centrifugation at 3000g at 4°C for 20 min. These cell pellets were subjected to DNA extraction in a buffer containing 0.1 M EDTA, 1% SDS essentially following Zhang and Lin (2002).

**PCR, cloning, and sequencing.** To PCR amplify cob from dinoflagellates, a set of primers was designed from the highly conserved regions of this gene based on previous data (Zhang and Lin 2002). The primer sequences were Dinocob1F (forward), 5'-ATGAAATCTCTATTACAWCATAT-CCTTGTCC-3’, and Dinocob1R (reverse), 5’-TCTCT-TGAGGGAATTTGWMACCTATCCA-3’. The PCR reaction was done using approximately 50 ng of genomic DNA for each species, and amplification was carried out with a single incubation for 1 min at 95°C, followed by 40 cycles of 20s at 94°C, 30 s at 55°C, and 40s at 72°C. For several dinoflagellates, the SSU rDNA was also amplified using the universal primers (18ScomF1 [forward], 5’-GCTTGTCTAGTAAAGATTGAAC-3’, 18ScomR1 [reverse], 5’-CACCTACGGAAACCTTGTTACGAC-3’) or the combination of the universal primers with the dinoflagellate-specific primers (dino18S5F1 [forward], 5’-AAGGCTGTCTTTATAGN-TACAGAAC-3’, dino18SR1 [reverse], 5’-GAGCCAGATRC-WCCACCGC-3’). The PCR products were purified using DNA Clean & Concentrator (Zymo Research, Orange, CA, USA) and directly sequenced on both strands using BigDye reagents and an ABI Prism automated sequencer (Perkin Elmer, Branchburg, NJ,
PHYLOGENY OF DINOFLAGELLATES

Table 1. Dinoflagellate species included in the phylogenetic analyses.

<table>
<thead>
<tr>
<th>Taxa</th>
<th>Strain and source</th>
<th>Trophic mode</th>
<th>SSU rDNA</th>
<th>cob</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Alexandrium tamarense</em></td>
<td>CB307; D.</td>
<td>P</td>
<td>AY456116</td>
<td></td>
</tr>
<tr>
<td>(Lebour) Balech</td>
<td>M. Anderson</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Akashiwo</em> sp. (Hirasaka)</td>
<td>LISI</td>
<td>P</td>
<td>AY456105</td>
<td></td>
</tr>
<tr>
<td><em>Hansen et Moestrup</em></td>
<td>P</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Ceratium</em> sp. Schrank</td>
<td>MLI</td>
<td>P</td>
<td>AY456107</td>
<td></td>
</tr>
<tr>
<td><em>Cryptothecium</em> cohni*</td>
<td>WHd; M. Gray</td>
<td>H</td>
<td>AF032290</td>
<td></td>
</tr>
<tr>
<td>Seligo in Grasse</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Heterocapsa</em> triquetra*</td>
<td>CCPM449d</td>
<td></td>
<td>AY481575</td>
<td></td>
</tr>
<tr>
<td>(Ehrenberg) Stein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Karoldinium</em> micrum*</td>
<td>D. K. Stoecker</td>
<td>M</td>
<td>AY345908</td>
<td></td>
</tr>
<tr>
<td>(Leadbeater et Dodge) Larsen</td>
<td>CCP</td>
<td>P</td>
<td>AY456104</td>
<td></td>
</tr>
<tr>
<td><em>Karenia</em> brevis* Hansen et</td>
<td>M2994d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moestrup</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Lingulodinium</em> polyedrum*</td>
<td>CCPM405d</td>
<td></td>
<td>AY456109</td>
<td></td>
</tr>
<tr>
<td><em>P. shumwayae</em> Glasgow et</td>
<td>CCPM1834d</td>
<td>H</td>
<td>AF357519</td>
<td></td>
</tr>
<tr>
<td>Burkholder</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>P. striata</em></td>
<td>CCPM1835d</td>
<td></td>
<td>AY456102</td>
<td></td>
</tr>
<tr>
<td><em>P. minimum</em></td>
<td>CCPM1827d</td>
<td>H</td>
<td>AY456117</td>
<td></td>
</tr>
<tr>
<td>(Pavillard) Schiller</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pyrodinium</em> bahamense*</td>
<td>CCMP696d</td>
<td>P</td>
<td>AY456111</td>
<td></td>
</tr>
<tr>
<td><em>Scotia</em> sp. <em>Balech</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Symbiodinium</em> microadriaticum*</td>
<td>CCPM830d</td>
<td></td>
<td>AY456111</td>
<td></td>
</tr>
<tr>
<td><em>Freudenthal</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Symbiodinium</em> sp.</td>
<td>CCPM832d</td>
<td>P</td>
<td>AY456112</td>
<td></td>
</tr>
<tr>
<td>Freudenthal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The cob sequences from the remaining taxa in Figure 1 are publicly available from GenBank.
*Isolated from Long Island Sound in April 2003. Isolates LISI and LIS2 of *Akashiwo* sp. were morphologically similar except that LISI is the smaller cell. Both resemble *Akashiwo sanguinea* based on the description in Steidinger and Tengen (1997) and SSU rDNA sequence. The LIS isolate of *Scrippsiella* sp. is most closely related to *S. tridichotoma* based on morphology (Steidinger and Tengen 1997) and SSU rDNA identity (99% [1738bp/1754bp] to AY421792).
*Isolated from Mirror Lake on the Storrs campus of the University of Connecticut; it is morphologically similar to *Ceratium hiurondella* based on Graham and Wilcox (2000).
P: photoautotrophic; H: heterotrophic; M: mixotrophic.
*Sequences obtained from other strains of the same species.
*Sequences determined in this study.

