J Mol Evol (2005) 61:765–775
DOI: 10.1007/s00239-005-0006-6

JOURNAL OF MOLECULAR EVOLUTION

Phylogeny of the Centrohelida Inferred from SSU rRNA, Tubulins, and Actin Genes

Miako Sakaguchi, Takeshi Nakayama, Tetsuo Hashimoto, Isao Inouye

Institute of Biological Sciences, University of Tsukuba, Tsukuba 305-8572, Japan

Received: 10 January 2005 / Accepted: 29 June 2005 [*ReviewingEditor:* Dr. Martin Kreitman]

Abstract. Amoeboid protists are major targets of recent molecular phylogeny in connection with reconstruction of global phylogeny of eukaryotes as well as the search for the root of eukaryotes. The Centrohelida are one of the major groups of Heliozoa, classified in the Actinopodida, whose evolutionary position is not well understood. To clarify the relationships between the Centrohelida and other eukaryotes, we sequenced SSU rRNA, α-tubulin, and β-tubulin genes from a centroheliozoan protist, *Raphidiophrys contractilis*. The SSU rRNA phylogeny showed that the Centrohelida are not closely related to other heliozoan groups, Actinophryida, Desmothoracida, or Taxopodida. Maximum likelihood analyses of the combined phylogeny using a concatenate model for an α- + β-tubulin + actin data set, and a separate model for SSU rRNA, α- and β-tubulin, and actin gene data sets revealed the best tree, in which the Centrohelida have a closer relationship to Rhodophyta than to other major eukaryotic groups. However, both weighted Shimodaira–Hasegawa and approximately unbiased tests for the concatenate protein phylogeny did not reject alternative trees in which Centrohelida were constrained to be sisters to the Amoebozoa. Moreover, alternative trees in which Centrohelida were placed at the node branching before and after Amoebozoa or Viridiplantae were not rejected by the WSH tests. These results narrowed the possibilities for the position of Centrohelida to a sister to the Rhodophyta, to the Amoebozoa, or to an independent branch between the branchings of Amoebozoa and Rhodophyta (or possibly Plantae) at the basal position within the bikonts clade in the eukaryotic tree.

Key words: Heliozoa — Centrohelida — Tubulin — Maximum likelihood — Concatenate phylogeny — Protists

Correspondence to: Miako Sakaguchi; *email:* miako@sakura.cc.tsukuba.ac.jp

Introduction

Single-gene analyses such as of SSU rDNA, actin, and β-tubulin, intensively performed for the last decade or so, have demonstrated the monophyly of major groups of eukaryotes, though evolutionary and phylogenetic relationships among many of these groups were unresolved. This circumstance was generally interpreted as the consequence of the Big Bang radiation of eukaryotes collectively called crown eukaryotes (e.g., Knoll 1992). However, recent phylogenetic analyses of combined multiple gene data have begun, with certain confidence, to clarify the eukaryote phylogeny by resolving relationships among major groups of the crown and have demonstrated the existence of several supergroups in which any eukaryote can be placed. For example, phylogenies based on combined mitochondrial genes (Burger et al. 1999) or nuclear genes (Moreira et al. 2000) supported that the Rhodophyta and Viridiplantae are sister groups; the Euglenozoa and Heterolobosea form a clade in combined analyses of

EF-1α, actin, and α- and β-tubulin (Baldauf et al. 2000); the Alveolata and stramenopiles have close relationships with weak support in Baldauf et al.'s (2000) analysis and with clear support in the combined phylogeny of eight genes including rRNAs (Arisue et al. 2002b); phylogenetic analyses of five plastid genes suggested the monophyly of plastids of the Chromista, a taxon comprising the Cryptophyta, Haptophyta, and photosynthetic stramenopiles originally defined by sharing four-membrane-bounded chloroplasts with ribosomes on the outermost membrane (Yoon et al. 2002); and combined phylogeny of 22 proteins demonstrated that the Diplomonadida and Parabasalia shared a common ancestor in the rooted eukaryotic tree (Arisue et al. 2005).

Amoeboid protists are the protistan assemblage for which the taxonomy and phylogenetic view have changed most drastically. In the last couple of years, molecular data have been reported for representatives of the major groups of the protozoan taxon Sarcodina, even though the number of data are still insufficient. The Sarcodina have now collapsed and split into eukaryotic supergroups such as Amoebozoa, Heterolobosea, Rhizaria, stramenopiles, and Opisthokonta. These amoeboid protists are probably the most important targets of recent phylogenetic studies, because they are of significance in search for the root of global eukaryotic tree. Many recent studies have suggested that accurate positioning of very diverse amoeboid protists would lead us to a comprehensive understanding of eukaryotic phylogeny (Stechmann and Cavalier-Smith 2002, 2003; Arisue et al. 2005).

Amoeboid protists having pseudopodia used for locomotion and food capture were assembled in the Sarcodina Schmarda 1871 (Levine et al. 1980) and classified into two groups, those possessing lobose or filose pseudopodia (the Rhizopoda) and those possessing radiating axopodia stiffened by a bundle of axonemal microtubules (the Actinopoda). In recent molecular phylogenetic studies, polyphyly of these amoeboid protists has become evident, and many have been split and placed into several supergroups. A phylogenetic analysis of combined data of SSU and LSU rRNAs, EF-1α, and EF-2 revealed that amitochondrial amoebozoan protists *Mastigamoeba* and *Entamoeba* are closely related to Mycetozoa (Arisue et al. 2002a), and a combined data set of over 100 genes demonstrated the monophyly of Conosa (Bapteste et al. 2002). An assemblage comprising of the cellular slime molds, pelobionts, and entamoebids (Cavalier-Smith 1998), all of which are placed in a supergroup, the Amoebozoa, possibly positioned near the root of eukaryotes (Stechmann and Cavalier-Smith 2002, 2003). Fahrni et al. (2003), based on actin and SSU rRNA phylogenies, suggested the monophyly of Amoebozoa as a supergroup including lobose amoebae, Archamoebae, and Mycetozoa. On the other hand, actin, SSU rRNA, or RPB1 phylogeny revealed that the Foraminifera, a huge group of Rhizopoda that are of ecological and paleontological importance, are close relatives of the Cercozoa, another supergroup recognized by molecular analyses (Keeling 2001; Berney and Pawlowski 2003; Longet et al. 2003).

