Inference of the Phylogenetic Position of Oxymonads Based on Nine Genes: Support for Metamonada and Excavata

Vladimír Hampl, David S. Horner,† Patricia Dyal,‡ Jaroslav Kulda,* Jaroslav Flegr,* Peter G. Foster,‡[1] and T. Martin Embley‡[1]*

*Department of Parasitology, Faculty of Science, Charles University in Prague, Prague, Czech Republic; †Department of Molecular Biosciences and Biotechnology, University of Milan, Milan, Italy; and ‡Department of Zoology, Natural History Museum, London, United Kingdom

Circumscribing major eukaryote groups and resolving higher order relationships between them are among the most challenging tasks facing molecular evolutionists. Recently, evidence suggesting a new supergroup (the Excavata) comprising a wide array of flagellates has been collected. This group consists of diplomonads, retortamonads, *Carpediemonas*, heteroloboseans, *Trimastix*, jakobids, and *Malawimonas*, all of which possess a particular type of ventral feeding groove that is proposed to be homologous. Euglenozoans, parabasalids, and oxymonads have also been associated with Excavata as their relationships to one or more core excavate taxa were demonstrated. However, the main barrier to the general acceptance of Excavata is that its existence is founded primarily on cytoskeletal similarities, without consistent support from molecular phylogenetics. In gene trees, Excavata are typically not recovered together. In this paper, we present an analysis of the phylogenetic position of oxymonads (genus *Monocercomonoides*) based on concatenation of eight protein sequences (α-tubulin, β-tubulin, γ-tubulin, EF-1α, EF-2, cytosolic (cyt) HSP70, HSP90, and ubiquitin) and 18S rRNA. We demonstrate that the genes are in conflict regarding the position of oxymonads. Concatenation of α- and β-tubulin placed oxymonads in the plant-chromist part of the tree, while the concatenation of other genes recovered a well-supported group of Metamonada (oxymonads, diplomonads, and parabasalids) that branched weakly with euglenozoans—connecting all four excavates included in the analyses and thus providing conditional support for the existence of Excavata.

Introduction

Reconstructing a comprehensive tree of all living organisms is the ultimate goal of molecular phylogenetics. Work over the last few decades has often resolved the branching pattern within major groups of eukaryotes, but the relationships between them remain largely unresolved. The shape of the tree suggested by small-subunit rRNA sequences—comprising a crown and several early offshoots (Sogin and Silberman 1998)—has been questioned following the demonstration that the position of many of the early branches (namely, Microsporidia, Parabasalia, and Diplomonadida) could be the result of the long-branch attraction artifact which can push fast-evolving eukaryotic sequences toward long-branch prokaryotic outgroups (Hirt et al. 1999; Stiller and Hall 1999; Philippe 2000; Dacks et al. 2002). The questioning of the existence of early branching taxa came hand in hand with the rejection of the archezoa hypothesis for the origin of eukaryotes. This hypothesis posited that the first eukaryotic lineages in the 18S rRNA tree, which do not harbor well-defined mitochondria, were descendants of primitively amitochondriate eukaryotes that diverged from the main trunk of eukaryote evolution before the acquisition of this organelle (Cavalier-Smith 1983). This view was abandoned after the demonstration that the genomes of the best-studied archezoans (microsporidians, trichomonads, diplomonads, *Entamoeba*) contain genes which cluster with mitochondrial orthologues (Clark and Roger 1995; Germot, Philippe, and Le Guyader 1996; Horner et al. 1996; Roger et al. 1998; Tachezy, Sánchez, and Müller 2001) and that key species contain highly derived organelles of mitochondrial origin (Mai et al. 1999; Tovar, Fisher, and Clark 1999; Williams et al. 2002; Tovar et al. 2003; Hrdy et al. 2004).

Although the early divergence of diplomonads and parabasalids is doubtful, available molecular data do not offer robustly supported alternative placements. The morphologically based "excavate hypothesis" of Simpson and Patterson (1999) postulates a relationship between the diplomonads, retortamonads, *Carpediemonas*, heteroloboseans, *Trimastix*, jakobids (*Jakoba*, *Reclinomonas*), and *Malawimonas*—the so-called "excavate taxa". All "excavate taxa" possess a ventral feeding groove through which pass one or more recurrent flagella. The hypothesis that these structures are homologous implies that the "excavate taxa" constitute either a clade or an evolutionary grade, for review see Simpson (2003). Molecular evidence has also suggested the affiliation of some "excavate taxa" to flagellates without a similar groove. There are indications that Parabasalia (which lack most if not all excavate characters) share a common ancestor with Diplomonadida. For example, *Trichomonas* and *Giardia* appear to have obtained genes for glucokinase, glucosephosphate isomerase, and two tRNA synthetase genes via horizontal gene transfer from a common source (Henze et al. 2001; Andersson, Sarchfield, and Roger 2004), and they are nearest neighbors in a tree constructed from concatenated genes (Baldauf et al. 2000; Arisue, Hasegawa, and Hashimoto 2005). The same concatenated protein data, some 18S rRNA analyses (Silberman et al. 2002; Cavalier-Smith 2003*a*; Nikolaev et al. 2004), and the sharing of unique discoid shape of mitochondrial cristae argue for the existence of Discicristata (Cavalier-Smith 1998)—comprising the excavate taxon Heterolobosea and nonexcavate Euglenozoa. Finally, phylogenetic tree based upon 18S rRNA sequences has recovered a robust relationship between the excavate taxon *Trimastix* and Oxymonadida (Dacks et al. 2001).

[1] Present address: School of Biology, University of Newcastle upon Tyne, Newcastle upon Tyne, United Kingdom.

Key words: phylogeny, concatenation, Excavata, Metamonada, oxymonads, *Monocercomonoides*.

E-mail: vlada@natur.cuni.cz.

Mol. Biol. Evol. 22(12):2508–2518. 2005
doi:10.1093/molbev/msi245
Advance Access publication August 24, 2005

Oxymonadida are a relatively little studied group of anaerobic amitochondrial flagellates that live as intestinal endosymbionts of insects and exceptionally of vertebrates (several species of the genus *Monocercomonoides*) (Brugerolle and Lee 2000). The ultrastructure of all five families of oxymonads is relatively well documented (Brugerolle and Lee 2000). A reinterpretation of the mastigont and associated cytoskeleton of *Monocercomonides hausmani*, conducted in the light of the oxymonad-*Trimastix* relationship, revealed similarities to "excavate taxa" including a rudimentary ventral groove (Simpson et al. 2002*a*). Although the existence of an oxymonad-*Trimastix* clade seems to be well corroborated, its precise position in the eukaryotic tree is unstable in published gene phylogenies (Moriya, Ohkuma, and Kudo 1998; Dacks and Roger 1999; Dacks et al. 2001; Moriya et al. 2001; Keeling and Leander 2003).

Cavalier-Smith (2002) combined morphological and molecular evidence for "excavate taxa" and their relatives and created the infrakingdom Excavata composed of the "excavate taxa," Parabasalia, Euglenozoa, and Oxymonadida. The concept of Excavata is currently one of the most disputed hypotheses concerning the early phylogeny of eukaryotes. With the exception of weak support for the monophyly of Excavata in two analyses of 18S rRNA (Cavalier-Smith 2003*a*; Nikolaev et al. 2004), the results of all other molecular studies favor their polyphyly (e.g., Dacks et al. 2001; Edgcomb et al. 2001; Simpson et al. 2002*b*; Arisue, Hasegawa, and Hashimoto 2005). Although the monophyly of the entire Excavata is uncertain, several more or less well-supported clusters of excavates can be recognized. Besides the aforementioned Discicristata, *Trimastix*-Oxymonadida, and Parabasalia-Diplomonadida groups, it has been shown in 18S rRNA and in α- and β-tubulin phylogenies that Retortamonadida (Silberman et al. 2002) and *Carpediemonas* (Simpson et al. 2002*b*) are close relatives of Diplomonadida and that *Jakoba libera* branches with *Reclinomonas americana* (Simpson et al. 2002*b*) Concatenation of α- and β-tubulin, HSP90, and cytosolic (cyt) HSP70 supported the relationship of *Reclinomonas* and Discicristata (Simpson and Roger 2004).

