J. Eukaryot. Microbiol., 49(1), 2002 pp. 30–37

Re-examining Alveolate Evolution Using Multiple Protein Molecular Phylogenies

NAOMI M. FAST,[a] LINGRU XUE,[b] SCOTT BINGHAM[b] and PATRICK J. KEELING[a]

[a]*Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, Vancouver, BC V6T 1Z4, Canada, and* [b]*Department of Plant Biology, Arizona State University, Tempe, Arizona 85287-1601, USA*

ABSTRACT. Alveolates are a diverse group of protists that includes three major lineages: ciliates, apicomplexa, and dinoflagellates. Among these three, it is thought that the apicomplexa and dinoflagellates are more closely related to one another than to ciliates. However, this conclusion is based almost entirely on results from ribosomal RNA phylogeny because very few morphological characters address this issue and scant molecular data are available from dinoflagellates. To better examine the relationships between the three major alveolate groups, we have sequenced six genes from the non-photosynthetic dinoflagellate, *Crypthecodinium cohnii*: actin, beta-tubulin, hsp70, BiP, hsp90, and mitochondrial hsp10. Beta-tubulin, hsp70, BiP, and hsp90 were found to be useful for intra-alveolate phylogeny, and trees were inferred from these genes individually and in combination. Trees inferred from individual genes generally supported the apicomplexa-dinoflagellate grouping, as did a combined analysis of all four genes. However, it was also found that the outgroup had a significant effect on the topology within alveolates when using certain methods of phylogenetic reconstruction, and an alternative topology clustering dinoflagellates and ciliates could not be rejected by the combined data. Altogether, these results support the sisterhood of apicomplexa and dinoflagellates, but point out that the relationship is not as strong as is often assumed.

Key Words. Actin, Apicomplexa, ciliates, dinoflagellates, evolution, heat shock protein, tubulin.

THE alveolates (or Alveolata) are a protist lineage composed of three major groups—ciliates, apicomplexa, and dinoflagellates—as well as a number of other genera, such as *Perkinsus*, *Colpodella*, and *Colponema*. Together, the alveolates are one of the largest and most diverse protist assemblages presently recognized, including over ten thousand named species and occupying practically every niche where protists are found (Lee, Hutner, and Bovee 1985; Levine 1988; Taylor 1987). At first glance, the three major alveolate groups share little in terms of morphology. Ciliates are well-known for their nuclear dimorphism, complex ciliature, and feeding apparatus (Lee, Hutner, and Bovee 1985), while dinoflagellates are characterized by the presence of the dinokaryon (permanently condensed chromosomes with a unique form of closed mitosis) and a highly specialized flagellar dimorphism (Taylor 1987). Apicomplexa are all obligate intracellular parasites and are primarily distinguished by their often complex life history as well as a suite of structures associated with infection (Levine 1988). Ciliates are completely non-photosynthetic (except in rare cases where plastids derived from food are retained), while about half of the characterized dinoflagellates are photosynthetic (Taylor 1987), and apicomplexa have been found to contain a relic, non-photosynthetic plastid (Köhler et al. 1997; McFadden et al. 1996; Wilson et al. 1996). The plastids of dinoflagellates and apicomplexa are both derived from secondary endosymbiosis (Fast et al. 2001; Waller et al. 1998; Zhang, Green, and Cavalier-Smith 1999). It has recently been shown that the plastid likely originated before the origin of all alveolates based on a gene replacement event involving glyceraldehyde-3-phosphate dehydrogenase, GAPDH (Fast et al. 2001) Thus, even though no evidence of a plastid has been found in ciliates, it appears that their ancestors did contain a plastid at one time.

It is clear that alveolates are a diverse group, but they do share several distinct structural features, the most predominant being the cortical alveoli (membrane-bound vesicles just under the cell membrane), and tubular mitochondrial cristae (the latter are also shared with a number of other protists) (Lee, Hutner, and Bovee 1985; Levine 1988; Taylor 1987). It was these features that led to the first recognition of a relationship between dinoflagellates and ciliates (Taylor 1976). However, their relationship with apicomplexa was only recognized following molecular phylogenetic analysis (Gajadhar et al. 1991; Wolters 1991), illustrating the great morphological diversity between these lineages. Presently, the support for an alveolate clade from molecular data is quite strong: ribosomal RNA (rRNA) phylogenies consistently yield a highly supported alveolate grouping (Van de Peer and De Wachter 1997, Van de Peer et al. 2000), as did a recent combined analysis of rRNA and protein data that included ciliates and apicomplexa (Baldauf et al. 2000). In addition, similarities in the divergent regions of the large subunit rRNA structure have been noted among ciliates, dinoflagellates, and apicomplexa (Gagnon, Bourbeau, and Levesque 1996). The exact position of the alveolates within the tree of eukaryotes is not so well-defined, however, but many molecular phylogenies suggest a relationship between alveolates and heterokonts (or stramenopiles) (e.g. Van de Peer and De Wachter 1997). These groups also share a unique plastid gene-replacement (Fast et al. 2001), further supporting the notion that they share a close common ancestry.