The Actinopoda have been traditionally divided on morphological basis into the Radiolaria and Heliozoa. Radiolaria are classified into three groups, Acantharea, Polycystinea, and Phaeodarea (Haeckel 1887), and they have a skeleton composed of silica or strontium sulfate and a central capsule (Lee et al. 2000). The central capsule separates endoplasm containing organelles such as nucleus (or nuclei) and mitochondria from frothy ectoplasm. The Heliozoa are distinguished from Radiolaria by lacking central capsules and intricate skeletal elements, and some heliozoan groups produce a lorica and stalk, scales, and spines (Febvre-Chevalier 1990). They are very diverse and morphologically classified into four orders: Actinophryida, Centrohelida, Desmothoracida, and Gymnosphaerida (Mikrjukov et al. 2000). *Sticholonche zanclea*, a unique amoeboid protist possessing mobile axopodia for rowing, has sometimes been classified into the Heliozoa and treated as a sole member of the fifth order, Taxopodida (Febvre-Chevalier 1990). Phylogeny of the Actinopoda has progressed in the last few years, and the members of Actinopoda are now phylogenetically distinct. Phylogenetic analyses of SSU rRNA gene sequences of radiolarian Acantharea and Polycystinea using environmental SSU rRNA data showed their monophyly (López-García et al. 2002), while the Phaeodarea is placed separately from the other two radiolarians in the tree (Polet et al. 2004). Recent phylogenetic analyses of heliozoan groups demonstrated probably accurate positioning of major groups of the Heliozoa into different supergroups, i.e., the Actinophryida fall within the stramenopiles with a posterior probability (PP) of 1.0 for SSU rRNA and actin genes and a bootstrap proportion (BP) of 71% and 98% for each of the genes, Desmothoracida branch among Cercozoa with a PP of 1.0 and a BP of 94% for the SSU rRNA gene, and Taxopodida branch between the radiolarian Acantharea and Polycystinea with a PP of 1.0 and a BP of 80% for the SSU rRNA gene; the two heliozoans form a larger clade, Rhizaria, with Radiolaria, Cercozoa, and Foraminifera, with a PP of 1.0 for both genes (Nikolaev et al. 2004). *Ciliophrys* and *Pteridomonas*, traditionally classified into the Actinophryida, are heterotrophic members of the Dyctiochophyceae and belong to the stramenopiles (Sekiguchi et al. 2002), and a *Dimorpha*-like strain, an organism sometimes placed in the Heliozoa, falls within the Cercozoa (Cavalier-Smith and Chao

2003b). The first SSU rRNA data were also recently reported from the Centrohelida and show that the Centrohelida are sisters to Haptophyta (distance trees) or to the apusozoan *Ancyromonas* (quartet-puzzling tree) as a low-supported group (Cavalier-Smith and Chao 2003a). In the Bayesian analyses, the SSU rRNA data weakly support the clade consisting of the Centrohelida and Cryptophyta + Glaucophyta (PP, 0.67; BP, < 50%) and the actin data show that the Centrohelida have no relationship to any member of the eukaryotes (Nikolaev et al. 2004). These analyses indicate that the Centrohelida form an independent lineage in the eukaryotic tree.

The discovery of fused genes encoding dihydrofolate reductase (DHFR) and thymidylate synthase (TS) distributed in a wide range of eukaryotes suggested the existence of a very large clade that comprises all eukaryotes except the Amoebozoa and Opisthokonta (Stechmann and Cavalier-Smith 2002). This clade is designated bikonts after their possible biflagellate ancestral state. It is suggested that any eukaryote could be classified in one of these three clades. More importantly, if the DHFR-TS gene fusion is really a derived character, it suggests that the possible position of the root of eukaryotic tree lies between the bikonts and the Opithokonta, in which DHFR and TS genes are encoded by separate genes (Stechmann and Cavalier-Smith 2002).

The position of Amoebozoa is still argued in the scope of search for the root (Simpson and Roger 2002; Stechmann and Cavalier-Smith 2002, 2003; Baldauf 2003). The gene fusion has not been found in Amoebozoa; instead, the separate DHFR and TS genes were found in the lobosean amoeba *Hartmannella cantabrigiensis* (Stechmann and Cavalier-Smith 2003). Stechmann and Cavalier-Smith (2002) showed that the centrohelid *Chlamydaster sterni* also has the gene fusion, and they belong to the bikonts. For rooting the eukaryotic tree and understanding the early diversification of eukaryotes, it is very important to clarify the position of these very diverse amoeboid protists. The Centrohelida are one of the protistan amoebae whose phylogenetic position is not convincingly resolved (Cavalier-Smith 2004). They are not closely related to any other eukaryotes in individual SSU rRNA or actin phylogeny and Bayesian consensus trees of the two genes (Cavalier-Smith and Chao 2003a; Nikolaev et al. 2004), indicating that the SSU rRNA and actin data sets currently available do not contain enough phylogenetic signal for resolving the position of Centrohelida.

In this study, we sequenced SSU rRNA, α-tubulin, and β-tubulin of the centroheliozoan *Raphidiophrys contractilis* and performed combined maximum likelihood (ML) analyses of these genes using the "concatenate" and the "separate" models for the estimation of branch lengths (Pupko et al. 2002). Combined ML analyses using both models revealed the best tree in which the Centrohelida have a closer relationship to the Rhodophyta than to any other eukaryotic groups. However, statistical comparison of the alternative trees by the weighted Shimodaira–Hasegawa (WSH) and approximately unbiased (AU) tests also suggested the possibility that the Centrohelida are closely related to Amoebozoa, or located between the divergences of Amoebozoa and Rhodophyta (or possibly Plantae), and thus were placed at the basal position within the bikonts clade.

Materials and Methods

PCR Amplication and Sequencing

The centroheliozoan *R. contractilis* was cultured monoxenically as described by Sakaguchi et al. (2001). Genomic DNA was extracted with chloroform:isoamyl alcohol (24:1) after cell lysis and precipitated with sodium acetate, isopropanol, and ethanol. SSU rRNA genes were amplified with primers previously designed by Nakayama et al. (1996), and α- and β-tubulin genes were amplified with primers as described by Edgcomb et al. (2001) and Leander et al. (2003). All amplifications of SSU rRNA and tubulin genes consisted of 30 cycles at a denaturing temperature of 94°C for 1 min, an annealing temperature of 50°C for 1 min, and an extension temperature of 72°C for 1 min. PCR products were purified using the Geneclean Kit (BIO101) and were used for direct sequencing using the DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences). Some PCR products were cloned into pGEM T-easy vector (Promega). Bidirectional sequencing was carried out on an automated ABI-310 sequencer. The new sequences of *R. contractilis* were deposited in GenBank as follows: SSU rRNA (AB196984), α-tubulin (AB196985), and β-tubulin (AB196986).

Alignments and Phylogenetic Analyses

The SSU rRNA sequence alignment was adjusted manually, taking into account the secondary structure information obtained from the rRNA database (http:// www.psb.ugent.be/rRNA/) (Wuyts et al. 2004). Some eukaryotic groups such as Diplomonadida and Parabasalia were excluded from the data set in order to avoid possible violation of the analysis affected by LBA. A total of 1251 unambiguously aligned positions from 60-OTU were used for phylogenetic analysis (the alignment is available from M.S. upon request). The ML analyses were performed by DNAML using PHYLIP version 3.6a (Felsenstein 2002), employing the HKY85 model of substitution (Hasegawa et al. 1985) with global rearrangements and one-jumble options. An among-site rate variation was modeled on a Γ distribution with eight rate categories. The Γ-shape parameter (α) estimated by PAML version 3.1 (Yang 1997) was 0.373.