The Excavata is an important group for understanding eukaryotic diversity, ecology, and evolution. It comprises flagellates with aerobic, anaerobic, parasitic, and free-living lifestyles. In some Excavata, no mitochondria-derived organelles have yet been identified, but in general the group harbors a functionally and morphologically diverse range of mitochondria and mitochondria-derived organelles, including fully functional aerobic mitochondria, the mitochondria of kinetoplastids with their specialized genome organization (kinetoplast), poorly characterized double membrane-bounded organelles in *Trimastix* and *Carpediemonas*, hydrogenosomes in parabasalids, and a tiny mitochondrial homologue or "mitosome" in *Giardia*. The mitochondrial genomes of *R. americana* and *J. libera* contain the largest number of mitochondrial genes of any sequenced mitochondrial genome (Lang, Gray, and Burger 1999). The Excavata thus have the potential to play a crucial role in understanding the evolution of eukaryotes and the mitochondrial organelle. Whether all Excavata are monophyletic and locating their position(s) in the eukaryotic tree are therefore important questions.

The increase in the amount of sequence data in public databases enables the use of more than one gene for phylogenetic inference (Baldauf et al. 2000; Bapteste et al. 2002; Arisue, Hasegawa, and Hashimoto 2005). This approach increases the amount of characters, often a critical factor for the robust inference of tree topology (Bapteste et al. 2002). In order to establish the affiliation of the oxymonads to other taxa, we have sequenced eight protein-coding gene sequences and the 18S rRNA gene from the oxymonad genus *Monocercomonoides* and performed phylogenetic analyses using an alignment of concatenated genes (4,347 characters). We demonstrated that the genes contain conflicting phylogenetic signals. The tests of phylogenetic hypotheses showed that two genes (α-, β-tubulin) place oxymonads in the plant-chromist part of the tree, six genes (γ-tubulin, EF-1α, EF-2, cytHSP70, ubiquitin, and 18S rRNA) place oxymonads together with diplomonads and parabasalids, and one gene (HSP90) did not differentiate between these two hypotheses.

Materials and Methods

Organism

Monocercomonoides strain PA203 was isolated from a *Chinchilla laniger* by J. Kulda in 1993. The agnotobiotic culture with bacteria, but no other eukaryote, was maintained by serial transfer every 2–3 days in modified Diamond TYSGM medium in 37°C with the addition of rice starch.

Filtration

The culture was filtered to remove the majority of bacteria before the isolation of DNA or RNA. In the first step, large contaminants were removed by filtration through filter paper. The filtrate, which contained the protozoa, was then filtered through a 3-μm cutoff polycarbonate filter (Whatman International Ltd., Maidstone, UK). Most bacteria appeared in the flow-through, while the trophozoites of *Monocercomonoides* remained in the medium above the filter. The suspension of trophozoites was washed by continuous addition of pure medium (approximately 3 volumes of the original culture) and finally transferred to a clean tube. Both filtrations were accelerated by application of a partial vacuum.

Isolation of Genomic DNA and Amplification of 18S rRNA, α-tubulin, and EF-1α Genes

Cells were lysed using 4 M guanidium isothiocyanate, proteins were extracted once with phenol-chloroform (1:1) and twice with chloroform, and the DNA was precipitated with ethanol.

An 800-kb fragment of 18S rRNA gene was amplified from genomic DNA using eukaryotic primers 5′F EUK (5′-AYCTGGTTGATYYTGCCAG-3′) (Medlin et al. 1988) and 577R (5′-ACCGCGGCKGCTGGC-3′). The polymerase chain reaction (PCR) mixture consisted of 75 mM Tris-HCl pH 8.8, 20 mM NH_4SO_4, 0.01% Tween 20, 2 mM $MgCl_2$, 0.25 mM deoxynucleoside triphosphate, 0.4 μM

of each primer, 2 ng/μl bovine serum albumin, and 0.04 U/μl *Taq* polymerase. The temperature profile consisted of initial denaturation at 94°C for 2 min followed by 35 cycles of 94°C for 1 min, 60°C for 2 min, and 72°C for 3 min with a final polymerization at 72°C for 15 min. The PCR products were purified from the gel and sequenced. A specific primer MspecF1 (5′-AAGGGGTGTAATGATATGTGTCC-3′) was designed on the basis of the determined sequence. The remaining part of the 18S rRNA gene (approximately 2,800 bp) was amplified using similar reaction conditions (but annealing at 50°C) using one specific primer MspecF1 and the universal eukaryotic primer 3′R EUK (5′-TGATCCATCTGCAGGTTCACCT-3′) (Medlin et al. 1988). The PCR products were cloned into the pCR2.1-TOPO vector using the TOPO TA Cloning kit, and both strands were sequenced using a gene-walking strategy. Specific primers for the whole 18S rRNA gene were designed on the basis of the complete sequence (MonidesSSUF—5′-GAAGTGATATGCTGTCTCAA-3′ and MonidesSSUR—5′-TCACCTACGGAAACCTT-3′). Using these primers, the complete 18S rRNA gene was amplified, cloned, and sequenced. The amplification conditions differed from the previously described in the annealing temperature (50°C) and the length of polymerization step in the cycle (2 min). Regions that differed in their sequence among subclones were sequenced directly from PCR products.

The α-tubulin gene was amplified from genomic DNA using nested PCR. The primary PCR was carried out with primers AtubA (5′-RGTNGGNAAYGCNTGYTGGGA-3′) and AtubB (5′-CCATNCCYTCNCCNACRTACCA-3′) (Edgcomb et al. 2001). The reaction mix was identical to the 18S rRNA PCR, excepting that primers were at 1.2 μM. The cycle consisted of an initial denaturation (94°C for 2 min), 35 cycles (94°C for 1 min, 45°C for 1.5 min, and 72°C for 2 min), and final polymerization (72°C for 15 min). The products of the primary PCR were purified using Microcon YM-100 and used as the template for the secondary PCR with primers AlphatubF1 (5′-TAYTGYYWNGARCAYGGNAT-3′) and AlphatubR1 (5′-ACRAANGCNCGYTTNGMRWACAT-3′) (Moriya et al. 2001). Secondary PCR was performed under the same conditions as the primary PCR, only the number of cycles was reduced to 25. The PCR products were cloned into the pGEM-T Easy Vector. Both strands of five subclones from two independent nested PCR were sequenced using the gene-walking strategy.

The EF-1α gene was amplified from genomic DNA using nested PCR. The procedure was similar to the amplification of α-tubulin. The combination of primers was EF-1alpha55F (5′-GTNATHGGNCAYGTNGAYDSNGGNAA-3′) (Moriya, Ohkuma, and Kudo 1998) and EF-1alphaR2 (5′-CKCATRTCNCGNACNGCRAANCKNCC-3′) for primary PCR and EF-1alpha110F (3′-CGNGARMGNGGNATNACNAT-5′) (Moriya, Ohkuma, and Kudo 1998) and EF-1alpha345R (TCNCGNACNGCRAAYCKYCCNAGNGGNG) (Moriya, Ohkuma, and Kudo 1998) for secondary PCR. The reaction conditions were the same as used for α-tubulin. Both strands of four subclones from two independent nested PCRs were sequenced with a gene-walking strategy. All sequencing was performed on an ABI PRISM 310.