USA). In some cases, the PCR products were cloned into a TA plasmid vector and the plasmid was propagated using XL1 blue cells (Stratagene, La Jolla, CA, USA). To identify potential PCR-related polymorphisms, plasmids were isolated from 5 to 10 colonies, sequenced over both strands, and the sequences compared with each other.

Phylogenetic analysis. Sequences were aligned using the CLUSTAL W (1.8) server at the DNA Data Bank of Japan (DDBJ) (http://www.ddbj.nig.ac.jp/Welcome-e.html) using the default values. The alignment was then manually adjusted using the SeaView program to maintain codon integrity and set of parameters for the GTR + I + G model were calculated with TREE-PUZZLE ver. 5.1 (Schmidt et al. 2002) and the same evolutionary model. Global rearrangements and randomized sequence input order with five replicates was used to find the best tree. To test the stability of monophyletic groups in the ML protein tree, 100 bootstrap replicates were analyzed with proml as described above. We also used Bayesian inference of the amino acid data (MrBayes ver. 3.0b4, Huelsenbeck and Ronquist 2001). The mtrev + G model was implemented in this analysis, and Metropolis-coupled Markov chain Monte Carlo from a random starting tree was run for 1,000,000 generations with trees sampled each 1000 cycles. Four chains were run simultaneously of which three were heated and one was cold, with the initial 500,000 cycles (500 trees) discarded as the burn-in. A consensus tree was made with the remaining 500 trees to determine the posterior probabilities at the different nodes in the ML tree.