The concatenate α- + β-tubulin and the tubulins + actin amino acid sequence data were aligned by eye (the alignments are available from M.S. upon request). The complete alignments contained 31 taxa and 771 sites and 28 taxa and 1028 sites, respectively, and were used for combined analysis with the concatenate model. Glaucophyta was excluded from concatenate protein data sets, because the α-tubulin sequence of *Cyanophora paradoxa* is much shorter than that of other eukaryotic groups. In order to increase the OTU number of representative eukaryotic groups, three com-

Table 1. Eight eukaryotic groups used for 10,395 possible constrained SSU rRNA trees

Amoebozoa
((*Physarum*, *Dictyostelium*), ((*Leptomyxa*, *Hartmannella*), (*Entamoeba*, *Mastigamoeba*)))
Centrohelida
(*Heterophrys*, *Chlamydaster*, (*Raphidiophrys contractilis*, *Raphidiophrys ambigua*))
Rhodophyta
(*Porphyra*, nucleomorph of *Hanusia*, (*Cyanidium*, *Cyanidioschyzon*))
Viridiplantae
((*Chlorella*, *Chlamydomonas*), (*Zea*, *Arabidopsis*))
Cercozoa
(*Cercomonas*, *Dimorpha*-like sp., *Chlorarachnion*, (*Hedriocystis*, *Clathrulina*))
Cryptomonada
(*Goniomonas*, (*Guillardia*, *Hanusia*))
Euglenozoa
(*Euglena*, (*Trypanosoma*, *Crithidia*))
Alveolata + stramenopiles
((*Tetrahymena*, (*Prorocentrum*, (*Cryptosporidium*, (*Toxoplasma*, *Plasmodium*)))), (*Phytophthora*, (*Pteridomonas*, *Fucus*, *Skeletonema*, *Actinosphaerium*))

posite taxa, "Cercomonas" (α-tubulin from *Cercomonas* sp. ATCC50319 and β-tubulin from *Cercomonas* sp. ATCC50316), "Leishmania" (α-tubulin from *Leishmania donovani* and β-tubulin from *Leishmania mexicana*), and "stramenopile (Phaeophyceae)" (α-tubulin from *Pelvetia fastigiata* and β-tubulin from *Ectocarpus variabilis*), were included in the alignment of tubulins, and five composite taxa, "Cercomonas" (tubulins composite data and actin from *Cercomonas* sp. ATCC50317), "Leishmania" (tubulins composite data and actin from *Leishmania major*), "stramenopile" (tubulins composite data and actin from *Fucus distichus*), "Bangiophyceae" (tubulins from *Cyanidium caldarium* and actin from *Porphyra yezoensis*), and "Raphidiophrys" (tubulins from *Raphidiophrys contractilis* and actin from *Raphidiophrys ambigua*), were included in the tubulins + actin sequence data set. The amino acid composition of each protein data set was analyzed using TREE-PUZZLE version 5.2, and there was no statistically significant compositional bias for any species analyzed. ML analyses were performed with PROML in PHYLIP, using the JTT-F substitution matrix (Jones et al. 1992) and assuming homogeneous site rates. Based on the best trees, the Γ-shape parameters (α) were estimated by the use of PAML for the tubulins and tubulins + actin data sets. These values were 0.777 and 0.748, respectively. Then using these α values, the PROML analyses employing the JTT-F + Γ model (among-site rate variation model with eight rate categories) were performed with the global rearrangements and the one-jumble options.

In the neighbor-joining (NJ) analyses of SSU rRNA and the concatenate protein data, distance matrices were calculated using DNADIST and PROTDIST in PHYLIP, based, respectively, on the HKY85 and JTT-F models with rate variation among sites allowed. The NJ trees were reconstructed from the distances using NEIGHBOR in PHYLIP. In the maximum parsimony (MP) analyses, the MP trees were searched by DNAPARS and PROTPARS in PHYLIP.

Support values for NJ, ML, and MP trees were obtained by bootstrapping (100 replicates for ML trees and 1000 replicates for NJ and MP trees) using SEQBOOT and CONSENSE in PHYLIP. In the bootstrap analyses of SSU rRNA, the global rearrangements option was employed and the input order was jumbled once. Alternative topologies were compared using the WSH (Shimodaira and Hasegawa 1999) and AU (Shimodaira 2002) tests with the program CONSEL (Shimodaira and Hasegawa 2001).

In addition to the combined analyses using the concatenate model, we performed the combined analysis of the four genes, SSU rRNA, α-tubulin, β-tubulin, and actin, based on the separate model. We computed log-likelihood values of 10,395 trees for eight eukaryotic groups (Amoebozoa, Cercozoa, Cryptomonada, Rhodophyta, Viridiplantae, Euglenozoa, Alveolata + stramenopiles, and Centrohelida) and summed the likelihood values of these genes for each tree, selecting the best tree with the highest log-likelihood value in total. The relationship among each eukaryotic group for each gene was constrained in advance according to the best tree of the ML analysis for each gene. Constraints were made based also on the generally accepted findings of the eukaryotic phylogeny (Table 1).

Results

SSU rRNA Phylogeny

The best tree of the ML analysis presented in an unrooted format is shown in Fig. 1. All major eukaryotic groups were recovered in the tree, except for the Alveolata. Euglenozoa appeared as sisters to *Plasmodium falciparum*, probably because of an artifact by extremely fast rate of evolution. Cercozoa, Radiolaria, and two groups of heliozoans (Desmothoracida and Taxopodida) formed the clade Rhizaria, and Rhizaria themselves were placed as sisters to the stramenopiles clade. *Raphidiophrys contractilis* branched within the Centrohelida clade, and their monophyly was supported with 95–100% BP support. The unclassified centroheliozoan "marine microheliozoan TCS-2002" was weakly associated with the Rhodophyta rather than with the Centrohelida. However, the monophyly of the marine microheliozoan TCS-2002 with the Centrohelida was also very likely at either of the two different positions of the backbone tree. The Centrohelida branched as a sister group to the Glaucophyta and Cryptomonada, and they formed a clade with the Rhodophyta and Viridiplantae. However, no significant BP support was obtained for the branching order of these groups.

To evaluate the significance of the position of the Centrohelida, it was moved to possible branching