Preparation of a cDNA Library

Total RNA was isolated from 250×10^6 filtered trophozoites using the single-step guanidine isolation method (Ausubel et al. 2001). mRNA was isolated from total RNA using the Dynabeads mRNA purification kit (DYNAL, Wirral, UK). A cDNA library was constructed from 5 μg of mRNA using the λ ZAP-cDNA Gigapack III Gold Cloning Kit according to the manufacturer's instructions (# 080012, Stratagene, La Jolla, Calif.).

Expressed Sequence Tag Sequencing

Recombinant pBluescript SK (−) plasmids were excised from the λ ZAP II cDNA library according to the manufacturer's instructions (Stratagene). Plasmid DNA was extracted from 400 randomly selected clones using the Montage Plasmid Miniprep$_{96}$ Kit (Millipore, Bedford, Mass.) and sequenced with the vector-based M13F primer. A BlastX search was performed on the cDNA sequences using the National Center for Biotechnology Information database. Clones that contained the sequences of genes suitable for molecular phylogeny (β-tubulin, γ-tubulin, cytHSP70, HSP90, EF-2, and ubiquitin) were selected, and both strands of one clone for each gene were completely sequenced using the gene-walking strategy.

Phylogenetic Analyses

Oxymonad 18S rRNA Phylogeny

The 18S rRNA sequences of oxymonads and *Trimastix* were aligned using ClustalX (Thompson et al. 1997) and manually refined in BioEdit 5.0.9 (Hall 1999). The 18S rRNA tree was constructed from a data set of 1,329 characters representing 24 taxa by maximum likelihood in PAUP*4.0beta (Swofford 1998) using the TrN + I + Γ model recommended by a hierarchical likelihood ratio test implemented in Modeltest 3.06 (Posada and Crandall 1998). Five hundred bootstrap replicates, each with 10 subreplicates of starting tree construction by random taxa addition and tree bisection-reconnection (TBR) swapping, were performed. Bootstrapping using distance (least squares method with LogDet distances and constant sites removed from the alignment) and maximum parsimony methods was also performed in PAUP*4.0beta with 1,000 replicates, each with 10 subreplicates of starting tree construction by random taxa addition and TBR swapping. Bayesian posterior probabilities under the GTR + I + Γ model were obtained using MrBayes 3.0 (Ronquist and Huelsenbeck 2003) after 2,000,000 generations of four parallel MCMCs sampled every 1,000 generations after discarding the first 1,001 as burn-in.

Eukaryotic Phylogeny

Eukaryotic α-tubulin, β-tubulin, γ-tubulin, EF-1α, EF-2, cytHSP70, HSP90, ubiquitin, and 18S rRNA sequences were downloaded from public databases. Preliminary sequence data for HSP90 of *Trichomonas vaginalis* were obtained from The Institute for Genomic Research through the Web site http://www.tigr.org. The set of taxa was compiled such that at least seven gene sequences were present for each taxon. Only sequences that passed the chi-square test of nucleotide or amino acid composition in Tree-Puzzle 5.0

(Strimmer and vonHaeseler 1996) were used. Because of the limited data, some higher taxa were represented by different species for individual genes as follows: Stramenopila (*Pelvetia fastigiata*, *Thalassiosira weissflogii*, *Blastocystis hominis*, *Achlya ambisexualis*, *Phytophthora infestans*), Rhodophyta (*Porphyra purpurea*, *Cyanidioschyzon merolae*, *Guillardia theta* nucleomorph), Diplomonadida (*Giardia intestinalis*, *Hexamita inflata*), *Trichomonas* (*Trichomonas vaginalis*, *Trichomonas tenax*), *Leishmania* (*Leishmania donovani*, *Leishmania major*, *Leishmania tarentolae*), *Trypanosoma* (*Trypanosoma cruzi*, *Trypanosoma brucei*), *Tetrahymena* (*Tetrahymena pyriformis*, *Tetrahymena thermophila*), *Euplotes* (*Euplotes octocarinatus*, *Euplotes eurystomus*), and Apicomplexa (*Plasmodium falciparum*, *Toxoplasma gondii*). The sequences were aligned in BioEdit 5.0.9 at the amino acid level, nucleotide sequences of 18S rRNA gene were aligned in ClustalX. All alignments are available as Supplementary Material online.

The analyses of all concatenates were performed using MrBayes 3.0 (Ronquist and Huelsenbeck 2003) (missing data coded as X). Specific models were assigned to each gene partition (covarion JTT + I + Γ for proteins and covarion GTR + I + Γ for nucleotides). Proportions of invariant sites and the shape parameter for the gamma distribution were optimized for each individual partition. Four MCMC chains were run for 2,000,000 generations and sampled every 1000th generation; the trees from the generations before the chain reached the plateau of likelihood value were discarded as burn-in.

Bootstrapping was performed using the concatenated data set after the exclusion of 18S rRNA sequences because the software cannot analyze mixed amino acid/nucleotide data. A total of 1,000 distance bootstrap replicates were performed using the Seqboot, Protdist (JTT + Γ substitution model), Fitch (10 replicates of random taxa addition), and Consense programs of the PHYLIP 3.6 package (Felsenstein 1993). The 1,000 maximum parsimony bootstrap replicates were performed using Seqboot, Protpars (10 replicates of random taxa addition), and Consense (majority rule extended) programs of the PHYLIP package (Felsenstein 1993). The 100 maximum likelihood bootstraps were performed using Seqboot, Phyml (Guindon and Gascuel 2003) (JTT substitution model, eight gamma rate categories with estimated α parameter, estimated proportion of invariant sites), and Consense.

Likelihood Mapping

Likelihood mapping (Strimmer and vonHaeseler 1997) was performed in Tree-Puzzle 5.0 (Strimmer and vonHaeseler 1996) using JTT + I + Γ model for proteins and TrN + I + Γ model for 18S rRNA with parameters estimated by the software. The taxa were divided into four groups: (1) *Monocercomonoides*; (2) diplomonads, *Trichomonas*; (3) *Saccharomyces*, *Neurospora*, *Caenorhabditis*, *Homo*; and (4) all remaining taxa.

Approximately Unbiased Test

Phylogenetic hypotheses were tested using the approximately unbiased (AU) test (Shimodaira 2002) implemented in the program Consel (Shimodaira and Hasegawa 2001). For each gene, the site likelihoods for the particular set of topologies were calculated under the appropriate model using P4 (Foster 2004). These site likelihoods were used as input data for Consel.

When comparing the results of different gene concatenates, the input file for Consel was prepared by concatenation of site likelihoods calculated separately for two or three partitions of the concatenate, which enabled the use of more specific models. The partitions were proteins and 18S for the two larger data sets (fig. 4*c* and *d*) and single genes for the smaller data sets (fig. 4*a* and *b*).

Results

Sequenced Genes

18S rRNA, α-tubulin, and EF-1α genes were sequenced from genomic DNA clones and β-tubulin, γ-tubulin, EF-2, cytHSP70, HSP90, and ubiquitin were sequenced from cDNA library clones.

The sequenced part of the 18S rRNA gene of *Monocercomonoides* (2,927 nt long) is among the longest known. The exceptional length results from expansions of several regions of the sequence (Supplementary Table, Supplementary Material online). While some expansions are *Monocercomonoides* specific, others appeared in the same parts of the molecule in other oxymonad genera, but were substantially longer in *Monocercomonoides*. At 11 positions, the nucleotide sequence of the *Monocercomonoides* 18S rRNA gene differed among clones. Corresponding ambiguous signal was also observed in direct sequencing of the PCR products. All these polymorphisms were located in the expanded regions and probably result from the presence of multiple copies of the 18S rRNA gene in the genome of *Monocercomonoides*.