In terms of relationships within the alveolates, small subunit (SSU) rRNA phylogenies have consistently shown that dinoflagellates and apicomplexa are more closely related to each other than either group is to ciliates (Van de Peer and De Wachter 1997). This has led to the widespread acceptance of this relationship, which in turn has had a tremendous influence on the interpretation of alveolate evolution, especially relating to the recent finding of a cryptic secondary plastid in apicomplexa (e.g. Saldarriaga et al. 2001; Taylor 1999; Zhang, Green, and Cavalier-Smith 2000). With the growing number of cases where trees from a single gene have turned out to be misleading, it is increasingly important that such relationships be supported by many different pieces of evidence, both molecular and otherwise. This is particularly true in instances such as this, where the question essentially addresses the position of the root of alveolates, as root positions are notoriously difficult to determine (Philippe and Laurent 1998). Unfortunately, however, careful review of the molecular evidence for an apicomplexa-dinoflagellate sister group reveals that there is very little strong evidence for this relationship. The only nuclear protein-coding genes that have been used to test the apicomplexa-dinoflagellate relationship are actin and GAPDH. In GAPDH phylogenies, the branching order within alveolates is not resolved and the apicomplexa do not form a clade with dinoflagellates (Fast et al. 2001). In actin phylogeny, which has been used extensively to examine the relationships within alveolates (Reece et al. 1997; Siddall et al. 1997), apicomplexa and dinoflagellates do form a

Corresponding Author: P. Keeling—Telephone number: + 1 (604) 822-4906; FAX number: + 1 (604) 822-6089; E-mail: pkeeling@interchange.ubc.ca

Nucleotide sequences reported in this paper have been submitted to GenBank under accession numbers AF421536-AF421541.

clade, but global actin phylogenies reveal that known ciliate genes are extremely divergent and do not actually branch with other alveolates (e.g. Keeling 2001). This being the case, it is impossible to make any strong claims about the overall branching order of the major alveolate groups, since the real position of the divergent ciliate sequences cannot be determined. The same can be said for trees inferred from the mitochondrial gene *cox*I, where apicomplexa and dinoflagellates form a clade but depending on the analysis, they do not branch with ciliates (Inagaki et al. 1997) or do branch with ciliates with weak bootstrap support (Norman and Gray 1997). In sum, although dinoflagellates and apicomplexa are generally assumed to be sisters, there is no phylogenetic evidence to support this claim aside from the SSU rRNA phylogeny.

From the perspective of molecular phylogeny, the weakness in this area is clearly the dinoflagellates: large numbers of genes have been characterized from both ciliates and apicomplexa, but only a handful of molecular sequences are known from dinoflagellates (and most of these relate to their unusual plastid). To begin to address this, we have sequenced six genes encoding proteins from three cellular compartments of the nonphotosynthetic, gonyaulacalean dinoflagellate, *Crypthecodinium cohnii*. These new sequences include both the cytosolic and endoplasmic reticulum orthologs of the heat shock protein 70 (hsp70 and BiP respectively), cytosolic hsp90, mitochondrial hsp10, and the cytoskeletal proteins, actin and beta-tubulin. We have used these to construct individual phylogenies, as well as a phylogeny of four concatenated genes, to provide independent evidence for the branching order of ciliates, dinoflagellates and apicomplexa. We have found that the bulk of the analyses favour the conclusion that apicomplexa and dinoflagellates are sisters. However, not all gene trees support this conclusion, and examining the effect of outgroup choice shows that the sister relationship of these two may be less certain than currently assumed, altogether reinforcing the notion that phylogenetic hypotheses need to be confirmed by many independent gene trees.

MATERIALS AND METHODS

Sequencing of *Crypthecodinium cohnii* actin, beta-tubulin, hsp70, BiP, hsp90, and hsp10. cDNA sequences of actin, beta-tubulin, hsp70, BiP, hsp90, and hsp10 were identified in an expressed sequence tag (EST) sequence survey of *C. cohnii*. The *C. cohnii* library was provided by Kirk Apt and Casey Lippmeier at Martek Bioscience. Briefly, *C. cohnii* (Martek Bioscience culture collection strain) was grown in 50 g/L glucose, 6 g/L yeast extract, and 10% artificial sea water at 27 °C and pH 6.7. RNA was isolated by the method of Apt et al. (1995) and cDNA was synthesized using Moloney murine leukemia virus (MMLV) reverse transcriptase and cloned into Lambda-Zap II according to the manufacturer's directions (Stratagene, La Jolla, CA). Random clones were sequenced at the 5′-end using the T7 primer and ABI BigDye chemistry (ABI, Foster City, CA). Clones corresponding to full-length transcripts of actin, beta-tubulin, hsp70, BiP, hsp90, and hsp10 were identified by comparisons with public databases and completely sequenced on both strands by primer walking.

Phylogenetic analysis. Inferred amino acid sequences from actin, beta-tubulin, hsp70, BiP, hsp90, and mitochondrial hsp10 were aligned with the homologous protein sequences from a representative sampling of eukaryotes, making an effort to construct data sets that were comparable in terms of taxon sampling. In each case, alignments were first constructed using the pattern-induced multi-sequence alignment program (PIMA) (Smith and Smith 1992) and were further edited manually. Extremely partial or redundant sequences were not included. Phylogenetic analyses were carried out on datasets of unambiguously aligned amino acid chararacters. For actin, beta-tubulin, hsp70, BiP, hsp90, and hsp10, these datasets were composed of 244, 428, 578, 578, 614, and 82 characters and 96, 52, 45, 30, 39, and 35 taxa respectively. Distances were calculated using PUZZLE 5.0 (Strimmer and von Haeseler 1996) with the JTT substitution matrix, the frequency of amino acid usage calculated from the data, and site-to-site rate variation modelled on a gamma distribution with 8 rate categories and invariable sites. The fraction of invariable sites and shape parameter were estimated from the data. Trees were constructed with BioNJ (Gascuel 1997), weighted neighbor-joining using Weighbor (Bruno, Socci, and Halpern 2000) and Fitch-Margoliash (Felsenstein 1993), in the latter case using global rearrangements and ten input order jumbles. In general, results for each individual gene from these three methods only differed in weakly supported aspects of the topology, so only Fitch-Margoliash trees are shown. One hundred resampled bootstrap datasets were created with SEQBOOT (Felsenstein 1993) and gamma-corrected distances were computed using the parameters above with puzzleboot (shell script by M. Holder and A. Roger: www.tree-puzzle.de). Bootstrap trees were inferred as described above (except that in Fitch-Margoliash analyses only five input order jumbles were carried out).