Thereafter, we analyzed dinoflagellate cob DNA (834 nt), COB protein (237 amino acids), and SSU rDNA (1469 nt) alignments as individual data sets as well as a concatenated data set of cob + rDNA (2503 nt). The apicomplexans, *Plasmodium* spp., were used to root these trees. The COB protein analysis was done as described above except that we also used the unweighted maximum parsimony (MP) method to calculate bootstrap (200 replicates) support values for nodes in the ML tree with PAUP (Swofford 2003). The COB protein sequence was deduced from the cob gene sequence with internal stop codons being treated as missing data; the internal stop codons, ranging 0–2, occurred in some species and they are transcriptionally edited to, or serve as, sensible codons (Lin et al. 2002, unpublished data). Ten heuristic searches were used in the MP analysis with random-addition-sequence starting trees (10 rounds), and tree bisection-reconnection branch rearrangements were done to find the optimal parsimony tree. Best scoring trees were held at each step. For each DNA data set, we implemented the ML method in PAUP (GTR + I + G model) and used the search strategy described above for the MP analysis. The parameters for the GTR + I + G model were calculated with
PAUP using a starting distance (minimum evolution [10 random additions]) tree in each case that was built with a matrix of HKY85 distances. We analyzed 100 bootstrap replicates with the ML method. The DNA data were also analyzed with bootstrap ME-GTR + I + G and unweighted MP methods (2000 replicates). We also did Bayesian analysis of the cob, COB, rDNA, and cob + rDNA data sets using the GTR + I + G model for DNA and the mtrev + I model for the COB protein data as described above, except that 5 million generations were run and the final 1000 trees were used to calculate the posterior probabilities. The third codon positions of cob were excluded in a set of DNA analyses that were run as described above. Finally, analysis of the cob + rDNA data set using the partition homogeneity test (ILD test in PAUP, 1000 replicates) showed significant incongruence between these data sets ($P < 0.001$). However, we chose to combine the data because of substantial controversy regarding the utility of this test (Barker and Lutzoni 2002, Hipp et al. 2004).

Testing the tree topology. To assess the positions of different dinoflagellates in the cob + rDNA ML tree, we generated an ML backbone tree of 13 taxa (from the cob + rDNA data set) that included representatives of the major dinoflagellate groups but excluded the fucoxanthin-containing taxa (K. brevis, K. micrum) and the closely related Akashiwo sp. (AY456107), Symbiodinium spp., and C. cohnii. These clades (except Akashiwo sp.) were then added separately using MacClade (Maddison and Maddison 2002) to each of the 23 available branches in the 13-taxon tree to generate three sets of topologies that addressed all the possible divergence points for the three dinoflagellate clades (i.e. independent of each other). We did the same analysis using the cob data alone for additional assessment of the divergence point of the fucoxanthin dinoflagellates. Finally, we also generated the same sets of topologies in an 11-taxon backbone ML tree that excluded the highly divergent Plasmodium spp. The site-by-site likelihoods for these pools of trees were calculated using the respective data sets and BaseML implemented in PAML ver. 3.13 (Yang 1997) with the GTR + I model and the default settings. The approximately unbiased (AU) test was implemented using CONSEL ver. 0.1f (Shimodaira and Hasegawa 2001) to assign probabilities to the different trees in each pool.

RESULTS

Gene sequence. Dinoflagellate cob is highly conserved (see branch lengths in Figs. 1 and 2), although the gene is more enriched in A+T content (70%) than the other eukaryotes in Figure 1 (about 60% A+T). The cob sequences for the two Long Island Sound isolates of Akashiwo sp. (Daugbjerg et al. 2000) were identical as were their SSU rDNA sequences. These cells were indistinguishable morphologically, although their cell sizes were quite different (length/C2 width, 48.5/C2 5.0/C2 m vs. 78.2/C2 7.9/C2 m; data from 50 cells for each isolate). The cob gene in the two Symbiodinium strains was surprisingly divergent (9% dissimilarity) in comparison with other congeners (e.g. 2% for

---

**Fig. 1.** The ML phylogenetic tree of dinoflagellates and other organisms based on the COB sequence. The results of an ML bootstrap analysis are shown above the branches, and the thicker branches denote a Bayesian posterior probability $>0.95$. Branch lengths are proportional to the number of substitutions per site (see scale bar). The COB sequences are available in GenBank; the accession numbers for those of dinoflagellates are shown in Table 1.