Considerable polymorphism was also detected among the five sequenced subclones of the α-tubulin gene. The closest subclones differed in 28 nt (all synonymous), and the most distant subclones differed in 78 nt (10 nonsynonymous changes). This suggests that most of the differences are real rather than *Taq* polymerase errors in the course of PCR amplification. In phylogenetic analyses (not shown), the subclones formed a single cluster with 100% bootstrap support and very short internal branches. The sequence of subclone 1/1 was used in all further phylogenetic analyses.

The sequences of four EF-1α subclones differed in 3–57 nt and 2–8 amino acids, and formed a single cluster in the preliminary analyses. The sequence of subclone 2/1 was used in subsequent analyses.

For β-tubulin, γ-tubulin, EF-2, cytHSP70, HSP90, and polyubiquitin genes, one cDNA library subclone, the longest found in expressed sequence tags, was fully sequenced. The lengths of these inserts were 1,440, 1,970, 1,674, 2,259, 2,409, and 965 nt, respectively. The transcript of polyubiquitin contained three identical tandem copies of ubiquitin. The obtained sequences are available under GenBank accession numbers AY831435–AY831450.

18S rRNA Phylogeny

The *Monocercomonoides* 18S rRNA sequence formed a robust cluster with the other oxymonad sequences in the

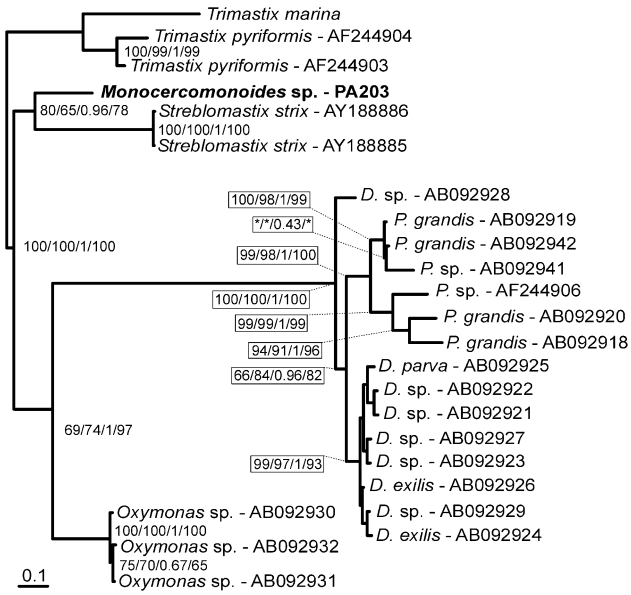

FIG. 1.—Bayesian tree constructed using the 18S rRNA sequences of oxymonads with *Trimastix* as out-group. The node values correspond to statistical evaluations of node support (distance bootstrap/maximum parsimony bootstrap/MrBayes posterior probability/maximum likelihood bootstraps). Asterisks indicate that the node was not recovered using a particular method. The genera *Dinenympha* and *Pyrsonympha* were abbreviated to *D.* and *P.*, respectively. The new sequence of *Monocercomonoides* is in bold.

phylogenetic tree that included representatives of all major eukaryotic taxa (not shown). The nearest sister taxon to oxymonads was the genus *Trimastix* (bootstrap support above 80% from all methods), consistent with previous findings (Dacks et al. 2001). The relationship of the oxymonads-*Trimastix* clade to other eukaryotic groups was unresolved. In the small-scale 18S rRNA phylogeny (fig. 1), including all available full-length oxymonad sequences and *Trimastix* as outgroup, *Monocercomonoides* and *Streblomastix strix* constituted the most basal oxymonad branch. The position of *Monocercomonoides* and the relationships among the main oxymonad groups were generally well supported with bootstrap values ranging from 65% to 80%.

Concatenation of 18S rRNA, α-tubulin, β-tubulin, γ-tubulin, EF-1α, EF-2, cytHSP70, HSP90, and Ubiquitin Genes

We have used all available sequence data for *Monocercomonoides* to elucidate the relationship of oxymonads to other eukaryotic groups. The Bayesian analysis using the concatenated data set resulted in a single tree with posterior probability 1 (fig. 2). In this tree, *Monocercomonoides* branched with the diplomonad and *Trichomonas* branch. The same placement was also recovered by the maximum likelihood and maximum parsimony methods (bootstrap 75 and 81, respectively). However, there were indications that some genes preferred a different topology. For example, in the Bayesian α-tubulin gene tree (Supplementary Material online), the diplomonad and *Trichomonas* lineages were recovered separately along with opisthokonts (animals and fungi). This placement of diplomonad and *Trichomonas* was the second most frequent among the maximum likelihood bootstrap trees (16%).

We used likelihood mapping (Strimmer and von Haeseler 1997) and the AU test (Shimodaira 2002) to examine the preferences of each gene for each of the three possible

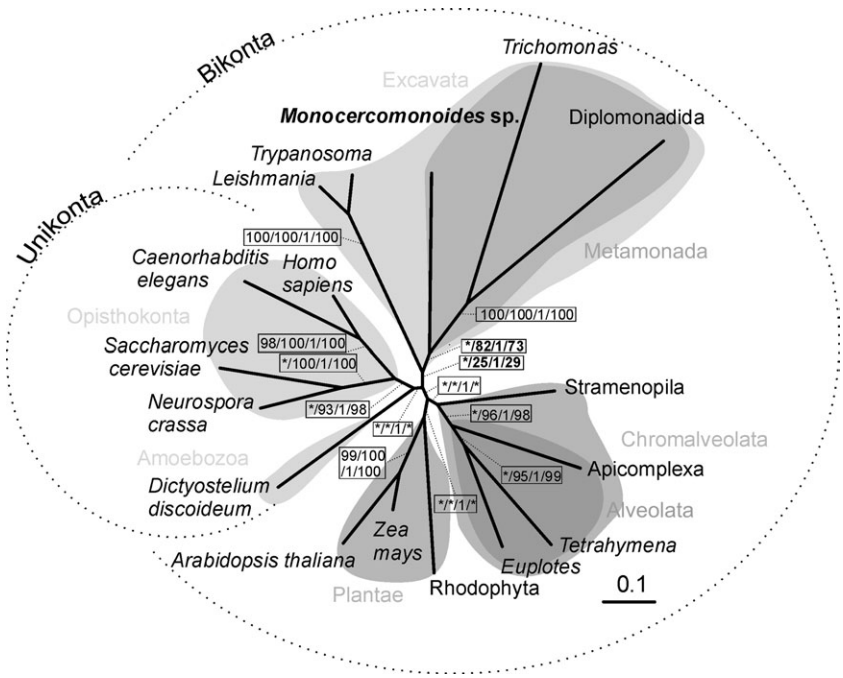

FIG. 2.—Bayesian tree constructed using the concatenate of α-tubulin, β-tubulin, γ-tubulin, EF-1α, EF-2, cytHSP70, HSP90, ubiquitin, and 18S rRNA. The node values correspond to statistical evaluations of node support (distance bootstrap/maximum parsimony bootstrap/MrBayes posterior probabilities/maximum likelihood bootstraps). Asterisks indicate that the branch leading to a given node was not recovered using a particular method.

