

Evaluating Support for the Current Classification of Eukaryotic Diversity

Laura Wegener Parfrey¹, Erika Barbero², Elyse Lasser², Micah Dunthorn¹, Debashish Bhattacharya^{3,4}, David J. Patterson⁵, Laura A. Katz^{1,2*}

1 Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst, Massachusetts, United States of America, **2** Department of Biological Sciences, Smith College, Northampton, Massachusetts, United States of America, **3** Department of Biological Sciences, University of Iowa, Iowa City, Iowa, United States of America, **4** Roy J. Carver Center for Comparative Genomics, University of Iowa, Iowa City, Iowa, United States of America, **5** Bay Paul Center for Genomics, Marine Biological Laboratory, Woods Hole, Massachusetts, United States of America

Perspectives on the classification of eukaryotic diversity have changed rapidly in recent years, as the four eukaryotic groups within the five-kingdom classification—plants, animals, fungi, and protists—have been transformed through numerous permutations into the current system of six “supergroups.” The intent of the supergroup classification system is to unite microbial and macroscopic eukaryotes based on phylogenetic inference. This supergroup approach is increasing in popularity in the literature and is appearing in introductory biology textbooks. We evaluate the stability and support for the current six-supergroup classification of eukaryotes based on molecular genealogies. We assess three aspects of each supergroup: (1) the stability of its taxonomy, (2) the support for monophyly (single evolutionary origin) in molecular analyses targeting a supergroup, and (3) the support for monophyly when a supergroup is included as an out-group in phylogenetic studies targeting other taxa. Our analysis demonstrates that supergroup taxonomies are unstable and that support for groups varies tremendously, indicating that the current classification scheme of eukaryotes is likely premature. We highlight several trends contributing to the instability and discuss the requirements for establishing robust clades within the eukaryotic tree of life.

Citation: Parfrey LW, Barbero E, Lasser E, Dunthorn M, Bhattacharya D, et al. (2006) Evaluating support for the current classification of eukaryotic diversity. *PLoS Genet* 2(12): e220. doi:10.1371/journal.pgen.0020220

Introduction

Biological research is based on the shared history of living things. Taxonomy—the science of classifying organismal diversity—is the scaffold on which biological knowledge is assembled and integrated into a cohesive structure. A comprehensive eukaryotic taxonomy is a powerful research tool in evolutionary genetics, medicine, and many other fields. As the foundation of much subsequent research, the framework must, however, be robust. Here we test the existing framework by evaluating the support for and stability of the classification of eukaryotic diversity into six supergroups.

Eukaryotes (organisms containing nuclei) encompass incredible morphological diversity from picoplankton of only two microns in size to the blue whale and giant sequoia that are eight orders of magnitude larger. Many evolutionary innovations are found only in eukaryotes, some of which are present in all lineages (e.g., the cytoskeleton, nucleus) and others that are restricted to a few lineages (e.g., multicellularity, photosynthetic organelles [plastids]). These and other eukaryotic features evolved within microbial eukaryotes (protists) that thrived for hundreds of millions of years before they gave rise independently to multicellular eukaryotes, the familiar plants, animals, and fungi [1]. Thus, elucidating the origins of novel eukaryotic traits requires a comprehensive phylogeny—an inference of organismal relationships—that includes the diverse microbial lineages.

Higher-level classifications have historically emphasized the visible diversity of large eukaryotes, as reflected by the establishment of the plant, animal, and fungal kingdoms. In these schemes the diverse microbial eukaryotes have gen-

erally been placed in one (Protista [2–4] or Protoctista [5]) or two (Protozoa and Chromista [6]) groups (Figure 1; but see also [7,8]). However, this historic distinction between macroscopic and microscopic eukaryotes does not adequately capture their complex evolutionary relationships or the vast diversity within the microbial world.

In the past decade, the emphasis in high-level taxonomy has shifted away from the historic kingdoms and toward a new system of six supergroups that aims to portray evolutionary relationships between microbial and macrobial lineages. The supergroup concept is gaining popularity as evidenced by several reviews [9,10] and inclusion in forthcoming editions of introductory biology textbooks. In addition, the International Society of Protozoologists recently proposed a formal reclassification of eukaryotes into six supergroups, though acknowledging uncertainty in some groups [7].