A number of cob indels were identified among the studied species. Three 6-bp insertions were found within about 80 bp of the 3’ terminus of cob: The first insertion was Symbiodinium spp. specific; the second was common to Symbiodinium spp., K. micrum, K. brevis, A. tamarense, and P. bahamense; whereas the third insertion occurred only in C. cohnii and Prorocentrum spp. A 15-bp deletion in the middle of cob in a cryptoperidiniopsoid taxon (strain CCMP1828) was absent in all other dinoflagellates, even in CCMP1827, which is thought to be closely related to CCMP1828.

Phylogenetic analyses. The COB amino acid sequence was used to obtain a global view of the dinoflagellate radiation relative to other eukaryotes. This
tree (Fig. 1) confirmed that the apicomplexans (e.g. *Plasmodium* spp.) are the closest relatives of dinoflagellates and identified *C. cohnii* as the earliest diverging dinoflagellate in our data set. A number of other eukaryotic groups were resolved as monophyletic lineages (animals, fungi, chlorophytes), suggesting that COB is a reasonable marker of eukaryotic phylogeny.

With *Plasmodium* spp. as the outgroup, cob ML trees were generated that focused on the dinoflagellates (Fig. 2). The two phylogenies, derived from DNA (Fig. 2A) and the deduced amino acid sequences (Fig. 2B), respectively, were generally congruent with each other, showing for example an early divergence of *C. cohnii* and the monophyly of *Symbiodinium* spp. and the *Prorocentrum* clades. In both trees, the Gymnodiniales formed a monophyletic group with strong bootstrap support, with a close evolutionary relationship being resolved between the fucoxanthin-containing taxa (*K. brevis*, *K. micrum*) and *Akashiwo* sp. The position of the *Symbiodinium* clade was, however, quite different in the cob DNA and protein trees with the former suggesting a later divergence position and the latter an early divergence (with bootstrap support) after *C. cohnii*. In the rDNA tree, *Symbiodinium* spp. appeared to be derived as well (Fig. 2). The Gymnodiniales were polyphyletic in the rDNA phylogeny (Fig. 3), although there was greater support for the monophyly of the *Pfiesteria* clade and for the Gonyaulacales with this data set. The combined cob + rDNA tree (Fig. 4) merged the results of the two individual data sets, showing moderate to strong support for Gymnodiniales, *Pfiesteria*, and Gonyaulacales monophyly with *Symbiodinium* as sister to the latter clade (however, without bootstrap support). Analysis of the first and second codon positions in the cob data resulted in essentially the same results as shown in Figure 2A (e.g. 89% MP bootstrap support [200 replicates]).

**AU test results.** The results of the AU test were entirely consistent with the ML tree topologies shown in Figures 1, 2, and 3 and with the associated bootstrap support values. The most highly favored position for the fucoxanthin dinoflagellates in the 13-and 11-taxa ML backbone trees in the analysis of the cob + rDNA and the cob data sets was as sister to *Akashiwo* sp. Placing *K. brevis* and *K. micrum* on virtually all other branches resulted in trees of significantly lower probability (Fig. 5, A, B, E, and F), except when they were placed at the base of *Prorocentrum* spp. with the cob + rDNA data and as sister to *H. triqueta* with the cob data. Exclusion of the divergent *Plasmodium* spp. markedly increased the probability of the favored position as sister to *Akashiwo* sp. (e.g. from $P = 0.483$ to $P = 0.984$ in the cob data, Fig. 5, B).
The early divergence of *C. cohnii* in our trees was very strongly supported with the AU test, with this position having a probability of 0.968 when *Plasmodium* spp. were included in the analysis. However, exclusion of this outgroup resulted in the change of the most favored position of *C. cohnii* to sister to *Pfiesteria* spp. (*P* ≤ 0.755, Fig. 5G). This surprising result leaves unresolved the phylogenetic position of *C. cohnii* in our analyses and suggests that the apicomplexans provide a potentially misleading signal when included in the trees. Interestingly, divergence at the base of the Gonyaulacales was the second most highly favored position for *C. cohnii* (*P* ≤ 0.583) in Figure 5G. Similarly, the position of *Symbiodinium* spp. was essentially unresolved with our data, and there were many alternative positions (i.e., to that favored as sister to *Prorocentrum* spp., *P* = 0.883, *P* = 0.807; Fig. 5, D and G) for this branch within the 5% confidence set of trees.