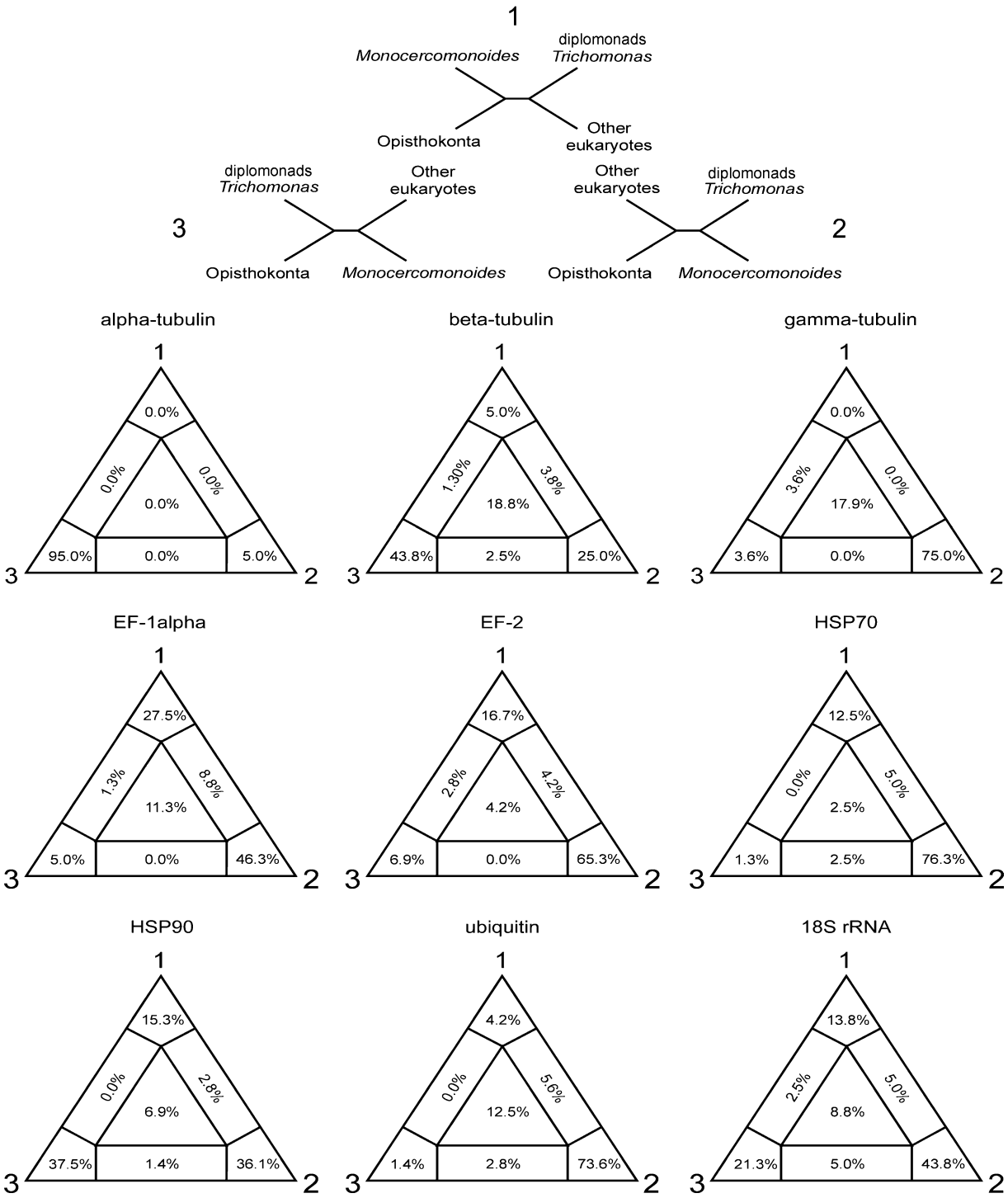

FIG. 3.—Likelihood mapping analysis. Taxa were divided into four groups: (1) *Monocercomonoides*; (2) diplomonads, *Trichomonas*; (3) opisthokonts; and (4) other eukaryotes. For all possible combinations of four taxa drawn from the four groups, likelihoods were calculated under the three possible topologies (top). These likelihoods determine the position of a particular quartet in a triangle where each apex represents one possible topology. The percentage of quartets that fall into the regions of the triangle is given for each gene.

relationships between the four groups of taxa: (1) *Monocercomonoides*; (2) *Trichomonas*, diplomonads; (3) opisthokonts; and (4) other taxa. Figure 3 summarizes the result of likelihood mapping analysis. In this analysis, likelihood of a four-taxon alignment, where each taxon is drawn from different group, is calculated under the three possible topologies. These three likelihoods determine the position of this quartet of taxa in a triangle, where each apex represents one topology of the four-group tree. This calculation is performed for all possible quartets. The graphs in figure 3 give for each gene the percentage of quartets that fall into each of the seven areas of the quartet map.

Table 1
Result of AU Test

Gene\Topology	1[a]	2	3
α-Tubulin	**0.001**	**4×10^{-5}**	Best
β-Tubulin	Best	0.421	0.232
γ-Tubulin	**0.031**	Best	0.052
EF-1α	**0.045**	Best	0.253
EF-2	0.260	Best	0.419
cytHSP70	0.263	Best	0.149
HSP90	Best	0.198	0.242
Ubiquitin	0.85	Best	**2×10^{-5}**
18S rRNA	0.295	Best	0.134

NOTE.—The hypothesis, under which the likelihood is highest, is designed as "best". Each value represents the significance (P value) of the difference between the maximum likelihood of individual gene alignments under particular topological hypothesis compared to the best hypothesis. The hypotheses that are significantly worse ($P < 0.05$) are in bold.

[a] The numbers of topological hypotheses correspond to figure 3.

Table 1 summarizes the results of AU tests. For each gene, we found (using MrBayes 2.0) the best topology under each of the three constraints of monophyly that corresponds to the topologies in figure 3: (1) monophyly of *Monocercomonoides* and opisthokonts; (2) monophyly of *Monocercomonoides*, diplomonads, and *Trichomonas*; or (3) monophyly of opisthokonts, diplomonads, and *Trichomonas*. For each gene, the significance of the differences between likelihoods under the three topologies was tested using the AU test.

The results of both likelihood mapping and the AU test showed that γ-tubulin, EF-1α, EF-2, cytHSP70, ubiquitin, and 18S rRNA genes preferred a topology in which *Monocercomonoides* formed a group with *Trichomonas* and diplomonads, but most genes could not reject the alternatives. α-tubulin sequences favored a topology in which *Trichomonas* and diplomonads formed a group with opisthokonts to the exclusion of *Monocercomonoides* and rejected both alternative topologies. β-tubulin also favored this topology in the likelihood mapping analysis, with the preference for this topology rising to 62.5% after the exclusion of the long-branches of fungi, rhodophytes, and *Dictyostelium*, but did not reject any topology in the AU test. HSP90 sequences were equivocal in their support for various hypotheses of relations.

In the knowledge of these conflicting preferences of individual genes, we performed phylogenetic analyses using four concatenates: (1) α-tubulin and β-tubulin; (2) α-tubulin, β-tubulin, and HSP90; (3) EF-1α, EF-2, γ-tubulin, cytHSP70, ubiquitin, and 18S rRNA; and (4) EF-1α, EF2, γ-tubulin, cytHSP70, ubiquitin, 18S rRNA, and HSP90. Bayesian analyses using these concatenates found, respectively, 84, 16, 6, and 1 topology with a cumulative posterior probability of 1. The consensus trees are given in figure 4.

We used the AU test to assess whether the results based on the four concatenates were significantly different. We tested the significance of differences between likelihoods of each concatenate calculated for all 107 topologies found in the four aforementioned Bayesian analyses. The test showed that all topologies resulting from the analysis of concatenates 1 and 2 were rejected by concatenates 3 and 4 and vice versa ($P \geq 0.006$). Some topologies found in the analysis of concatenate 1 were not rejected by concatenate 2 and vice versa. Similarly the results based on concatenates 3 and 4 were not significantly different.