The Supergroups

Below we introduce the six supergroups in alphabetical order (Figure 2). The supergroup “Amoebozoa” was proposed in 1996 [11]. Original evidence for the group was drawn from

Editor: David M. Hillis, University of Texas, United States of America

Received May 11, 2006; **Accepted** November 9, 2006; **Published** December 22, 2006

A previous version of this article appeared as an Early Online Release on November 13, 2006 (doi:10.1371/journal.pgen.0020220.eor).

Copyright: © 2006 Parfrey et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

* To whom correspondence should be addressed. E-mail: lkatz@smith.edu

Synopsis

Evolutionary perspectives, including the classification of living organisms, provide the unifying scaffold on which biological knowledge is assembled. Researchers in many areas of biology use evolutionary classifications (taxonomy) in many ways, including as a means for interpreting the origin of evolutionary innovations, as a framework for comparative genetics/genomics, and as the basis for drawing broad conclusions about the diversity of living organisms. Thus, it is essential that taxonomy be robust. Here the authors evaluate the stability of and support for the current classification system of eukaryotic cells (cells with nuclei) in which eukaryotes are divided into six kingdom level categories, or supergroups. These six supergroups unite diverse microbial and macrobial eukaryotic lineages, including the well-known groups of plants, animals, and fungi. The authors assess the stability of supergroup classifications through time and reveal a rapidly changing taxonomic landscape that is difficult to navigate for the specialist and generalist alike. Additionally, the authors find variable support for each of the supergroups in published analyses based on DNA sequence variation. The support for supergroups differs according to the taxonomic area under study and the origin of the genes (e.g., nuclear, plastid) used in the analysis. Encouragingly, combining a conservative approach to taxonomy with increased sampling of microbial eukaryotes and the use of multiple types of data is likely to produce a robust scaffold for the eukaryotic tree of life.

molecular genealogies and morphological characters such as eruptive pseudopodia and branched tubular mitochondrial cristae. However, no clear synapomorphy—shared derived character—exists for “Amoebozoa.” In fact, amoeboid organisms are not restricted to the “Amoebozoa,” but are found in at least four of the six supergroups.

The “Amoebozoa” include a diversity of predominantly amoeboid members such as *Dictyostelium discoideum* (cellular slime mold), which is a model for understanding multicellularity [12]. Another member, *Entamoeba histolytica*, is an amitochondriate amoeba (Pelobiont) and is the cause of amoebic dysentery, an intestinal infection with global health consequences [13].

“Chromalveolata” was introduced as a parsimonious, albeit controversial, explanation for the presence of plastids of red algal origin in photosynthetic members of the “Alveolata” and “Chromista” [14]. Under this hypothesis, the last common ancestor of the chromalveolates was a heterotroph that acquired photosynthesis by engulfing a red alga and retaining it as a plastid [15,16]. The “Alveolata” include ciliates, dinoflagellates, and apicomplexa, and its monophyly is well supported by morphology and molecules. “Chromista” was created as a kingdom to unite diverse microbial lineages with red algal plastids (and their nonphotosynthetic descendants) [6,17], but no clear synapomorphy unites this clade.

The supergroup “Chromalveolata” includes microbes with critical roles in the environment and in human health. Numerous key discoveries emerged from studies of the model organism *Tetrahymena* (ciliate: “Alveolata”), including self-splicing RNAs and the presence of telomeres [18]. *Phytophthora* (stramenopile: “Chromista”), a soil-dwelling organism, is the causative agent of the Irish Potato Famine [19], whereas *Plasmodium* (Apicomplexa: “Alveolata”) is the causative agent of malaria [20].