**DISCUSSION**

Mitochondrial *cob* has proven to be useful for phylogenetic studies of various organisms (Serizawa et al. 2000, Rathore et al. 2001). Our results suggest that this gene is a useful addition to SSU rDNA for inferring dinoflagellate phylogeny.

**Sequence conservation.** The *cob* gene is highly conserved in general (Zhang et al. 1999, Conway et al. 2000, Taylor and Hellberg 2003). Our results indicate that this gene is also conserved in dinoflagellates. Furthermore, dinoflagellate *cob* has several unique features. First, it is more highly enriched in As and Ts (70%) than the homolog in many other organisms (about 60%). The high A+T content of *cob* has also been found in insects, nematodes, and eumycotes and is thought to result from high A+T pressure (Jermini et al. 1994). The high A+T content of *cob* potentially would reduce translation efficiency, but the problem is ameliorated in dinoflagellates by RNA editing, which results in an increase in G+C content (Lin et al. 2002).

*Cob* phylogeny. The *cob* and rDNA trees provide congruent results in many respects, with both showing an affiliation of dinoflagellates with other alveolates such as *Plasmodium* spp. and similar branching patterns within the Dinophyceae. This is in concert with neighbor joining and MP (not shown). The comparison of *cob* and rDNA trees provides valuable insight, especially for resolving lineages for which previous analyses show conflicting results. Combination of the two genes can potentially resolve many nodes in the dinoflagellate phylogenetic tree.

The different phylogenetic trees suggest an early divergence of *C. cohnii* within the Dinophyceae, but this result is not supported by the AU test when the divergent *Plasmodium* spp. are excluded from the analysis. Therefore, our data do not exclude kinship of *C. cohnii* to the other Gonyaulacales in the tree; rather, a position as sister to the Gonyaulacales has a high, albeit
nonsignificant, probability of $P = 0.583$. Interestingly, the divergent apicomplexans appear to have a strong effect on the position of *C. cohnii*, with the latter being attracted to the former when they are included in the trees as an outgroup (Fig. 5C) but essentially not having a single highly supported position when they are removed (Fig. 5G). *Cryptothecodinium cohnii* has been placed within Gonyaulacales based on characteristics of thecal plate tabulation and some SSU rDNA data (Fensome et al. 1993, Saldarriaga et al. 2001) but is considered an early branch of dinoflagellate evolution by other analyses (Saunders et al. 1997, Litaker et al. 1999). Conflicting results still remain in more recent studies based on multiple protein encoding genes (Saldarriaga et al. 2003, Leander and Keeling 2004). Our analyses of *cob* do not convincingly resolve this issue, suggesting that more taxa (and/or data) are needed to determine the divergence point of *C. cohnii*.

*Symbiodinium* is a highly complex lineage whose members are resolved to eight phylotypes (clades) on
the basis of rDNA and chloroplast DNA (Ishikura et al. 2004, LaJeunesse 2004), but few species have been formally described. The two strains (CCMP830, 832) analyzed in this study form a monophyletic group strongly supported in all our trees, although SSU rDNA analyses indicate they belong to two different clades (B and A for CCMP830 and 832, respectively; unpublished data). However, affiliation of this lineage to other dinoflagellates remains unresolved (Fig. 5, D and H). A close evolutionary relationship between 

Symbiodinium and Gymnodiniales has been suggested on the basis of a similar degree of thecal plate development (Steidinger and Tengen 1997). In contrast, a more recent tree based on actin and β-tubulin indicated an early divergence of this taxon (Saldarriaga et al. 2003), as seen in Figure 2B in our study. These data suggest that similarity in thecal plate development may not reflect the degree of genetic relatedness and calls for additional analyses including more species of Sym-

biodinium and related dinoflagellates to determine the precise divergence point of this important lineage. More studies are also needed to address the relationship among the phylotypes. The greater difference in cob sequence between the two 

Symbiodinium strains than in other congeneric dinoflagellate species suggests a possibility of their affiliation to different genera. The possibility of resolving 

Symbiodinium into different genera has been suggested (Rodriguez-Lanetty 2003, LaJeunesse 2005).