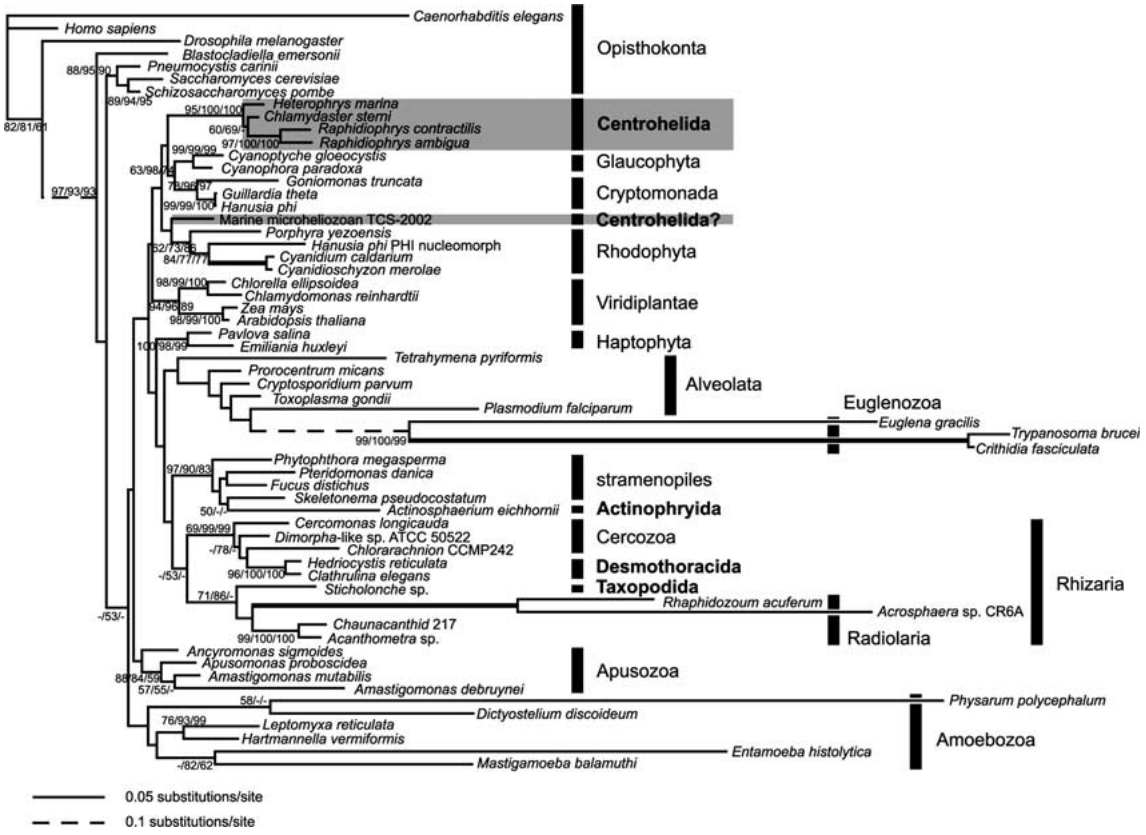


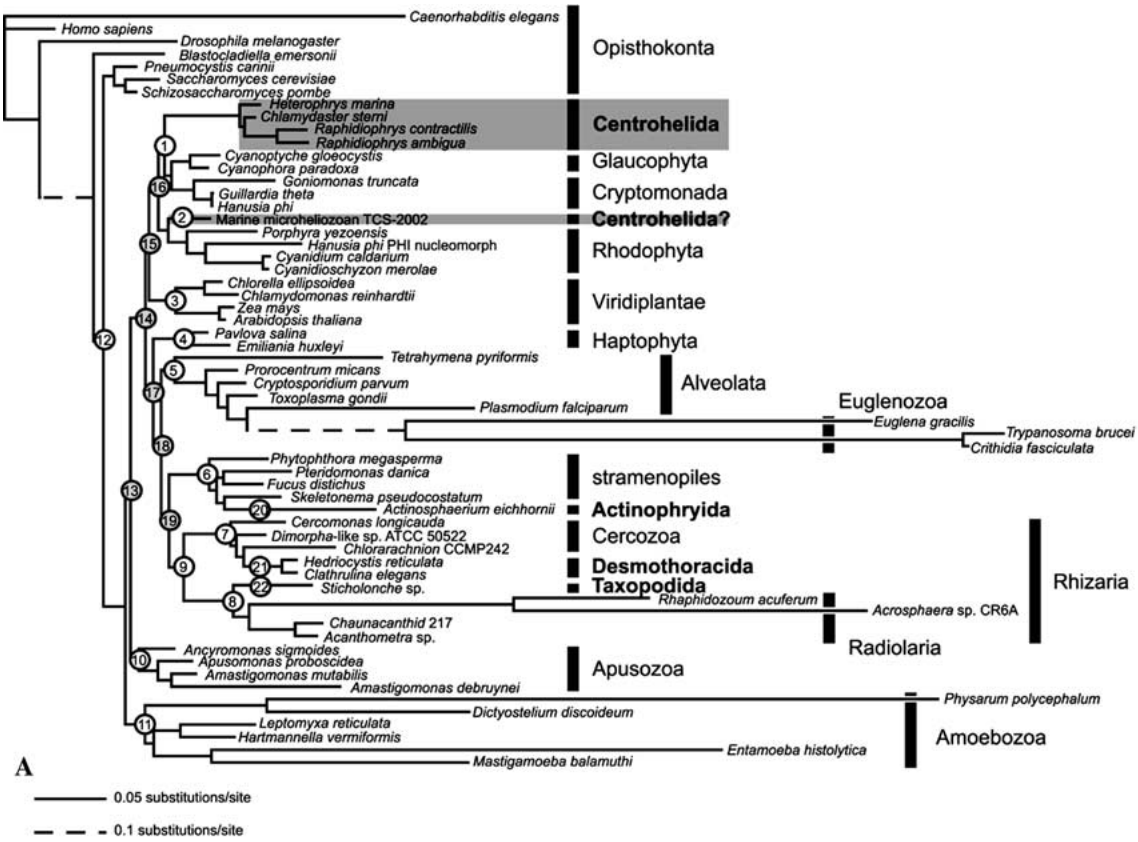
Fig. 1. Unrooted tree based on SSU rRNA sequences from 60 eukaryotes inferred from a ML analysis using 1251 unambiguously aligned positions. Numbers at internal nodes represent BP support values of ML, NJ, and MP analyses after 100 replicates, respectively. Tick nodes indicate that support values of three analyses are all 100%. Dashes indicate values under 50%, and support values that are all under 50% are omitted. NJ and MP analyses did not reveal the same topology as the best tree. The scale bars indicate a distance of 0.05 and 0.1 substitutions per sites, respectively.

positions in the backbone tree in Fig. 1. WSH and AU tests were used for excluding alternative trees, the log-likelihoods of which are significantly different from that of the best tree (Fig. 2). The WSH tests rejected the trees in which Centrohelida were sisters to Actinophryida (node 20; $p < 0.001$), to Desmothoracida (node 21; $p < 0.001$), or to Taxopodida (node 22; $p < 0.001$), while other alternatives were not significantly different from the best tree. The AU tests further rejected ($p < 0.05$) the trees in which Centrohelida were constrained as sisters to Viridiplantae (node 3), Cercozoa + Desmothoracida (node 7), Taxopodida + Radiolaria (node 8), Apusozoa (node 10), Opisthokonta (node 12), or other heliozoan groups (nodes 20–22). Other alternative trees in which Centrohelida were placed at the nodes branching before and after Apusozoa were also rejected by AU tests (nodes 13 and 14; $p < 0.05$).

α- + β-*Tubulin Phylogeny with the Concatenate Model*

In the α- + β-tubulin phylogeny (Fig. 3A), most of the major groups were recovered in agreement with other combined tubulin phylogenies (Edgcomb et al. 2001; Simpson et al. 2002). Diplomonadida and Parabasalia were placed between Opisthokonta and Amoebozoa, demonstrating the typical feature of the tubulin phylogeny (Arisue et al. 2005). Centrohelida (*R. contractilis*) branched together with the Rhodophyta (Cyanidiales) with 71–83% BP support in different phylogenetic methods. Although the position of Amoebozoa in the eukaryote tree is still ambiguous, the clade of Centrohelida + Rhodophyta was positioned basally within the bikonts clade.