A combined dataset of 27 taxa and 2198 characters was also constructed by concatenating the sequence alignments of beta-tubulin, hsp70, BiP, and hsp90. Actin was not included in the combined analysis for two reasons. First, ciliate sequences for actin are extremely divergent and do not branch with the other alveolates (see discussion below), a situation that could significantly bias a dinoflagellate-apicomplexan relationship in a combined analysis. Second, the sampling for actin is very different than the other protein coding genes examined; its inclusion would remove half the apicomplexan taxa from the dataset and leave only one ciliate. Taxa were included if they were represented by a minimum of three out of four of the genes and, with one exception, all genera are represented by the same species: *Eimeria* sequences for BiP, beta-tubulin and hsp90 come from *Eimeria tenella*, whereas the hsp70 sequence is from *Eimeria maxima*. Those taxa missing genes are: for beta-tubulin: *Lycopersicon*; for hsp70: *Tetrahymena*; for BiP: *Babesia*, *Dictyostelium*, *Candida*; and for hsp90: *Toxoplasma*, *Entamoeba*, *Giardia*, *Trichomonas*, *Glycine*. In order to maximize the information content of these analyses, partial sequences were also included where no full-length sequence was available. Included partial sequences are as follows: for beta-tubulin: *Entamoeba* (66%); for hsp70: *Paramecium* (35%); and for BiP: *Paramecium* (60%). Trees and 100 bootstrap trees were inferred for all methods described above for the individual gene phylogenies. In addition, this smaller dataset allowed for a computationally manageable protein maximum likelihood analysis using ProML (Felsenstein 1993). Here, trees were searched with global rearrangements and one input order jumble, site-to-site rate variation was modeled using the -r option with 9 categories (8 categories of substitution rates plus invariable sites) with rates and their respective probabilities estimated from the data using PUZZLE (Strimmer and von Haeseler 1996). ProML bootstrapping was not feasible due to computing limitations.

In order to assess whether the choice of genes included in the combined analysis could be biasing the results, analyses were also carried out on six datasets that included all possible combinations of pairs of the four genes, along with analyses of four datasets that included all possible combinations of three out of the four genes. Taxa with non-overlapping data were removed and, with one exception, resulting trees gave alveolate topologies identical to trees inferred from the dataset including all four genes. In the Hsp70+Hsp90 analysis, the dinoflagellate