“Excavata” is a supergroup composed predominately of

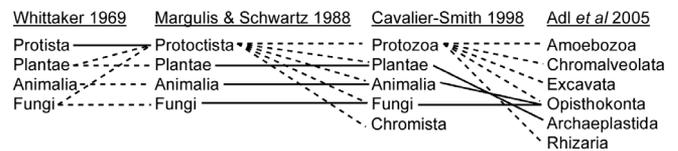


Figure 1. Trends in the Taxonomy of Eukaryotes

A comparison of four representative taxonomies illustrates trends within eukaryotic taxonomy over the past 50 years [2,5–7]. Movement of taxa is traced from earlier to more recent taxonomies with solid and dashed lines. A solid line indicates all members of a group (left of line) are incorporated into the subsequent group (right of line). Dashed lines indicate that a subset of members (left) is incorporated into subsequent groups (right).

doi:10.1371/journal.pgen.0020220.g001

heterotrophic flagellates whose ancestor is postulated to have had a synapomorphy of a conserved ventral feeding groove [21]. Most members of “Excavata” are free-living heterotrophs, but there are notable exceptions that are pathogens. For example, *Giardia* (Diplomonada) causes the intestinal infection giardiasis, and *Trichomonas vaginalis* (Parabasalida) is the causative agent of a sexually transmitted disease [22]. Kinetoplastids, such as *Trypanosoma* (Euglenozoa), have unique molecular features such as extensive RNA editing of mitochondrial genes that is templated by minicircle DNA [23].

“Opisthokonta” includes animals, fungi, and their microbial relatives. This supergroup emerged from molecular gene trees [24] and is united by the presence of a single posterior flagellum in many constituent lineages [25]. Molecular studies have expanded microbial membership of the group and revealed a potential molecular synapomorphy, an insertion in the Elongation Factor 1 α gene in lineages containing this ortholog [26,27].

“Opisthokonts” include many biological model organisms (*Drosophila*, *Saccharomyces*). Vast amounts of research have been conducted on members of this supergroup and much textbook science is based on inferences from these lineages. Other notable opisthokonts include *Encephalitozoon* (Microsporidia: Fungi), a causative agent of diarrhea, which has one of the smallest known nuclear genomes at 2.9 MB [28]. Also included within the “Opisthokonta” are the choanoflagellates (e.g., *Monosiga*), which are the sister to animals [29].

The supergroup “Plantae” was erected as a kingdom in 1981 [30] to unite the three lineages with primary plastids: green algae (including land plants), rhodophytes, and glaucophytes. Under this hypothesis a single ancestral primary endosymbiosis of a cyanobacterium gave rise to the plastid in this supergroup [31]. The term “Plantae” has been used to describe numerous subsets of photosynthetic organisms, but in this manuscript will only be used in reference to the supergroup.

Well-known “Plantae” genera include *Arabidopsis*, a model angiosperm, and *Porphyra* (red alga), the edible seaweed nori. Within the “Plantae” there have been numerous independent origins of multicellularity including: *Volvox* (Chlorophyta) [32], the land plants, and red algae.

“Rhizaria” emerged from molecular data in 2002 to unite a heterogeneous group of flagellates and amoebae including: cercozoans, foraminifera, diverse testate amoebae, and former members of the polyphyletic radiolaria [33]. “Rhizaria” is an expansion of the “Cercozoa” [6] that was also recognized from molecular data [34,35]. “Cercozoa” and

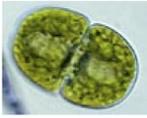
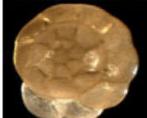
Supergroup	Groups	Representatives Genera	Assessment		
			Nuclear	Plastid	
"Amoebozoa"		Lobose amoebae	<i>Amoeba</i>	+	
		Slime molds	<i>Dictyostelium</i>		
		Pelobionts	<i>Entamoeba</i>		
"Chromalveolata"		Ciliates	<i>Tetrahymena</i>	-	+
		Stramenopiles	<i>Phytophthora</i>		
		Apicomplexa	<i>Plasmodium</i>		
"Excavata"		Diplomonads	<i>Giardia</i>	-	
		Euglenozoa	<i>Trypanosoma</i>		
		Parabasalids	<i>Trichomonas</i>		
"Opisthokonta"		Animals	<i>Drosophila</i>	+++	
		Fungi	<i>Encephalitozoon</i>		
		Choanoflagellates	<i>Monosiga</i>		
"Plantae"		Green algae	<i>Arabidopsis</i>	+	+++
		Red algae	<i>Porphyra</i>		
		Glaucophytes	<i>Cyanophora</i>		
"Rhizaria"		Cercomonads	<i>Cercomonas</i>	+	
		Foraminifera	<i>Allogromia</i>		
		Euglyphids	<i>Paulinella</i>		