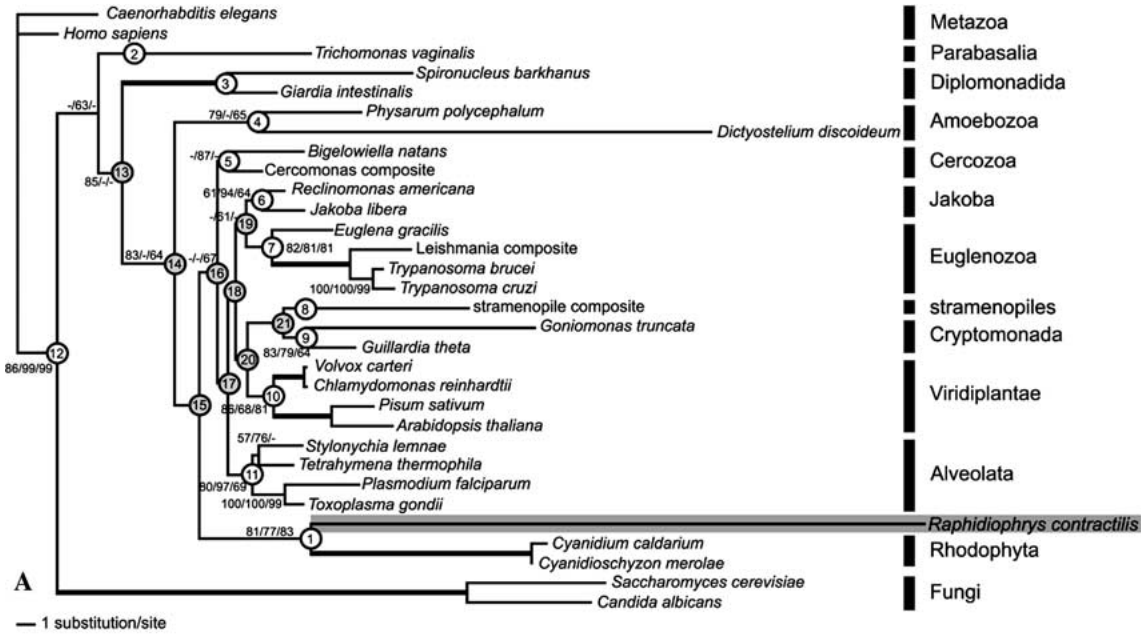
Comparison among alternative trees was performed by WSH and AU tests (Fig. 3B). Only two trees were not rejected by the AU test ($p > 0.05$): the tree in which Centrohelida were constrained to be sisters to Amoebozoa (node 4) and the tree in which Centrohelida branched with stramenopile composite + Cryptomonada (node 21). However, in the WSH test, trees in which Centrohelida were constrained as sisters to Amoebozoa (node 4), Cercozoa (node 5), Cryptomonada (node 9), or Viridiplantae (node 10) were not rejected ($p > 0.05$), while other trees in which Centrohelida were placed at the nodes branching before or after Amoebozoa (nodes 14 and 15), branching after Alveolata (node 18), or united with Jakobids + Euglenozoa (node 19), with stramenopile composite +



Node	Position of Centrohelida	Δli	*p*-value WSH	*p*-value AU
1	sisters to Glaucophyta + Cryptomonada (best tree)	(-23821.9)	0.990	0.740
2	sisters to Marine microheliozoan TCS-2002	3.5	0.744	0.419
3	sisters to Viridiplantae	11.7	0.227	0.024*
4	sisters to Haptophyta	4.3	0.929	0.477
5	sisters to Alveolata + Euglenozoa	6.2	0.832	0.309
6	sisters to stramenopiles	11.4	0.519	0.127
7	sisters to Cercozoa + Desmothoracida	21.2	0.097	0.010*
8	sisters to Taxopodida + Radiolaria	21.2	0.097	0.010*
9	sisters to Rhizaria	12.1	0.406	0.078
10	sisters to Apusozoa	20.5	0.124	0.002**
11	sisters to Amoebozoa	13.4	0.426	0.120
12	sisters to Opisthokonta	16.3	0.304	0.036*
13	branch after Amoebozoa	16.3	0.307	0.037*
14	branch after Apusozoa	11.7	0.390	0.025*
15	unite Glaucophyta + Cryptomonada + Rhodophyta + Viridiplantae	10.4	0.419	0.064
16	unite Glaucophyta + Cryptomonada + Rhodophyta	1.0	0.870	0.423
17	unite Haptophyta + Alveolata + Euglenozoa + stramenopiles + Rhizaria	4.3	0.944	0.480
18	unite Alveolata + Euglenozoa + stramenopiles + Rhizaria	4.4	0.963	0.713
19	unite stramenopiles + Rhizaria	6.2	0.832	0.305
20	sisters to Actinophryida (Heliozoa)	70.6	0.001**	0.000**
21	sisters to Desmothoracida (Heliozoa)	101.2	0.000**	0.018*
22	sisters to Taxopodida (Heliozoa)	63.3	0.000**	0.004**

B

Fig. 2. Unrooted tree based on SSU rRNA sequences and the comparison of alternative topologies by WSH and AU tests. **A** The SSU rRNA ML tree as shown in Fig. 1. Circles indicate the alternative branching positions of Centrohelida as shown in B. The numbers in the white circles indicate the sister relationships between Centrohelida and other eukaryotic groups, and others within the gray circles indicate the relocation of Centrohelida to the internal branching positions. The scale bars indicate a distance of 0.05 and 0.1 substitution per sites, respectively. **B** Comparison of alternative topologies for the relationship between Centrohelida and other eukaryotic groups. Δli indicates log-likelihood differences of alternative topologies from that of the best ML tree (A). *p*-values of the WSH and AU tests were estimated by CONSEL (Shimodaira and Hasegawa 2001). Topologies rejected by each test at the *5% and **1% levels, respectively.



Node	Position of Centrohelida	Δ*li*	*p*-value WSH	*p*-value AU
1	sisters to Rhodophyta (best tree)	(-11819.1)	0.989	0.941
2	sisters to Parabasalia	52.0	0.001**	0.001**
3	sisters to Diplomonadida	47.9	0.003**	0.002**
4	sisters to Amoebozoa	13.0	0.379	0.126
5	sisters to Cercozoa	23.0	0.061	0.017*
6	sisters to Jakoba	26.8	0.043*	0.007**
7	sisters to Euglenozoa	27.8	0.040*	0.007**
8	sisters to stramenopiles	28.5	0.048*	0.013*
9	sisters to Cryptomonada	28.8	0.051	0.008**
10	sisters to Viridiplantae	24.1	0.095	0.031*
11	sisters to Alveolata	26.2	0.049*	0.016*
12	sisters to Opisthokonta	52.6	0.001**	0.000**
13	branch before Diplomonadida	47.8	0.002**	0.002**
14	branch before Amoebozoa	20.6	0.096	0.014*
15	branch before Rhodophyta	18.3	0.112	0.023*
16	branch before Cercozoa	21.0	0.038*	0.003**
17	branch before Alveolata	25.0	0.039*	0.022*
18	branch after Alveolata	25.0	0.064	0.005**
19	unite Jakoba + Euglenozoa	25.0	0.064	0.006**
20	unite stramenopiles + Cryptomonada+ Viridiplantae	23.7	0.080	0.040*
21	unite stramenopiles + Cryptomonada	23.8	0.095	0.050

B

Fig. 3. Unrooted tree based on concatenate α-tubulin + β-tubulin amino acid sequences and comparison of alternative topologies by WSH and AU tests. **A** Unrooted tree based on concatenate α-tubulin + β-tubulin amino acid sequences from 31 representatives of eukaryotes inferred from a ML analysis using 771 unambiguously aligned positions. Numbers at internal nodes represent BP support values of ML, NJ, and MP analyses after 100 replicates, respectively. Tick nodes indicate that support values of three analyses are all 100%. Dashes indicate values under 50%, and support values that are all under 50% are omitted. NJ and MP analyses did not reveal the same topology as the best tree. Circles indicate the alternative branching positions of Centrohelida as shown in B. The numbers in the white circles indicate the sister relationships between Centrohelida and other eukaryotic groups, and others within the gray circles indicate the relocation of Centrohelida to the internal branching positions. The scale bar indicates a distance of 1 substitution per site. **B** Comparison of alternative topologies for the relationship between Centrohelida and other eukaryotic groups. Δ*li* indicates log-likelihood differences of alternative topologies from that of the best ML tree (A). *p*-values of the WSH and AU tests were estimated by CONSEL (Shimodaira and Hasegawa 2001). Topologies rejected by each test at the *5% and **1% levels, respectively.