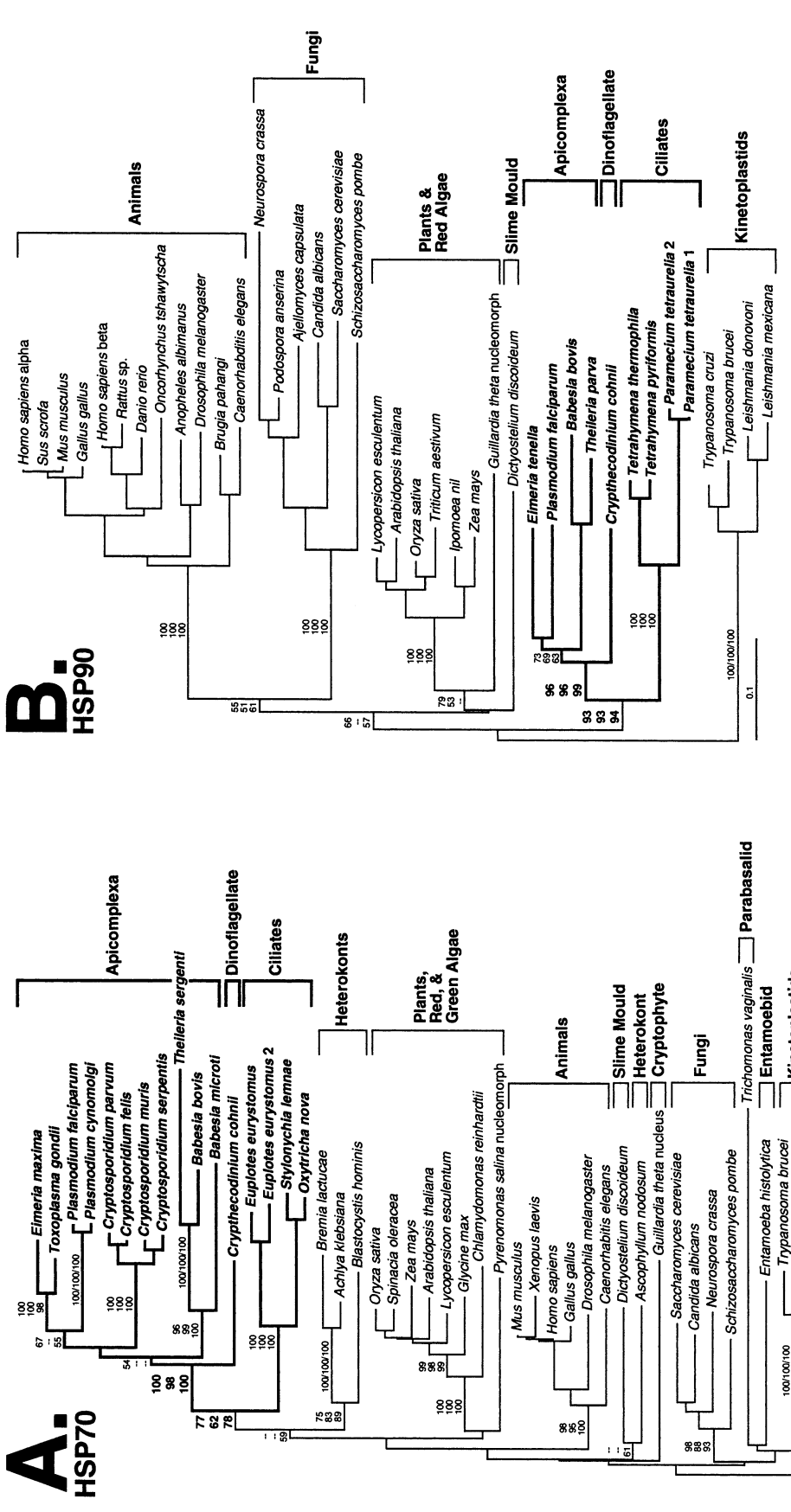

Fig. 1. Individual gene trees for (A) heat shock protein 70 (hsp70), (B) hsp90, (C) the endoplasmic reticulum ortholog of hsp70, or BiP, and (D) beta-tubulin. All trees are Fitch-Margoliash trees constructed from distances corrected for site-to-site rate variation. Numbers at nodes correspond to bootstrap percentages for Fitch-Margoliash (top), weighted neighbor-joining (centre), and BioNJ (bottom). Major eukaryotic lineages are bracketed to the right, with alveolate lineages apicomplexa, dinoflagellates, and ciliates indicated in bold.

Fig. 1. Continued.

sequence was nested within the apicomplexan sequences. In an attempt to maximize the information content, the complete four gene dataset was used for all further analyses.

Kishino-Hasegawa (K-H) tests (Kishino and Hasegawa 1989) to assess the statistical likelihood of alternative tree topologies were carried out on the combined dataset using PUZZLE with rate variation parameters estimated from the data. Test trees were constructed by collapsing all major groups to polytomies, and constraining all possible relationships among apicomplexa, dinoflagellates, and ciliates. Tests were carried out with both pair-wise and multiple comparisons. The relationships within the alveolates were tested using the rate category parameters described above and topologies were rejected if they failed at the 5% level. The level of confidence was based on the K-H test, but was also calculated as a one-tailed test to determine if a topology would likely be rejected in a Shimodaira-Hasegawa test (Goldman, Anderson and Rodrigo 2000; Shimodaira and Hasegawa 1999), and in all cases topologies rejected at the 5% level for two-tailed tests were also rejected as one-tailed tests.

In order to test the effect of the outgroup on the relationships within the alveolates in the combined analysis, 100 datasets were created, where each included the same seven alveolate sequences and a random pair of sequences as the outgroup: 100 pairs of sequences were chosen randomly from the non-alveolate sequences using the Jack2 perl script by W. Fischer. Trees were constructed from these 100 datasets to examine the strength of the intra-alveolate relationships. Protein maximum likelihood trees were constructed with ProML, using uniform rates and global rearrangements, and gamma-corrected distance trees were inferred for each of the 100 datasets as described for bootstrapping.

RESULTS AND DISCUSSION

Characterization of six new genes from *Crypthecodinium*. Complete sequences were obtained for full-length cDNAs from the dinoflagellate *Crypthecodinium cohnii* of actin, beta-tubulin, hsp70, BiP, hsp90, and mitochondrial hsp10. Clones ranged in size from 618 bp (for hsp10) to 2540 bp (for hsp90), with upstream and downstream untranslated regions of ~ 55–80 bp and ~ 200–340 bp, respectively. Mitochondrial hsp10 is the first mitochondrion-targeted gene from a dinoflagellate, and is thus of interest as molecular data from dinoflagellate mitochondria are rare (Inagaki et al. 1997; Norman and Gray 1997). The N-terminal sequence of the *C. cohnii* hsp10 was analyzed with iPSORT (Bannai et al. 2001), and is predicted to have a short mitochondrial transit peptide of nine amino acids, which is similar to the size of other eukaryotic leaders for this protein.

Single gene phylogenies. Phylogenies were first constructed from each gene individually. Unfortunately, hsp10 was found to be too short and too poorly sampled to be phylogenetically useful (e.g. no complete apicomplexan hsp10 sequence is available), and was not analysed further. Actin phylogenies were inferred (not shown), but were not found to differ from published actin trees including other dinoflagellate sequences (Reece et al. 1997; Siddall et al. 1997). Actin was not analyzed further either, since the ciliate sequences are very divergent and consistently branch outside the alveolate clade with the other long-branch, ''deep'' lineages (Keeling 2001). For this reason, actin is not suitable to determine the overall branching pattern within alveolates. In contrast, the alveolates form a monophyletic clade in all phylogenies of hsp70, BiP, hsp90 and beta-tubulin, so these four genes were subjected to further analysis.