Figure 2. Summary of Eukaryotic Supergroups

Assessment based on our analysis of molecular genealogies. +++, well supported; +, some support; -, support missing or very limited. Nuclear, genealogies based on nuclear genes. Plastid, genealogies based on chloroplast genes. Pictured organisms: *Lesquereusia*, *Thalassionema*, *Jakoba*, *Proterospongia*, *Cosmarium*, *Ammonia*. (Images: micro*scope, <http://starcentral.mbl.edu/microscope>). doi:10.1371/journal.pgen.0020220.g002

foraminifera appear to share a unique insertion in ubiquitin [36], but there is a paucity of non-molecular characters uniting members of "Rhizaria."

"Rhizaria" encompasses a diversity of forms, including a heterotrophic flagellate *Cercomonas* (Cercomonada: "Cercosoa") and a photosynthetic amoeba *Paulinella chromatophora*, (Silicoflosea: "Cercosoa"). The latter likely represents a recent endosymbiosis of a cyanobacterium [37,38]. Some members of the "Rhizaria," notably the shelled foraminifera, also have a substantial fossil record that can be used to determine the age of sediments [39].

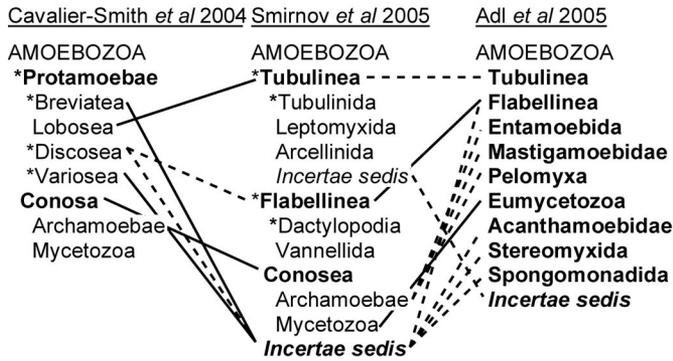
Our Approach

To assess the robustness of the six proposed supergroups, we compare formal taxonomies and track group composition and nomenclature across time (Figures 1 and 3). We also evaluate support for the six supergroups by analyzing published molecular genealogies that either target a specific supergroup or aim to survey all supergroups. Our focus on molecular genealogies is limited. We recognize that supergroups have, in many cases, been defined by suites of

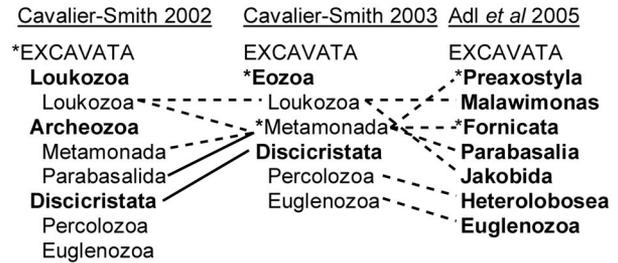
characters such as flagellar apparatus in "Excavata" [33,40] and "Opisthokonta"[25], and that groups are more robust when supported by multiple data types (see Discussion). Use of genealogies is further complicated because a genealogy is the reconstruction of the history of a gene, and may or may not be congruent with phylogenies, which depict the history of organisms [41,42]. Despite these factors, our treatment of molecular genealogies is warranted given the prevalence of molecular analyses in the literature that seeks support for supergroups and the reliance on these gene trees in establishing taxonomy.