Phylogenetic status of gymnodiniales and plastid evolution of Dinophyceae. The most interesting result in this study is the clear support for the monophyly of the fucoxanthin-containing taxa 

K. brevis and 

K. micrum with the peridinin-containing 

Akashiwo sp. (Daugbjerg et al. 2000) within the Gymnodiniales. It has been proposed that the fucoxanthin-containing taxa may share a common origin (Tengs et al. 2000), but their mon-

ophyly has received variable degrees of bootstrap sup-

port from analyses of nuclear and plastid SSU rDNA (Daugbjerg et al. 2000, Tengs et al. 2000, de Salas et al. 2003). The evolutionary position of these dinoflagel-

lates relative to peridinin-containing taxa was therefore not firmly established. In a study using the micricle-

encoded plastid gene 

psbA, Yoon et al. (2002) suggested that peridinin and fucoxanthin dinoflagellate might share an ancestral haptophyte plastid that resulted from a tertiary endosymbiosis in their common ances-

tor. Our results (given that cob is a vertically inherited marker of host phylogeny) contradict this latter hy-

pothesis (Inagaki et al. 2004) and provide clear molecular phylogenetic evidence for the position of 

fucoxanthin dinoflagellates as a derived clade within peridinin-containing taxa.

The derived position of the fucoxanthin-containing Gymnodiniales relative to peridinin-containing taxa (including members of this order) strongly suggests that the haptophyte-type plastid is a derived trait that was acquired in 

Karenia spp., 

K. micrum, and, putative-

ly, 

Takayama spp. (de Salas et al. 2003). These results are consistent with the finding of a 

psbO (Ishida and

Green 2002) and a plastid targeted GAPDH (unpub-

lished data) gene of haptophyte origin in the nuclear genome of 

K. brevis (Ishida and Green 2002). Under this scenario, the haptophyte 

psbO and 

GAPDH sequences, which are presumably of tertiary endosymbiotic origin, replaced the nuclear encoded genes of red algal (i.e. chromalveolate) provenance that are nor-

mally found in peridinin dinoflagellates (A. tamarense) (Hackett et al. 2004b). Furthermore, a broader multi-

gene phylogenetic analysis of five minicircle proteins now also strongly supports the haptophyte origin of the dinoflagellate fucoxanthin plastid through tertiary endosymbiosis (unpublished data).

We thank D. M. Anderson (WHOI), P. Tester (National Ocean Services), D. K. Stoeker (Horn Point Environmental Labora-

tory), and R. Azanza-Corales (Marine Science Institute, Uni-

cersity of the Philippines at Diliman) for providing cultures of 

Alexandrium tamarense, P. Ehrenbergia shumwayae, Karodinium micrum, and 

Pyrodinium behamense, respectively. Brett Branco and John Bean collected water samples from Mirror Lake, Storr's cam-

pus of the University of Connecticut, from which 

Ceratium sp. was isolated. W. Litaker and P. Tester provided SSU rDNA se-

quences for strain CCMP1828 and CCMP1835. T. N. Feinstein and K. Perkins provided assistance with growing the cultures, and L. Miranda measured the cell size of 

Akashiwo sp. The critical comments of two anonymous reviewers are also deeply appreciated. This research was supported by NSF grant DEB 0541486 (to S. L. and H. Z.), NOAA-ECOHAB grants NA860P0491 and NA16OP2797 (to S. L.), and NSF grants DEB 0107754 and MCB 0236651 (to D. B.). This is ECOHAB publication no. 121.