Cryptomonada + Viridiplantae (node 20), or with stramenopile composite + Cryptomonada (node 21) were not rejected either ($p > 0.05$).

Tubulins + Actin Phylogeny with the Concatenate Model

In the tubulins + actin tree (Fig. 4A), the Centrohelida (Raphidiophrys composite) and Rhodophyta formed a clade with 61–88% BP support.

The alternative tree in which the Centrohelida were sisters to Amoebozoa (node 3) was not rejected, while all other trees were rejected by AU test ($p < 0.05$) (Fig. 4B). However, by WSH tests, alternative trees in which the Centrohelida were sisters to Amoebozoa (node 3) or Viridiplantae (node 6) were not rejected ($p > 0.05$), and other trees in which the Centrohelida were placed at the node branching before or after Amoebozoa (nodes 11 and 12) or Viridiplantae (nodes 16 and 17) were not rejected either ($p > 0.05$).

Combined SSU rRNA and Protein Phylogeny with the Separate Model

To further evaluate the phylogenetic position of the Centrohelida among eukaryotes, the combined analysis of the SSU rRNA, α-tubulin, β-tubulin, and actin genes was performed with the separate model. Because of the limitation of our computational facility, Opisthokonta and Parabasalia + Diplomonadida were excluded from the analysis. The best tree was (Amoebozoa, (Centrohelida, Rhodophyta), ((Alveolata + stramenopiles, Cercozoa), (Cryptomonada, (Viridiplantae, Euglenozoa)))), again suggesting the relationship between the Centrohelida and the Rhodophyta as shown in the concatenate protein phylogeny (Fig. 4A). However, the branching order of other major eukaryotic groups was unclear.

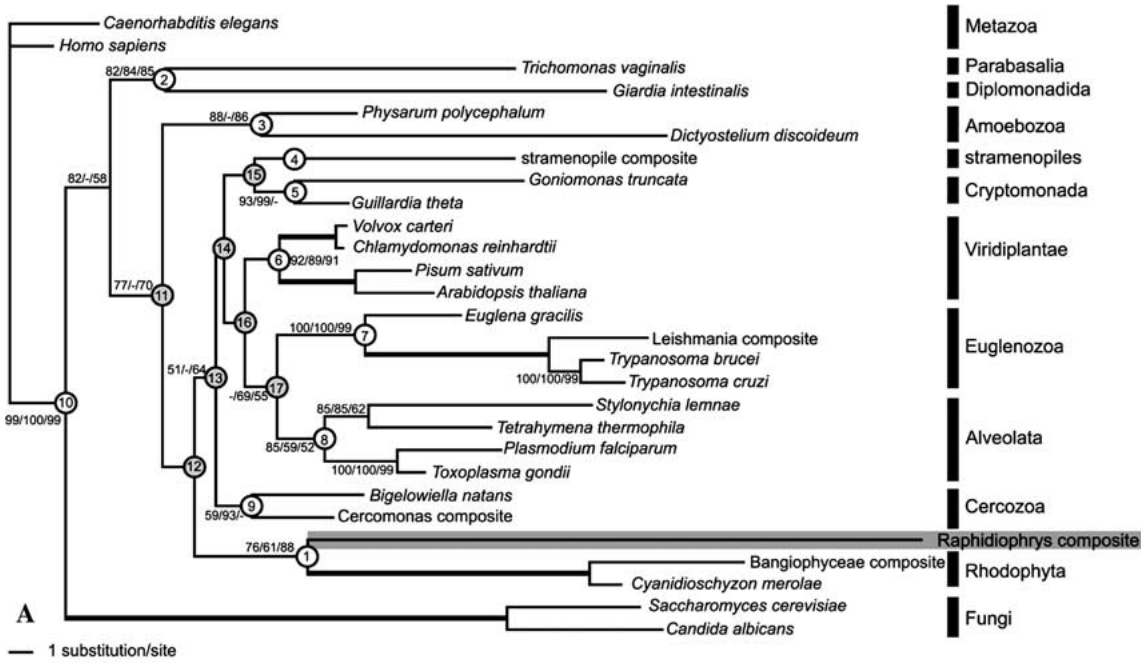
Discussion

Molecular phylogenetic analyses performed for the last couple of years have succeeded in placing various amoeboid protists in the global eukaryotic tree. For some amoeboid members, however, previous studies have failed to point to their accurate phylogenetic positions. The centroheliozoan protists are one of these. Probably, these protists are deep branches of eukaryotes, and available data did not provide information to resolve their position within the global tree. In this study, in order to obtain more reliable views on the phylogenetic position of the centroheliozoa and their relationships with other groups of eukaryotes, we sequenced SSU rRNA, α-tubulin, and β-tubulin genes of *R. contractilis* and performed detailed phylogenetic analyses on the basis of these genes and the actin gene of other species of the same genus, *Raphidiophrys ambigua*. These trials provided better answers for the above-mentioned questions on the centroheliozoa.

The best tree of the ML analysis for SSU rRNA indicated a sister relationship between the Centrohelida and the Glaucophyta + Cryptomonada cluster, though its BP support was weak. This relationship was also suggested in the SSU rRNA tree by Nikolaev et al. (2004). In our tree, moreover, these three eukaryotic groups branched with the Rhodophyta, and the common ancestor of these groups was located as sisters to Viridiplantae, forming a large clade, though, once again, their BP support was weak. All members of this clade have flat mitochondrial cristae, and both Rhodophyta and Centrohelida completely lack flagella. At present, phylogenetic relationships among major eukaryotic groups including these have not been resolved in SSU rRNA analysis but should be gauged by adding more gene sequences as suggested by Nikolaev et al. (2004).

The microheliozoan TCS-2002, deposited in GenBank by the name of unclassified Centrohelida, did not branch with the Centrohelida in the best tree but appeared as a sister group of the Rhodophyta. Cavalier-Smith and Chao (2003a) mentioned, based on preliminary ultrastructural examination, that this microheliozoan is not a centrohelid. If more detailed morphological analyses show that this protist is actually a centrohelid or closely related to Centrohelida, the node 2 in Fig. 2B with monophyly of the Centrohelida and the microheliozoa might be more likely than the best tree. The log-likelihood difference between these two trees was only 3.5, and the difference was not significant. The tree (node 2 in Fig. 2B) is reconciled with the combined protein trees which reconstruct the monophyly of the Centrohelida with Rhophyta. The SSU rRNA phylogeny supported the fact that the Centrohelida is not closely related to other heliozoans, being congruent with morphological data (Febvre-Chevalier 1990) and the previous SSU rRNA tree (Nikolaev et al. 2004). This study also supports the fact that axopodia possess an analogous structure resulting from convergent evolution.

Combined analyses of three different combinations of the genes, tubulins, tubulins + actin, and tubulins + actin + SSU rRNA, consistently suggested that Centrohelida are likely the sisters of the Rhodophyta. However, on the other hand, the alternative trees were not significantly rejected in the combined protein phylogenies (Figs. 3 and 4). The Rhodophyta have no flagella and possess flattened mitochondrial cristae like Centrohelida. The Amoebozoa, other possible sisters to Centrohelida shown by AU tests in Fig. 4B, use pseudopodia to move and engulf prey, but they have mitochondria with tubular cristae.