For each genealogy we evaluate the taxon sampling for the targeted supergroup (Membership; Figures 4–9) and the monophyly of all supergroups with at least two member taxa (Supergroup monophyly; Figures 4–9). Monophyletic clades, those that include an ancestor and all of its descendants [43], are scored (+; Figures 4–9). We assess support for supergroups when they are targeted by specific studies and when they are included as out-groups in studies targeting other supergroups. A conservative measure of out-group monophyly was

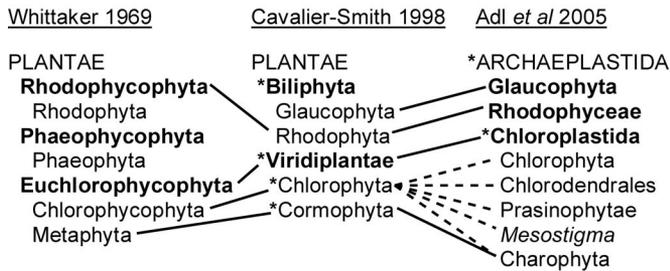
A “Amoebozoa”



B “Excavata”



C “Plantae”



D “Rhizaria”

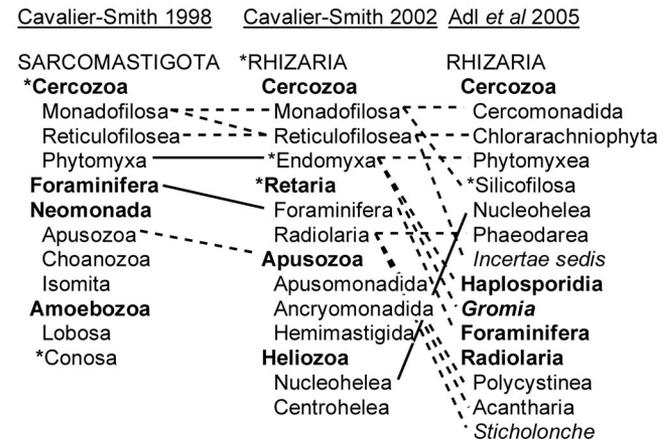


Figure 3. Trends in Supergroup Taxonomy

A comparison of three formal classifications illustrates trends within (A) “Amoebozoa” [7,45,47]; (B) “Excavata” [7,33,60]; (C) “Plantae” [2,6,7]; and (D) “Rhizaria” [6,7,33]. A majority of solid, horizontal lines would indicate temporal stability of supergroup classification. For visual simplicity we do not indicate groups newly included in the supergroups or taxonomic restructuring within subgroups. Asterisk indicates a newly introduced term. “Chromalveolata” and “Opisthokonta” are not included because only one formal taxonomy exists for both groups. See Figure 1 for further notes. doi:10.1371/journal.pgen.0020220.g003

Year	Gene	Ref	Membership										Supergroup Monophyly						
			My	Dc	Tu	Am	Fl	Pe	Ma	En	Rs	Br	A	C	E	O	P	R	
1995	SSU	[25]	○	○	○	○				○	○		—	—		+	+	—	
1997	SSU	[114]	○	○	○	○				○	○		—	—		+	—	—	
2001	SSU	[84]	●	●	●	●	●			●	●	●	○	+	—	—	+	—	—
2002	Multi	[97]		●						●	●		+	+	—	+	+	—	
2003	Actin	[48]	●	●	●	●	●	●	●	●	●		+	+	+	+	—	—	
2003	SSU	[48]	●	●	●	●	●	●	●	●	●	●	+	—	+	+	—	—	
2004	SSU	[45]	●	●	●	●	●	●	●	●	●	○	+	—	—	+	—	+	
2000	Multi	[115]	●	●		●							+	+	—	+	+	—	
2002	SSU	[33]	●	●	●	●				●	●		+	—	+	+	—	—	
2004	SSU	[49]	○	○	●	●	○	●	●	●	○	○	—	—	—	+	—	+	

Figure 4. Support for Membership and Supergroup Monophyly from “Amoebozoa”-Targeted Molecular Genealogies