Cytosolic hsp70 trees (Fig. 1A) fail to resolve the branching order of major eukaryotic lineages with any strong statistical support. However, the alveolate group is recovered with moderate bootstrap support (62–78%). Within the alveolates, the apicomplexa and dinoflagellates branch together to the exclusion of ciliates with extremely high support (98–100% bootstrap support). However, the apicomplexa only form a weak monophyletic group in the Fitch-Margoliash analysis; in trees where apicomplexa are not monophyletic, the dinoflagellate sequence branches at various positions within the apicomplexa. In both the Fitch-Margoliash and BioNJ analyses, heterokonts are recovered as the next nearest relatives of the alveolates with weak support. In weighted neighbor-joining trees, plants are found to be the sister group to alveolates, but the heterokont sister group is still present in 50% of the bootstrap trees. This finding is in agreement with other evidence pointing to a heterokont-alveolate sister relationship (Van de Peer and De Wachter 1997).

In the phylogeny of hsp90 (Fig. 1B), the overall branching order of major eukaryotic groups is resolved with much more support than hsp70, although the sampling is far less comprehensive. Here, a number of relationships widely accepted based on other data are recovered, including the sisterhood of animals and fungi, the relationship between plants and red algae (represented by the *Guillardia* nucleomorph), and the alveolates. In contrast to hsp70 trees, the alveolate group and the branching order within the alveolates are both recovered with high support: alveolate monophyly is supported by 93% to 94% bootstrap, while apicomplexa and dinoflagellates branch together at the exclusion of ciliates with 95% to 99% bootstrap support, depending on the analysis.

In BiP trees (Fig. 1C) the well-established groups of animals, plants, and fungi are well supported. However, the branching order between these groups and with other eukaryotes is not well-resolved. For example, although animals and fungi are now generally accepted to be related, they only branch together in the Fitch-Margoliash analysis, and this group also includes the long-branched *Trichomonas* sequence. In addition, the monophyly of alveolates is only weakly supported and, although heterokonts increasingly appear to be the next nearest relatives of alveolates, this group consistently branches with plants. In contrast, the relationships within the alveolates are well-supported, with apicomplexa and dinoflagellates branching together to the exclusion of the ciliate with 78% to 88% bootstrap support.

Lastly, beta-tubulin phylogeny (Fig. 1D) was restricted to a subset of eukaryotes due to the comparatively large sequence sampling for this gene. The general characteristics of beta-tubulin phylogeny are well known (Keeling et al. 1999), allowing us to focus on the clade that consistently contains alveolates, and excluding unrelated clades such as animals, fungi, and a number of the more divergent protist lineages (e.g. diplomonads and parabasalia). As expected, major well-defined lineages are resolved in our analyses, as are other proposed relationships, such as those between cercomonads and chlorarachniophytes, and Heterolobosea and Euglenozoa (Keeling, Deane, and McFadden 1998). Plants and green algae also form a strongly supported group that includes *Helicosporidium*, a parasite that has recently been identified as a green alga (Tartar et al. 2001). The alveolates are recovered consistently, but without any support, and are sister to heterokonts once more, but again with very weak support. There is also very little support for the relationships within the alveolates. The ciliates are consistently paraphyletic and the dinoflagellate sequence does not branch with the apicomplexa. The lack of support for branching patterns within the alveolates suggests that beta-tubulin lacks sufficient information to adequately resolve the relationships among alveolates, and it is doubtful that this result reveals a real relationship between dinoflagellates and ciliates.