Node	Position of Centrohelida	Δ*li*	*p*-value WSH	*p*-value AU
1	sisters to Rhodophyta (best tree)	(-16311.0)	0.984	0.905
2	sisters to Parabasalia + Diplomonadida	52.2	0.008**	0.000**
3	sisters to Amoebozoa	13.6	0.434	0.163
4	sisters to stramenopiles	35.5	0.014*	0.001**
5	sisters to Cryptomonada	36.9	0.011*	0.001**
6	sisters to Viridiplantae	27.6	0.074	0.026*
7	sisters to Euglenozoa	37.0	0.021*	0.004**
8	sisters to Alveolata	37.2	0.026*	0.004**
9	sisters to Cercozoa	31.6	0.017*	0.007**
10	sisters to Opisthokonta	51.8	0.005**	0.000**
11	branch before Amoebozoa	22.1	0.155	0.004**
12	branch before Rhodophyta	18.3	0.220	0.034*
13	branch before Cercozoa	24.4	0.025*	0.012*
14	branch after Cercozoa	30.5	0.022*	0.003**
15	unite stramenopiles + Cryptomonada	30.9	0.032*	0.005**
16	unite Viridiplantae + Euglenozoa + Alveolata	27.7	0.059	0.013*
17	unite Euglenozoa + Alveolata	28.9	0.073	0.012*

B

Fig. 4. Unrooted tree based on concatenate tubulins + actin amino acid sequences and comparison of alternative topologies by WSH and AU tests. **A** Unrooted tree based on concatenate α-tubulin + β-tubulin + actin amino acid sequences from 28 representatives of eukaryotes inferred from a ML analysis using 1028 unambiguously aligned positions. Numbers at internal nodes represent BP support values of ML, NJ, and MP analyses after 100 replicates, respectively. Tick nodes indicate that support values of three analyses are all 100%. Dashes indicate values under 50%, and support values that are all under 50% are omitted. NJ and MP analyses did not reveal the same topology as the best tree. Circles indicate the alternative branching positions of Centrohelida as shown in B. The numbers in the white circles indicate the sister relationships between Centrohelida and other eukaryotic groups, and others within the gray circles indicate the relocation of Centrohelida to the internal branching positions. The scale bar indicates a distance of 1 substitution per site. **B** Comparison of alternative topologies for the relationship between Centrohelida and other eukaryotic groups. Δ*li* indicates log-likelihood differences of alternative topologies from that of the best ML tree (A). *p*-values of the WSH and AU tests were estimated by CONSEL (Shimodaira and Hasegawa 2001). Topologies rejected by each test at the *5% and **1% levels, respectively.

Though there are few common ultrastructural features among these eukaryotic groups, comparison of these trees with respective *p*-values suggested with some confidence that the Centrohelida could be positioned to a deeply branched sister of the Rhodophyta or the Amoebozoa, or to an independent branch between the branchings of Amoebozoa and Rhodophyta.

The monophyly of Rhodophyta and Viridiplantae has been established by using concatenate mitochondrial gene (Burger et al. 1999) and nuclear gene (Moreira et al. 2000) data sets. In contrast, Nozaki et al. (2003) have shown by analysis of concatenate protein data sets that the Rhodophyta are placed at the basal position within the bikonts and do not form a monophyletic group with the Viridiplantae. Their conclusion was based mainly on the data set including EF1α, α- and β-tubulin, and actin. Also, in our phylogenetic analyses of concatenate tubulin and actin sequences, the monophyly of the Rhodophyta and Viridiplantae was not reconstructed in the best tree. However, some constrained trees in which the Centrohelida + Rhodophyta cluster is a sister to Viridiplantae and the clade is located at the base of bikonts were not rejected by WSH tests ($p > 0.05$), and AU tests also did not reject ($p > 0.05$) the constrained trees when the clade uniting the Centrohelida + Rhodophyta and Viridiplantae was located at the position that is different from the earliest branch of the bikonts clade (data not shown). These results might be caused by the weak signal that unites Rhodophyta with Viridiplantae as suggested by mitochondrial and other nuclear genes (Burger et al. 1999; Moreira et al. 2000).

Some eukaryotes lacking flagella or those being temporarily flagellated such as Fungi, Amoebozoa, and Rhodophyta have highly divergent tubulin genes. Branches of the tubulin trees in these groups are sometimes very long so that the groups seem to be affected by long branch attraction (LBA) artifacts. As our protein trees show that the Centrohelida, also with a long branch, are sisters to Rhodophyta, we cannot exclude the possibility that the sister relationship between them was caused by LBA effects. However, the ML tree excluding *Dictyostelium* and Rhodophyta, both with long branches, revealed that the Centrohelida were not attracted to the Parabasalia + Diplomonadida clade with long branches but were located at the sisters to *Physarum* (data not shown). The branch length of *Physarum* is much shorter than those of *Dictyostelium* and Rhodophyta. Therefore, if Centrohelida were correctly located at the second most likely position, the sisters to the Amoebozoa, in the absence of Rhodophyta, then we could exclude the possibility that Centrohelida was randomly attracted to be located at other long branches in an artificial fashion. As there are few nuclear gene data for Rhodophyta at present, we have analyzed the tubulin and actin phylogeny using the Bangiophyceae (Cyanidiales and *Porphyra*) as representatives of the Rhodophyta. If the sequences of other red algae that represent short branches become available, we can reexamine the relationship between the Centrohelida and the Rhodophyta more precisely.

Though the position of Amoebozoa is obscure at present, it has been argued that the possible position of the root of eukaryotic tree lies between the Opisthokonta and the bikonts based on multiple-gene analyses (Arisue et al. 2005) and the distribution of fused DHFR-TS genes in the bikonts (Simpson and Roger 2002; Stechmann and Cavalier-Smith 2002, 2003; Baldauf 2003). Stechmann and Cavalier-Smith (2002) reported DHFR-TS gene fusion in the Centrohelida. It is likely that the Centrohelida branched after Amoebozoa in the eukaryotic tree and are placed at the basal position within the bikonts clade. It is also likely that the Rhodophyta (or possibly Plantae) branched after the Centrohelida, as Centrohelida may have some relationship to the Rhodophyta and Viridiplantae.

In our concatenate protein phylogenies, some eukaryotic groups such as Haptophyta and Rhizaria (excluding Cercozoa) were not included due to the lack of their protein sequence data. As mentioned by Cavalier-Smith and Chao (2003a), *Ancyromonas*, a member of Apusozoa, has mitochondria with flat cristae and kinetocysts like Centrohelida. Though the sister relationship between the Centrohelida and the Apusozoa was not supported in the SSU rRNA tree, we cannot exclude the possible relationship between the two groups because of sharing ultrastructural features. More molecular data will be needed from the Centrohelida and other eukaryotes to resolve the branching order and clarify the precise evolutionary position of the Centrohelida in the big picture of eukaryote phylogeny.

Acknowledgments. M.S. was supported by a JSPS Research Fellowship for Young Scientists. This study was supported in part by JSPS Grant JSPS RFTF 00L0162. We would like to thank Dr. T. Suzaki (Kobe University) for generously providing samples used in this study.