Membership: ● indicates the member taxon falls within the supergroup Amoebozoa; ○ indicates that the member taxon is excluded from the Amoebozoa clade, or no clade is formed. Papers below blank line survey eukaryotic diversity [33,49,115] and are included in all analyses. Member taxa: My, Mycetozoa; Dc, Dictyosteliids; Tu, Tubulinea (*Lobosea*, *Gymnamoebia sensu stricto*); Am, Acanthamoebidae; Fl, Flabellina (*Discosea*, *Glycostylea*); Pe, *Pelomyxa*; Ma, Mastigamoebidae; En, Entamoebidae; Rs, residua; Br, *Breviata*, “*Mastigamoeba invertans sensu* NCBI.” **Supergroup Monophyly**, + indicates monophyly; — indicates group is para- or polyphyletic, and blank indicates insufficient data available. Supergroup definition based on Adl et al. 2005 [4]: A, Amoebozoa; C, Chromalveolata; E, Excavata; O, Opisthokonta; P, Plantae; R, Rhizaria. The position of *Breviata*, Br, was not considered when scoring the monophyly of Amoebozoa as this organism was misidentified and affiliations are unknown (see text). Some nodes were constrained in reference [97]. References cited in this figure are [25,33,45,48,49,84,97,114,115]. doi:10.1371/journal.pgen.0020220.g004

Year	Gene	Loc	Ref	Membership				Supergroup Monophyly					
				Al	St	Ha	Cr	A	C	E	O	P	R
1999	Multi	Pla	[116]	○	○	○		—					+
2002	Multi	Pla	[117]		●	●	●		+				+
2003	GAPDH	Pla	[118]	●	●	●	●		+				+
2004	Multi	Pla	[119]	●	●	●	○		—				+
2004	Multi	Pla	[119]	●	●	●	●		+				+
2004	FBA II	Pla	[120]	●	●	●	●		+				—
2005	Multi	Pla	[121]	●	●	●	●		+				+
2005	PsbA	Pla	[121]	●	●	●	○		—				+
2005	Multi	Pla	[107]	●	●	●	●		+				+
2003	GAPDH	Nuc	[118]	●	●	○	○		—	—	—	—	—
2003	Hsp90	Nuc	[122]	○	○		○		—	—	—	+	—
2005	Hsp90	Nuc	[56]	●	●	○	○		—	—	—	—	—
2005	Multi	Nuc	[56]	●	●	○	○		—	—	—	+	—
2005	β-tub	Nuc	[55]	●	●		○		—	—	—	—	—
2005	SSU	Nuc	[55]	○	○	○	○		—	—	—	—	—
2000	Multi	Nuc	[115]	●	●				+	+	—	+	+
2002	SSU	Nuc	[33]	●	●	○	○		+	—	+	+	—
2004	SSU	Nuc	[49]	○	○	○	○		—	—	—	+	—

Figure 5. Support for Membership and Supergroup Monophyly from “Chromalveolata”-Targeted Molecular Genealogies

Member taxa: Al, Alveolata; St, Stramenopiles (Heterokonts); Ha, Haptophyta; Cr, Cryptophyceae. Monophyletic “Plantae” from plastid genealogies includes secondarily derived plastids. See Figure 4 for further notes. References cited in this figure are [33,49,55,56,107,108,115–122]. Loc, location (genome) from which the gene of interest originated; Pla, plastid genome; Nuc, nuclear genome; Mit, mitochondrial genome. doi:10.1371/journal.pgen.0020220.g005

used because we required only two member lineages be present. In contrast, focal supergroups had broader taxonomic sampling.

Results

Taxonomic Instability

There is considerable instability in taxonomies of the six putative supergroups (Figure 3). Causes of the rapid revisions

in eukaryotic taxonomy over short time periods include: (1) nomenclatural ambiguity, (2) ephemeral and poorly supported higher-level taxa, and (3) classification schemes erected under differing taxonomic philosophies. For example, taxonomy of the “Amoebozoa,” a term originally introduced by Lühe in 1913 [44] to encompass a very different assemblage of organisms, has changed considerably in ten years (Figure 3A). “Varioseae” was created as a subclade within the “Amoebozoa” in 2004 to group taxonomically unplaced genera of amoebae with “exceptionally varied phenotype” [45]. Rarely supported by morphology or molecular evidence [46–49], this taxon was excluded from subsequent classifications [7,47] but is still discussed in the literature [46]. Similarly, the excavate taxon “Loukozoa” [6] has been continually redefined to include a variety of taxa bearing a ventral groove (Figure 3B) and finally abandoned [40]. The taxonomy of “Rhizaria” has emerged largely from molecular genealogies and has varied partly in response to shifting topology of gene trees that change with taxon sampling and the method of tree construction [6,33,50,51] (Figure 3D).