Taken together, the single gene analyses tend to confirm a

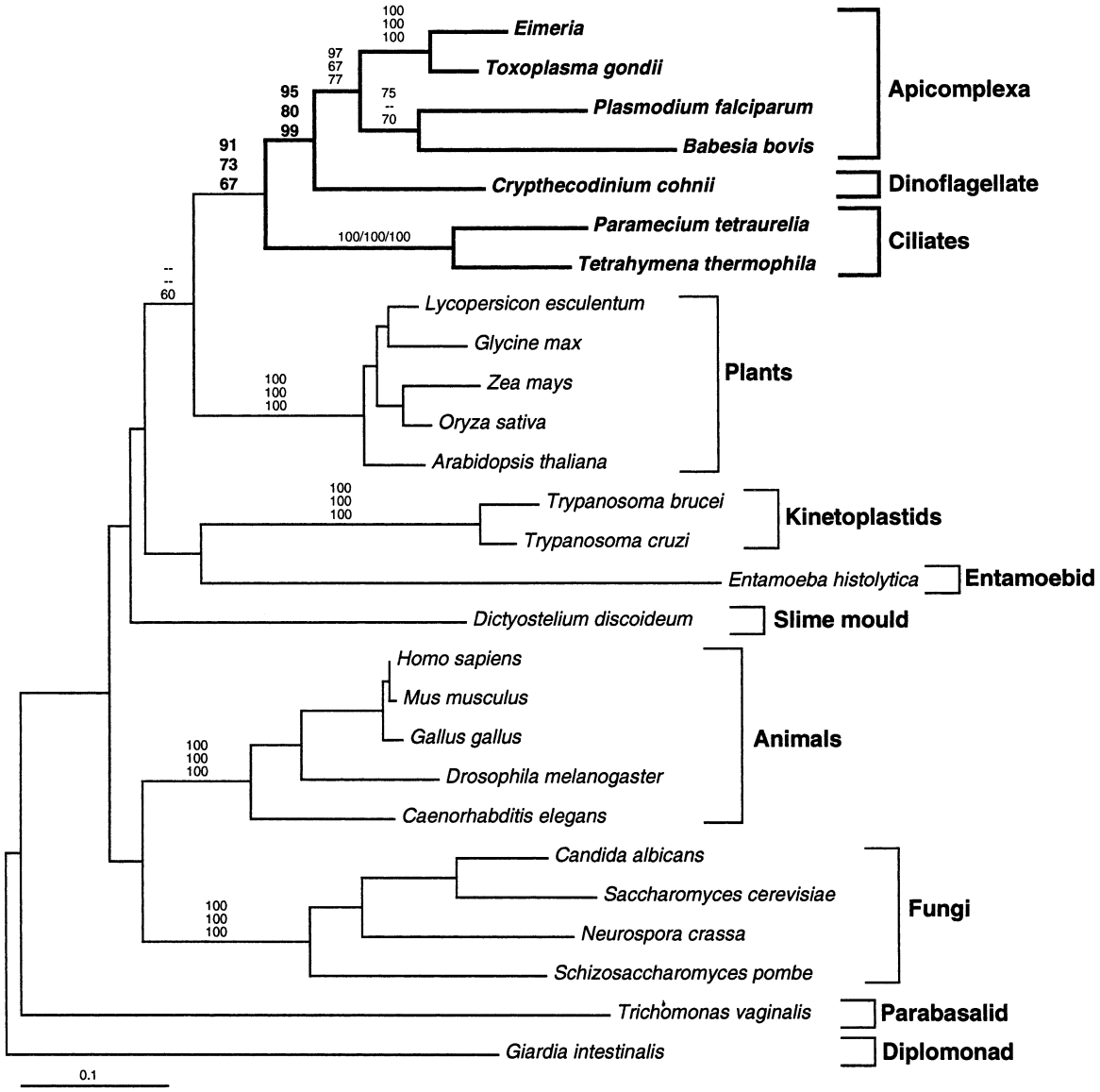

Fig. 2. Protein maximum likelihood phylogeny of combined data from heat shock protein 70 (hsp70), hsp90, the endoplasmic reticulum ortholog of hsp70, or BiP, and beta-tubulin. Numbers at nodes correspond to bootstrap percentages for Fitch-Margoliash (top), weighted neighbor-joining (centre), and BioNJ (bottom). Major eukaryotic lineages are bracketed to the right, with alveolate lineages apicomplexa, dinoflagellates, and ciliates indicated in bold.

closer relationship between apicomplexa and dinoflagellates, but the data are not uniformly strong or entirely consistent. In hsp90, BiP, and hsp70 the sisterhood of apicomplexa and dinoflagellates is supported, while beta-tubulin trees show the dinoflagellate sequence weakly affiliated with a subgroup of ciliate sequences.

Combined phylogenetic analysis. Previously, combined protein phylogenies could not include dinoflagellates, as actin and GAPDH are the only widely sampled, non-plastid proteins characterized from dinoflagellates and used for phylogeny (Fagan, Hastings, and Morse 1998; Fast et al. 2001; Reece et al. 1997; Siddall et al. 1997). In addition, the only recent analysis of global eukaryotic phylogeny using multiple protein-coding genes was based on four genes (Baldauf et al. 2000), of which only beta-tubulin is shared with the set of genes analyzed here. In an effort to increase the amount of information available from the single gene phylogenies, and to compare the overall relationships between the eukaryotic groups with a different combined set of protein-coding genes, a phylogenetic analysis based on the concatenated sequences of BiP, hsp70, hsp90, and beta-tubulin was carried out. Unfortunately, heterokonts are not evenly sampled and could not be included in the analysis, so we could gather no further information about their relationship to alveolates.

The Protein maximum likelihood (Fig. 2), Fitch-Margoliash, Weighted neighbor-joining and BioNJ trees all share the same basic topology, with some differences in the weakly supported backbone of the tree. In all analyses the alveolates are consistently monophyletic and are most closely related to plants. Although bootstrapping of the corrected protein maximum likelihood tree was too computationally intensive to be undertaken, 100 bootstrap replicates were inferred for the distance trees (Fig. 2). As a group, alveolates are moderately well-supported with 67% to 91% bootstrap support, while the topology within

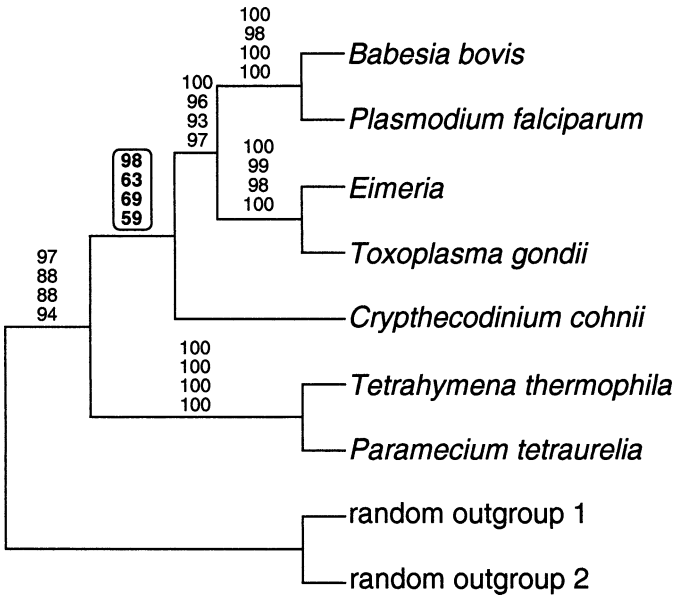

Fig. **3**. Outgroup effect on combined data of four proteins from alveolates. The topology of seven alveolate ''sequences'' was reconstructed with 100 random pairs of outgroup sequences. Numbers at nodes correspond to the percent occurrence of that branch in the 100 datasets using protein maximum likelihood (top), Fitch-Margoliash (second), weighted neighbor-joining (third), and BioNJ (bottom).

the alveolates is extremely well supported, with 80% to 99% bootstrap support for placing the dinoflagellate and apicomplexan sequences together to the exclusion of ciliates.