Discussion

Brugerolle (1991) and Moriya et. al (2003) suggested a basal placement for the family Polymastigidae in the oxymonad evolutionary tree. It was also reported that *Monocercomonoides*, a representative of the family Polymastigidae, shares some cytoskeletal homologies with *Trimastix*, the closest relative of oxymonads (Simpson et al. 2002*a*). In our tree (fig. 1), *Monocercomonoides* together with *Streblomastix strix* forms the basal branch. Consistent with their affinity in the tree, the 18S rRNA sequences of both *Streblomastix strix* (2,406 nt) and *Monocercomonoides* (2,927 nt) are longer than other oxymonad sequences, and some expanded regions, though different in sequence, were shared between the two 18S rRNA genes. On the other hand, the attachment organelle (holdfast) and contractile axostyle of *Streblomastix* is similar to Pyrsonymphidae and Oxymonadidae, while *Monocercomonoides* lacks both structures. If the 18S rRNA tree topology is correct, these characters must either have originated independently in *Streblomastix* on one side and Pyrsonymphidae, Oxymonadidae on the other, or they must have been secondarily lost in *Monocercomonoides*.

A noncanonical genetic code was observed in *Streblomastix*, in which the stop codons TAA and TAG encode the amino acid glutamine (Keeling and Leander 2003). A similar genetic code was also reported in some, but not all, species of diplomonads (Keeling and Doolitle 1996). Such deviations from the standard code may represent an autapomorphy of *Streblomastix* because they have neither been reported in other oxymonad genera nor do we observe such modifications in *Monocercomonoides*.

To infer the position of oxymonads in the eukaryotic tree, we analyzed all available data for *Monocercomonoides*. The analysis of the concatenate of all genes gave a tree in which oxymonads robustly branched with *Trichomonas* and diplomonads and in which several suggested eukaryotic supergroups were recovered (fig. 2). However, it also appeared that some genes disagreed with this topology. Likelihood mapping and the AU test showed that the topology was supported by γ-tubulin, EF-1α, EF-2, cytHSP70, ubiquitin, and 18S rRNA genes, while α-tubulin, and to a lesser extent β-tubulin, preferred a different placement for *Monocercomonoides*. The HSP90 gene was indecisive on this issue. Unsurprisingly, phylogenetic analyses based on the concatenate of α-tubulin and β-tubulin and the concatenate of γ-tubulin, EF-1α, EF-2, cytHSP70, ubiquitin, and 18S rRNA produced significantly different topologies (fig. 4*a* and *c*). The inclusion of HSP90 sequences in either concatenate did not change the topology of robustly supported nodes, but it did improve support in both cases for another contentious supergroup—the chromalveolates. The analysis of six genes plus HSP90 (fig. 4*d*) produced the same topology as the analysis based on all genes (fig. 2) with a cluster comprising the green plants + rhodophytes and a cluster including all the excavates—*Trypanosoma*, *Leishmania*, *Monocercomonoides*, *Trichomonas*, and diplomonads. The similar outcome of the two analyses apparently results from the fact that, in the all-gene

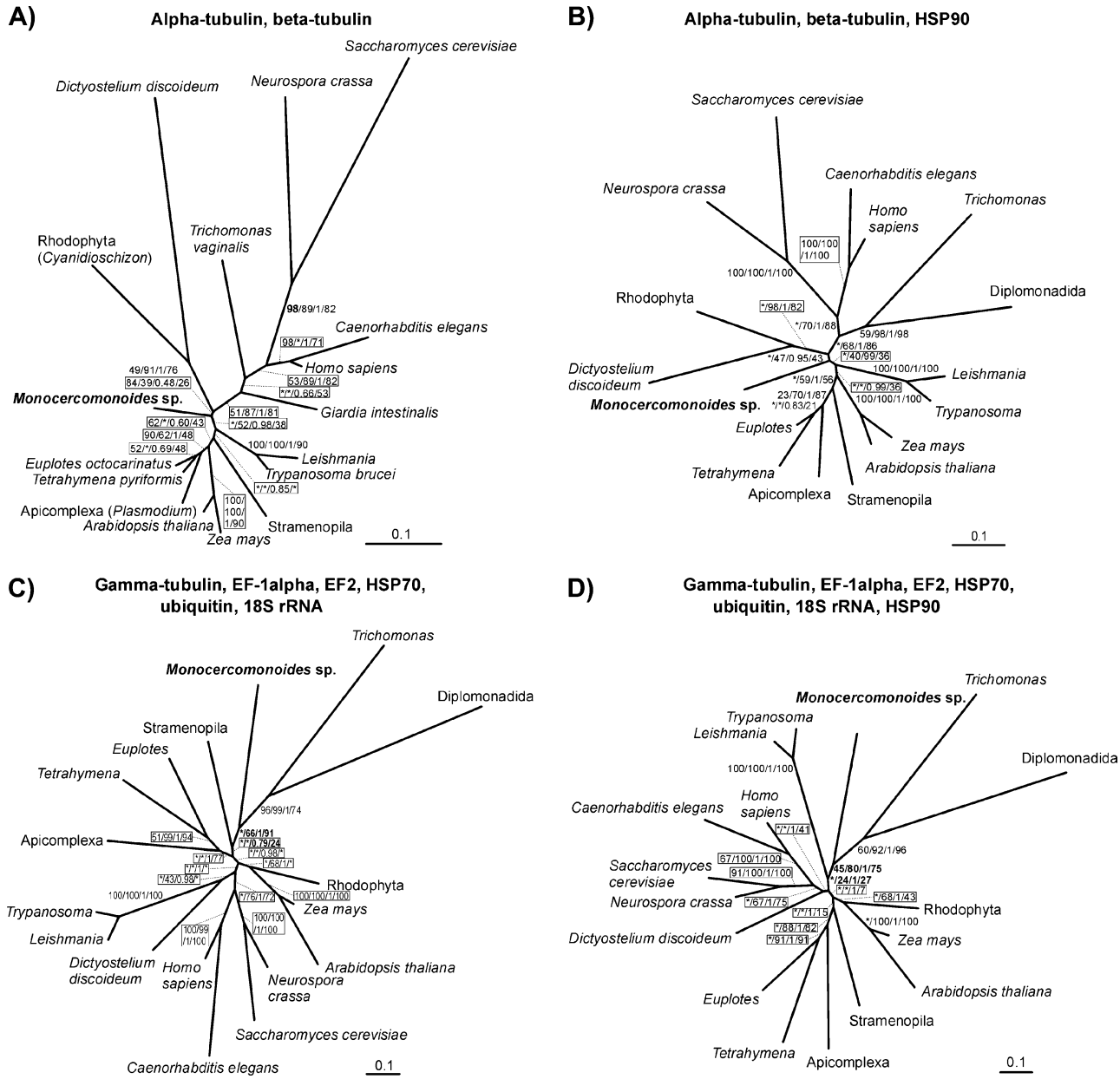

FIG. 4.—Bayesian trees constructed on the basis of four different concatenates. The node values correspond to statistical evaluations of node support (distance bootstrap/maximum parsimony bootstrap/MrBayes posterior probabilities/maximum likelihood bootstraps). Asterisks indicate that the branch leading to a given node was not recovered using a particular method.

analysis, the signal contained in γ-tubulin, EF-1α, EF-2, cytHSP70, ubiquitin, and 18S rRNA genes was greater than the signal from α- and β-tubulins.