References

Arisue N, Hashimoto T, Lee JA, Moore DV, Gordon P, Sensen CW, Gaasterland T, Hasegawa M, Müller M (2002a) The phylogenetic position of the pelobiont *Mastigamoeba balamuthi* based on sequences of rDNA and translation elongation factors EF-1α and EF-2. J Eukaryot Microbiol 49:1–10

Arisue N, Hashimoto T, Yoshikawa H, Nakamura Y, Nakamura G, Nakamura F, Yano TA, Hasegawa M (2002b) Phylogenetic position of *Blastocystis hominis* and of stramenopiles inferred from multiple molecular sequence data. J Eukaryot Microbiol 49:42–53

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

Baldauf SL (2003) The deep roots of eukaryotes. Science 300:1703–1706

Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittile WF (2000) A kingdom-level phylogeny of eukaryotes based on combined protein data. Science 290:972–977

Bapteste E, Brinkmann H, Lee JA, Moore DV, Sensen CW, Gordon P, Duruflé L, Gaasterland T, Lopez P, Müller M,

Philippe H (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

Berney C, Pawlowski J (2003) Revised small subunit rRNA analysis provides further evidence that Foraminifera are related to Cercozoa. J Mol Evol 57:S120–S127

Burger G, Saint-Louis D, Gray MW, Lang BF (1999) Complete sequence of the mitochondrial DNA of the red alga *Porphyra purpurea*: Cyanobacterial introns and shared ancestry of red and green algae. Plant Cell 11:1675–1694

Cavalier-Smith T (1998) A revised six-kingdom system of life. Biol Rev Cambr Philos Soc 73:203–266

Cavalier-Smith T (2004) Only six kingdoms of life. Proc R Soc Lond B 271:1251–1262

Cavalier-Smith T, Chao EE (2003a) Molecular phylogeny of centrohelid Heliozoa, a novel lineage of bikont eukaryotes that arose by ciliary loss. J Mol Evol 56:387–396

Cavalier-Smith T, Chao EE (2003b) Phylogeny and classification of phylum Cercozoa (Protozoa). Protist 154:341–358

Edgcomb VP, Roger AJ, Simpson AGB, Kysela DT, Sogin ML (2001) Evoltionary relationships among ''jakobid'' flagellates as indicated by alpha- and beta-tubulin phylogenies. Mol Biol Evol 18:514–522

Fahrni JF, Bolivar I, Berney C, Nassonova E, Smirnov A, Pawlowski J (2003) Phylogeny of lobose amoebae based on actin and small-subunit ribosomal RNA genes. Mol Biol Evol 20:1881–1886

Febvre Chevalier C (1990) Phylum Actinopoda: Class Heliozoa. In: Margulis L, Corliss JO, Melkonian M, Chapman DJ (eds) Handbook of Protoctista. Jones and Bartlett, Boston, pp 347–362

Felsenstein J (2002) PHYLIP (phylogeny inference package), version 3.6a. Distributed by the author, University of Washington, Seattle

Haeckel E (1887) Report on radiolaria collected by HMS Challenger during the years 1873–1876. Rep Sci ResVoyage HMS Challenger 1873–1876 18:1–1803

Hasegawa M, Kishino H, Yano T (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282

Keeling PJ (2001) Foraminifera and Cercozoa are related in actin phylogeny: Two orphans find a home? Mol Biol Evol 18:1551–1557

Knoll AH (1992) The early evolution of eukaryotes: a geological perspective. Science 256:622–627

Leander BS, Clopton RE, Keeling PJ (2003) Phylogeny of gregarines (Apicomplexa) as inferred from small-subunit rDNA and β-tubulin. Int J Syst Evol Microbiol 53:345–354

Lee JJ, Leedale GF, Bradbury P (2000) The illustrated guide to the Protozoa, 2nd ed. Society of Protozoologists, Lawrence, KS

Levine ND, Corliss JO, Cox FEG, Deroux G, Grain J, Honigberg BM, Leedale GF, Loeblich III AR, Lom J, Lynn D, Merinfeld EG, Page FC, Poljansky G, Sprague V, Vavra J, Wallace FG (1980) A newly revised classification of the Protozoa. J Protozool 27:37–58

Longet D, Archibald JM, Keeling PJ, Pawlowski J (2003) Foraminifera and Cercozoa share a common origin according to RNA polymerase II phylogenies. Int J Syst Evol Microbiol 53:1735–1739

López-García P, Rodríguez-Valera F, Moreira D (2002) Toward the monophyly of Haeckel's Radiolaria: 18S rRNA environmental data support the sisterhood of Polycystinea and Acantharea. Mol Biol Evol 19:118–121

Mikrjukow KA, Siemensma FJ, Patterson DJ (2000) Phylum Heliozoa. In: Lee JJ, Leedale GF, Bradbury P (eds) The illustrated guide to the Protozoa. 2nd ed. Society of Protozoologists, Lawrence, KS, pp 860–871

Moreira D, Le Guyader H, Philippe H (2000) The origin of red algae and the evolution of chloroplasts. Nature 405:69–72

Nakayama T, Watanabe S, Mitsui K, Uchida H, Inouye I (1996) The phylogenetic relationship between the Chlamydomonadales and Chlorococcales inferred from 18SrRNA sequence data. Phycol Res 44:47–55

Nikolaev SI, Berney C, Fahrni JF, Bolivar I, Polet S, Mylnikov AP, Aleshin VV, Petrov NB, Pawlowski J (2004) The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. Proc Natl Acad Sci USA 101:8066–8071

Nozaki H, Matsuzaki M, Takahara M, Misumi O, Kuroiwa H, Hasegawa M, Shin-I T, Kohara Y, Ogasawara N, Kuroiwa T (2003) The phylogenetic position of red algae revealed by multiple nuclear genes from mitochondria-containing eukaryotes and an alternative hypothesis on the origin of plastids. J Mol Evol 56:485–497

Polet S, Berney C, Fahrni J, Pawlowski J (2004) Small-subunit ribosomal RNA gene sequences of Phaeodarea challenge the monophyly of Haeckel's Radiolaria. Protist 155:53–63

Pupko T, Huchon D, Cao Y, Okada N, Hasegawa M (2002) Combining multiple data sets in a likelihood analysis: which models are the best? Mol Biol Evol 19:2294–2307

Sakaguchi M, Suzaki T, Kahn SMMK, Hausmann K (2001) Food capture by kinetocysts in the heliozoon *Raphidiophrys contractilis*. Eur J Protistol 37:453–458

Sekiguchi H, Moriya M, Nakayama T, Inouye I (2002) Vestigial chloroplasts in heterotrophic stramenopiles *Pteridomonas danica* and *Ciliophrys infusionum* (Dictyochophyceae). Protist 153:157–167

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

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

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

Simpson AGB, Roger AJ (2002) Eukaryotic evolution: Getting to the root of the problem. Curr Biol 12:R691–R695

Simpson AGB, Roger AJ, Silberman JD, Leipe DD, Edgcomb VP, Jermiin LS, Patterson DJ, Sogin ML (2002) Evolutionary history of ''early-diverging'' eukaryotes: the excavate taxon *Carpediemonas* is a close relative of *Giardia*. Mol Biol Evol 19:1782–1791

Sogin ML (1991) Early evolution and the origin of eukaryotes. Curr Opin Genet Dev 1:457–463

Stechmann A, Cavalier-Smith T (2002) Rooting the eukaryote tree by using a derived gene fusion. Science 297:89–91

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

Wuyts J, Perriere G, Van de Peer Y (2004) The European ribosomal RNA database. Nucleic Acids Res 32:D101–D103

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 13:555–556

Yoon HS, Hackett JD, Pinto G, Bhattacharya D (2002) The single, ancient origin of chromist plastids. Proc Natl Acad Sci USA 99:15507–15512