The intra-alveolate relationships were also assessed with Kishino-Hasegawa tests by collapsing all major groups to polytomies (including the relationships within apicomplexa) and testing the three possible branching orders within alveolates: apicomplexa-dinoflagellate sisterhood, apicomplexa-ciliate sisterhood, and ciliate-dinoflagellate sisterhood. These tests were conducted between all possible pairs of trees and also comparing all combinations simultaneously. The best tree consistently included apicomplexa and dinoflagellates as sisters, the ciliate-dinoflagellate sisterhood was not significantly worse, and the apicomplexa-ciliate sisterhood was consistently rejected at the 5% level in all tests.

Lastly, the effect of the outgroup was examined by creating a sample of random outgroups. In all cases, the majority of the trees recovered an apicomplexa-dinoflagellate clade (Fig. 3). However, the frequency of trees supporting this relationship varied tremendously, ranging from only 59% in the BioNJ analysis to 98% using Protein maximum likelihood.

Concluding remarks. With four new protein sequences from *Crypthecodinium*, we were able to more substantially address the issue of alveolate relationships by constructing single gene phylogenies, combined gene phylogenies, and carrying out statistical tests of tree topologies. With varying degrees of support, these analyses back up the conclusion that dinoflagellates and apicomplexa are sisters, but the results are not universally clear-cut since K-H tests failed to reject the sisterhood of dinoflagellates and ciliates in combined data analysis, and the support for the dinoflagellate-apicomplexa grouping is variable in the single gene trees and analysis of outgroup effect.

ACKNOWLEDGMENTS

This work was supported by grants from the Canadian Institutes for Health Research to P.J.K. (MOP-42517) and from Project Ingenhouz at Arizona State University to S.B.. N.M.F. was supported by a fellowship from the Natural Sciences and Engineering Research Council of Canada. P.J.K. is a Scholar of the Canadian Institute for Advanced Research. We also thank Kirk Apt and Casey Lippmeier at Martek Bioscience Corp. for providing the *Crypthecodinium* library and W. Fischer for the use of Jack2.

LITERATURE CITED

Apt, K. E., Clendennen, S. K., Powers, D. A. & Grossman, A. R. 1995. The gene family encoding the fucoxanthin chlorophyll proteins from the brown alga *Macrocystis pyrifera*. *Mol. Gen. Genet.,* **246**:455–464.

Baldauf, S. L., Roger, A. J., Wenk-Siefert, I. & Doolittle, W. F. 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science,* **290**:972–977.

Bannai, H., Tamada, Y., Maruyama, O., Nakai, K. & Miyano, S. 2001. Views: Fundamental Building Blocks in the Process of Knowledge Discovery, *In:* Proceedings of the 14th International FLAIRS Conference.

Bruno, W. J., Socci, N. D. & Halpern, A. L. 2000. Weighted neighbor joining: a likelihood-based approach to distance-based phylogeny reconstruction. *Mol. Biol. Evol.,* **17**:189–197.

Fagan, T., Hastings, J. W. & Morse, D. 1998. The phylogeny of glyceraldehyde-3-phosphate dehydrogenase indicates lateral gene transfer from cryptomonads to dinoflagellates. *J. Mol. Evol.,* **47**:633–639.

Fast, N. M., Kissinger, J. C., Roos, D. S. & Keeling, P. J. 2001. Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. *Mol. Biol. Evol.,* **18**:418–426.

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package). J. Felsenstein, University of Washington, Seattle.

Gagnon, S., Bourbeau, D. & Levesque, R. C. 1996. Secondary structure and features of the 18S, 5.8S and 26S ribosomal RNAs from the Apicomplexan parasite *Toxoplasma gondii*. *Gene,* **173**:129–135.

Gajadhar, A. A., Marquardt, W. C., Hall, R., Gunderson, J., Ariztia-Carmona, E. V. & Sogin, M. L. 1991. Ribosomal RNA sequences of *Sarcocystis muris, Theileria annulata* and *Crypthecodinium cohnii* reveal evolutionary relationships among apicomplexans, dinoflagellates and ciliates. *Mol. Biochem. Parasitol.,* **45**:147–154.

Gascuel, O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol. Biol. Evol.,* **14**:685–695.

Goldman, N., Anderson, J. P. & Rodrigo, A. G. 2000. Likelihood-based tests of topologies in phylogenetics. *Syst. Biol.,* **49**:652–670.

Inagaki, Y., Hayashi-Ishimaru, Y., Ehara, M., Igarashi, I. & Ohama, T. 1997. Algae or protozoa: phylogenetic position of euglenophytes and dinoflagellates as inferred from mitochondrial sequences. *J. Mol. Evol.,* **45**:295–300.

Keeling, P. J. 2001. Foraminifera and Cercozoa are related in actin phylogeny: two orphans find a home? *Mol. Biol. Evol.,* **18**:1551–1557.

Keeling, P. J., Deane, J. A. & McFadden, G. I. 1998. The phylogenetic position of alpha- and beta-tubulins from the *Chlorarachnion* host and *Cercomonas* (Cercozoa). *J. Eukaryot. Microbiol.,* **45**:561–570.