It is clear that α-tubulin and β-tubulin favor a topology, and in particular a placement for oxymonads, which is significantly incongruent with most of the other genes. Similar incongruence was observed by Moriya et al. (2001) on a smaller sampling of genes (α-tubulin versus EF-1α). This disagreement between α-, β-tubulin and other proteins and rRNA genes was also reported by Arisue, Hasegawa, and Hashimoto (2005), who analyzed 24 genes together. Although we cannot determine which (if either) of these groups of genes reflects the real evolutionary history of the organisms, we favor the tree based upon the γ-tubulin, EF-1α, EF-2, cytHSP70, ubiquitin, 18S rRNA, and HSP90 because it contains the greatest number of functionally independent proteins and rRNA. This data set also supports the relationship of plants and red algae, which is consistent with other data (McFadden and van Dooren 2004), while in the α-, β-tubulin tree, red algae form a well-supported but probably incorrect group with *Dictyostelium*. The incongruence of the two data sets could be explained either by the presence of tree construction artifact that obscures the result of one of the analyses or by the fact that the evolutionary history of gene(s) in one set differs from the history of other genes, for example if some gene(s) had undergone horizontal transfer or if there was hidden paralogy. In either case, the two-gene analysis is more likely to be affected by these

phenomena, particularly if the two proteins, like tubulins, are functionally dependent.

The results of the seven- and nine-gene analyses (figs. 2 and 4*d*) moderately support the existence of clusters of opisthokonts, green plants + rhodophytes, and alveolates. Bayesian analyses (the only analyses that included the 18S rDNA gene and the partition-specific covarion models) also supported the existence of the proposed supergroups chromalveolates (Cavalier-Smith 1999) and corticates (Cavalier-Smith 2003*b*) and the unikont/bikont split (Stechmann and Cavalier-Smith 2003). The relationship of *Monocercomonoides* with diplomonads and *Trichomonas* supports the monophyly of the phylum Metamonada amended by Cavalier-Smith (2003*a*) which comprises: Eopharyngia (diplomonads, enteromonads, retortamonads), *Carpediemonas*, Anaeromonada (oxymonads and *Trimastix*), and Parabasalia. Moreover, the maximum parsimony– and maximum likelihood–based phylogenies suggest, albeit rather weakly, the existence of an excavate supergroup. The sampling of metamonads and excavates in our tree is rather sparse, but considering multiple lines of evidence indicating that retortamonads and *Carpediemonas* are related to diplomonads (Silberman et al. 2002; Simpson et al. 2002*b*), *Trimastix* to oxymonads (Dacks et al. 2001), and Heterolobosea and *Reclinomonas* to Euglenozoa (Baldauf et al. 2000; Simpson and Roger 2004), we can conclude that this analysis presents robust molecular evidence for the monophyly of Metamonada and suggests, albeit weakly, the monophyly of at least nine excavate groups (diplomonads, parabasalids, Heterolobosea, *Reclinomonas*, Euglenozoa, *Trimastix*, oxymonads, retortamonads, and *Carpediemonas*).

Conclusions

Careful analysis of new sequence data from *Monocercomonoides* has considerably improved our understanding of phylogenetic relationships among oxymonads as well as the position of oxymonads in the eukaryotic tree. *Monocercomonoides* formed the most basal branch in the oxymonad tree together with the genus *Streblomastix*. We detected some conflicting signal among genes for the position of oxymonads. This conflict seems to originate from the α-tubulin and less strongly β-tubulin gene sequences. After exclusion of α- and β-tubulin, phylogenetic analyses strongly supported the monophyly of Metamonada comprising oxymonads, diplomonads, and parabasalids in our analyses. In addition, likelihood- and parsimony-based methods recovered all the included excavates together. Although support for these relationships was not high, it is nevertheless some of the strongest data supporting the excavate hypothesis. The results of our analyses demonstrate that concatenation of data can improve the signal for relationships and also that care is needed to investigate possible incongruence between partitions.

Supplementary Material

Supplementary Table with an overview of the expanded regions of *Monocercomonoides* 18S rRNA, alignments, and gene trees are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The work was supported by the grants GAČR 204/03/1243, and MŠMT project 0021620828. Preliminary sequence data was obtained from The Institute for Genomic Research through the Web site http://www.tigr.org. Sequencing of *Trichomonas vaginalis* genome was accomplished with support from National Institute of Allergy and Infectious Diseases.

Literature Cited

Andersson, J. O., S. W. Sarchfield, and A. J. Roger. 2004. Gene transfer from Nanoarchaeota to an ancestor of diplomonads and parabasalids. Mol. Biol. Evol. **21**:1–6.

Arisue, N., M. Hasegawa, and T. Hashimoto. 2005. Root of the eukaryota tree as inferred from combined likelihood analyses of multiple molecular sequence data. Mol. Biol. Evol. **22**: 409–420.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. More, J. G. Seidman, J. A. Smith, and K. Struhl. 2001. Current protocols in molecular biology. John Wiley & Sons Inc., Stafford, Australia.

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

Bapteste, E., H. Brinkmann, J. A. Lee, D. V. Moore, C. W. Sensen, P. Gordon, L. Durufle, T. Gaasterland, P. Lopez, M. Muller, and H. Philippe. 2002. The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. Proc. Natl. Acad. Sci. USA **99**:1414–1419.

Brugerolle, G. 1991. Flagellar and cytoskeletal systems in amitochondrial flagellates—Archamoeba, Metamonada and Parabasala. Protoplasma **164**:70–90.

Brugerolle, G., and J. J. Lee. 2000. Order Oxymonadida. Pp. 1186–1195 *in* J. Lee, G. F. Leedale, and P. Bradbury, eds. The illustrated guide to the Protozoa. J Allen Press Inc., Lawrence, Kans.

Cavalier-Smith, T. 1983. Endosymbiotic origin of the mitochondrial envelope. Pp. 265–279 *in* W. Schwemmler and H. E. A. Schenk, eds. Endocytobiology II. de Gruyter, Berlin, Germany.

———. 1998. A revised six-kingdom system of life. Biol. Rev. Camb. Philos. Soc. **73**:203–266.

———. 1999. Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J. Eukaryot. Microbiol. **46**:347–366.

———. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. Int. J. Syst. Evol. Microbiol. **52**:297–354.

———. 2003*a*. The excavate protozoan phyla Metamonada Grasse emend. (Anaeromonadea, Parabasalia, *Carpediemonas*, Eopharyngia) and Loukozoa emend. (Jakobea, *Malawimonas*): their evolutionary affinities and new higher taxa. Int. J. Syst. Evol. Microbiol. **53**:1741–1758.

———. 2003*b*. Protist phylogeny and the high-level classification of Protozoa. Eur. J. Protistol. **39**: 338–348.

Clark, C. G., and A. J. Roger. 1995. Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. Proc. Natl. Acad. Sci. USA **92**:6518–6521.

Dacks, J., and A. J. Roger. 1999. The first sexual lineage and the relevance of facultative sex. J. Mol. Evol. **48**:779–783.

Dacks, J. B., A. Marinets, W. F. Doolittle, T. Cavalier-Smith, and J. M. Logsdon. 2002. Analyses of RNA polymerase II genes

from free-living protists: phylogeny, long branch attraction, and the eukaryotic big bang. Mol. Biol. Evol. **19**:830–840.

Dacks, J. B., J. D. Silberman, A. G. B. Simpson, S. Moriya, T. Kudo, M. Ohkuma, and R. J. Redfield. 2001. Oxymonads are closely related to the excavate taxon *Trimastix*. Mol. Biol. Evol. **18**:1034–1044.

Edgcomb, V. P., A. J. Roger, A. G. B. Simpson, D. T. Kysela, and M. L. Sogin. 2001. Evolutionary relationships among "jakobid" flagellates as indicated by alpha- and beta-tubulin phylogenies. Mol. Biol. Evol. **18**:514–522.

Felsenstein, J. 1993. PHYLIP (phylogeny inference package). Version 3.6a2. Distributed by the author, Department of Genetics, University of Washington, Seattle.

Foster, P.G. 2004. Modeling compositional heterogeneity. Syst. Biol. **53**:485–495.

Germot, A., H. Philippe, and H. Le Guyader. 1996. Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. Proc. Natl. Acad. Sci. USA **93**:14614–14617.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**:696–704.

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

Henze, K., D. S. Horner, S. Suguri, D. V. Moore, L. B. Sánchez, M. Müller, and T. M. Embley. 2001. Unique phylogenetic relationships of glucokinase and glucosephosphate isomerase of the amitochondriate eukaryotes *Giardia intestinalis*, *Spironucleus barkhanus* and *Trichomonas vaginalis*. Gene **281**: 123–131.

Hirt, R. P., J. M. Logsdon, B. Healy, M. W. Dorey, W. F. Doolittle, and T. M. Embley. 1999. Microsporidia are related to Fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proc. Natl. Acad. Sci. USA **96**:580–585.

Horner, D. S., R. P. Hirt, S. Kilvington, D. Lloyd, and T. M. Embley. 1996. Molecular data suggest an early acquisition of the mitochondrion endosymbiont. Proc. R. Soc. Lond. B Biol. Sci. **263**:1053–1059.

Hrdy, I., R. P. Hirt, P. Dolezal, L. Bardonová, P. G. Foster, J. Tachezy, and T. M. Embley. 2004. *Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. Nature **432**:618–622.

Keeling, P. J., and W. F. Doolitle. 1996. A non-canonical genetic code in an early diverging eukaryotic lineage. EMBO J. **15**:2285–2290.

Keeling, P. J., and B. S. Leander. 2003. Characterisation of a non-canonical genetic code in the oxymonad *Streblomastix strix*. J. Mol. Biol. **326**:1337–1349.

Lang, B. F., M. W. Gray, and G. Burger. 1999. Mitochondrial genome evolution and the origin of eukaryotes. Annu. Rev. Genet. **33**:351–397.

Mai, Z. M., S. Ghosh, M. Frisardi, B. Rosenthal, R. Rogers, and J. Samuelson. 1999. Hsp60 is targeted to a cryptic mitochondrion-derived organelle ("crypton") in the microaerophilic protozoan parasite *Entamoeba histolytica*. Mol. Cell. Biol. **19**:2198–2205.

McFadden, G. I., and G. G. van Dooren. 2004. Evolution: red algal genome affirms a common origin of all plastids. Curr. Biol. **14**:R514–R516.

Medlin, L., H. J. Elwood, S. Stickel, and M. L. Sogin. 1988. The characterisation of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene **71**:491–499.

Moriya, S., J. B. Dacks, A. Takagi, S. Noda, M. Ohkuma, W. F. Doolittle, and T. Kudo. 2003. Molecular phylogeny of three oxymonad genera: *Pyrsonympha*, *Dinenympha* and *Oxymonas*. J. Eukaryot. Microbiol. **50**:190–197.

Moriya S., M. Ohkuma, and T. Kudo. 1998. Phylogenetic position of symbiotic protist *Dinenympha* [correction of *Dinemympha*] exilis in the hindgut of the termite *Reticulitermes speratus* inferred from the protein phylogeny of elongation factor-1α. Gene **210**:221–227.

Moriya, S., K. Tanaka, M. Ohkuma, S. Sugano, and T. Kudo. 2001. Diversification of the microtubule system in the early stage of eukaryote evolution: elongation factor-1α and α-tubulin protein phylogeny of termite symbiotic oxymonad and hypermastigote protists. J. Mol. Evol. **52**: 6–16.

Nikolaev, S. I., C. Berney, J. F. Fahrni, I. Bolivar, S. Polet, A. P. Mylnikov, V. V. Aleshin, N. B. Petrov, and J. Pawlowski. 2004. The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. Proc. Natl. Acad. Sci. USA **101**:8066–8071.

Philippe, H. 2000. Opinion: long branch attraction and protist phylogeny. Protist **151**:307–316.

Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics **14**:817–818.

Roger, A. J., S. G. Svard, J. Tovar, C. G. Clark, M. W. Smith, F. D. Gillin, and M. L. Sogin. 1998. A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. Proc. Natl. Acad. Sci. USA **95**:229–234.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572–1574.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. **51**:492–508.

Shimodaira, H., and M. Hasegawa. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics **17**:1246–1247.

Silberman, J. D., A. G. B. Simpson, J. Kulda, I. Cepicka, V. Hampl, P. J. Johnson, and A. J. Roger. 2002. Retortamonad flagellates are closely related to diplomonads—implications for the history of mitochondrial function in eukaryote evolution. Mol. Biol. Evol. **19**:777–786.

Simpson, A. G. B. 2003. Cytoskeletal organization, phylogenetic affinities and systematics in the contentious taxon Excavata (Eukaryota). Int. J. Syst. Evol. Microbiol. **53**: 1759–1777.

Simpson, A. G. B., and D. J. Patterson. 1999. The ultrastructure of *Carpediemonas membranifera* (Eukaryota) with reference to the "excavate hypothesis". Eur. J. Protistol. **35**:353–370.

Simpson, A. G. B., R. Radek, J. B. Dacks, and C. J. O'Kelly. 2002*a*. How oxymonads lost their groove: an ultrastructural comparison of *Monocercomonoides* and excavate taxa. J. Eukaryot. Microbiol. **49**:239–248.

Simpson, A. G. B., and A. J. Roger. 2004. Excavata and the origin of amitochondriate eukaryotes. Pp. 27–53 *in* R. Hirt and D. S. Horner, eds. Organelles, genomes and eukaryote phylogeny. CRC Press, London, UK.

Simpson, A. G. B., A. J. Roger, J. D. Silberman, D. D. Leipe, V. P. Edgcomb, L. S. Jermiin, D. J. Patterson, and M. L. Sogin. 2002*b*. Evolutionary history of "early-diverging" eukaryotes: the excavate taxon *Carpediemonas* is a close relative of *Giardia*. Mol. Biol. Evol. **19**:1782–1791.

Sogin, M. L., and J. D. Silberman. 1998. Evolution of the protists and protistan parasites from the perspective of molecular systematics. Int. J. Parasitol. **28**:11–20.

Stechmann, A., and T. Cavalier-Smith. 2003. The root of the eukaryote tree pinpointed. Curr. Biol. **13**: R665–R666.

Stiller, J. W., and B. D. Hall. 1999. Long-branch attraction and the rDNA model of early eukaryotic evolution. Mol. Biol. Evol. **16**:1270–1279.

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

———. 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc. Natl. Acad. Sci. USA **94**:6815–6819.

Swofford, D. L. 1998. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Mass.

Tachezy, J., L. B. Sánchez, and M. Müller. 2001. Mitochondrial type iron-sulfur cluster assembly in the amitochondriate eukaryotes *Trichomonas vaginalis* and *Giardia intestinalis*, as indicated by the phylogeny of IscS. Mol. Biol. Evol. **18**:1919–1928.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**:4876–4882.

Tovar, J., A. Fischer, and C. G. Clark. 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. Mol. Microbiol. **32**: 1013–1021.

Tovar, J., G. Leon-Avila, L. B. Sanchez, R. Sutak, J. Tachezy, M. van der Giezen, M. Hernandez, M. Muller, and J. M. Lucocq. 2003. Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. Nature **426**:172–176.

Williams, B. A. P., R. P. Hirt, J. M. Lucocq, and T. M. Embley. 2002. A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. Nature **418**:865–869.

Geoffrey McFadden, Associate Editor

Accepted August 15, 2005