Keeling, P. J., Deane, J. A., Hink-Schauer, C., Douglas, S. E., Maier, U. G. & McFadden, G. I. 1999. The secondary endosymbiont of the cryptomonad *Guillardia theta* contains alpha-, beta-, and gamma-tubulin genes. *Mol. Biol. Evol.,* **16**:1308–1313.

Kishino, H. & Hasegawa, M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *J. Mol. Evol.,* **29**:170–179.

Köhler, S., Delwiche, C. F., Denny, P. W., Tilney, L. G., Webster, P., Wilson, R. J., Palmer, J. D. & Roos, D. S. 1997. A plastid of probable green algal origin in Apicomplexan parasites. *Science,* **275**:1485–1489.

Lee, J. J., Hutner, S. H. & Bovee, E. C. 1985. Illustrated Guide to the Protozoa. Society of Protozoologists, Lawrence, Kansas. p. 393–575.

Levine, N. D. 1988. The Protozoan Phylum Apicomplexa. CRC Press, Boca Raton, Fl.

McFadden, G. I., Reith, M. E., Munholland, J. & Lang-Unnasch, N. 1996. Plastids in human parasites. *Nature,* **381**:482.

Norman, J. E. & Gray, M. W. 1997. The cytochrome oxidase subunit 1 gene (*cox*1) from the dinoflagellate, *Crypthecodinium cohnii*. *FEBS Lett.* **413**:333–338.

Philippe, H. & Laurent, J. 1998. How good are deep phylogenetic trees? *Curr. Opin. Genet. Dev.,* **8**:616–623.

Reece, K. S., Siddall, M. E., Burreson, E. M. & Graves, J. E. 1997. Phylogenetic analysis of *Perkinsus* based on actin gene sequences. *J. Parasitol.,* **83**:417–423.

Saldarriaga, J. F., Taylor, F. J. R., Keeling, P. J. & Cavalier-Smith, T. 2001. Dinoflagellate nuclear SSU rRNA phylogeny suggests multiple plastid losses and replacements. *J. Mol. Evol.,* **53**:204–213.

Shimodaira, H. & Hasegawa, M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.,* **16**:1114–1116.

Siddall, M. E., Reece, K. S., Graves, J. E. & Burreson, E. M. 1997. 'Total Evidence' refutes the inclusion of *Perkinsus* species in the phylum Apicomplexa. *Parasitology,* **115**:165–176.

Smith, R. F. & Smith, T. F. 1992. Pattern-induced multi-sequence alignment (PIMA) algorithm employing secondary structure-dependent gap penalties for use in comparative protein modelling. *Protein Eng.,* **5**:35–41.

Strimmer, K. & Von Haeseler, A. 1996. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.,* **13**:964–969.

Tartar, A., Boucias, D. G., Adams, B. J. & Becnel, J. J. 2002. Phylogenetic analysis identifies the invertebrate pathogen *Helicosporidium* sp. as a green algae (Chlorophyta). *Int. J. Syst. Evol. Microbiol.,* **52**: 273–279.

Taylor, F. J. R. 1976. Flagellate phylogeny: a study in conflicts. *J. Protozool.,* **23**:28–40.

Taylor, F. J. R. 1987. The biology of dinoflagellates. Botanical Monographs. Vol. 21. Blackwell Scientific Publications, Oxford.

Taylor, F. J. R. 1999. Morphology (tabulation) and molecular evidence for dinoflagellate phylogeny reinforce each other. *J. Phycol.,* **35**:1–6.

Van de Peer, Y. & De Wachter, R. 1997. Evolutionary relationships among the eukaryotic crown taxa taking into account site-to-site rate variation in 18S rRNA. *J. Mol. Evol.,* **45**:619–630.

Van de Peer, Y., Van der Auwera, G. & De Wachter, R. 1996. The evolution of stramenopiles and alveolates as derived by ''substitution rate calibration'' of small ribosomal subunit RNA. *J. Mol. Evol.,* **42**: 201–210.

Van de Peer, Y., Baldauf, S. L., Doolittle, W. F. & Meyer, A. 2000. An updated and comprehensive rRNA phylogeny of (Crown) eukaryotes based on rate-calibrated evolutionary distances. *J. Mol. Evol.,* **51**:565–576.

Waller, R. F., Keeling, P. J., Donald, R. G., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A. F., Besra, G. S., Roos, D. S. & McFadden, G. I. 1998. Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA,* **95**:12352–12357.

Wilson, R. J., Denny, P. W., Preiser, P. R., Rangachari, K., Roberts, K., Roy, A., Whyte, A., Strath, M., Moore, D. J., Moore, P. W. & Williamson, D. H. 1996. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.,* **261**: 155–172.

Wolters, J. 1991. The troublesome parasites—molecular and morphological evidence that Apicomplexa belong to the dinoflagellate-ciliate clade. *Biosystems,* **25**:75–83.

Zhang, Z., Green, B. R. & Cavalier-Smith, T. 1999. Single gene circles in dinoflagellate chloroplast genomes. *Nature,* **400**:155–159.

Zhang, Z., Green, B. R. & Cavalier-Smith, T. 2000. Phylogeny of ultrarapidly evolving dinoflagellate chloroplast genes: a possible common origin for sporozoan and dinoflagellate plastids. *J. Mol. Evol.,* **51**:26–40.

Received 07/20/01, 09/24/01, 10/25/01; accepted 10/25/01

WEBSITES

Society of Protozoologists Home Page

http://www.uga.edu/protozoa

The Journal of Eukaryotic Microbiology Home Page

Abstracts for Section Meetings, 2001

http://www.jeukmic.org/