The taxonomy of “Plantae” is destabilized by the complex history of the term. Used since Haeckel’s time [52], “Plantae” has been redefined numerous times to describe various collections of photosynthetic organisms, leading to major discrepancies between taxonomic schemes (Figure 3C; e.g., [2,5]). The term “Archaeplastida” was recently introduced to alleviate confusion over “Plantae,” but this synonym is not widely used.

The stability of two supergroups, “Chromalveolata” and “Opisthokonta,” cannot be assessed at this time because only a single formal taxonomy exists [7]. Other classification schemes of eukaryotes segregate animals and fungi as separate kingdoms and place microbial opisthokonts in the kingdom Protozoa (Figure 1) [6,33]. Similarly, chromalveolate members are often divided between the polyphyletic kingdoms “Chromista” and “Protozoa” (Figure 1) [33,49].

Year	Gene Ref	Membership										Supergroup Monophyly						
		Di	Rt	Cp	Tr	Ox	Ht	Eu	MI	Jk	Pa	Dy	A	C	E	O	P	R
2001	SSU [123]	○			■	■	○	○					—	—	—	+	+	
2002	SSU [124]	●	●		■	■	◆	◆					—	—	—	+	—	+
2002	SSU [125]	●	●	●	○		○	○	○	○	○		—	—	—	+	—	+
2002	Tub [125]	●	●	●			◆	◆	○	○	○		—	—	—	+	—	—
2003	SSU [40]	●	●	●	■	■	◆	◆	○	○	○		—	—	—	+	—	+
2003	SSU [60]		●	●	■	■	◆	◆	●	●	●	○	+	—	+	+	—	+
2003	SSU [126]	○			■	■	○	○					—		—	+	—	
2005	Multi [127]	●	●	●	■	■	◆	◆	○	○	○		+	+	—	+	+	
2005	Multi [128]	●						●			●			+	+	+	+	
2000	Multi [115]	○					◆	◆			○		+	+	—	+	+	
2002	SSU [33]	●					●	●		●	●		+	—	+	+	—	—
2004	SSU [49]		○		■	■	○	○	○	○	○		—	—	—	+	—	+

Figure 6. Support for Membership and Supergroup Monophyly from “Excavata”-Targeted Molecular Genealogies

Member taxa: Di, Diplomonadida; Rt, Retortamonadida; Cp, *Carpodomonas*; Tr, *Trimastix*; Ox, Oxymonadida; Ht, Heterolobosea; Eu, Euglenozoa; MI, *Malawimonas*; Jk, Jakobida; Pa, Parabasalia; Dy, *Diphyllaea*. Hypothesized subgroups: ● Fornicata clade (Di + Rt + Cp) monophyletic, ■ Preaxostyla clade (Ox + Tr) monophyletic, ◆ Discicristata clade (Ht + Eu) monophyletic. The position of *Diphyllaea*, Dy, was not considered when scoring the monophyly of “Excavata” as the inclusion of this organism within “Excavata” is controversial and has been removed from recent classifications (see text). See Figure 4 for further notes. References cited in this figure are [33,40,49,60,115,123–128]. doi:10.1371/journal.pgen.0020220.g006

Year	Gene	Ref	Membership									Supergroup Monophyly						
			Mt	Fu	Cf	Cy	Ic	Co	Nu	Mi	Ap	A	C	E	O	P	R	
1993	SSU	[24]	●	●	●								—	+		+		
1993	βTub	[27]	●	●											—	+		
1993	Actin	[27]	●	●									—	—	—	+		
1995	SSU	[25]	●	●	●						●		—	—		+	+	—
1996	SSU	[129]	●	●	●	●	●	●			●		—	—		+	+	—
1998	SSU	[130]	●	●	●		●				○					+	—	—
2003	SSU	[131]	●	●	●	●	●	●	●	●	○		—	—		+	—	+
2003	HSP90	[122]	●	●				●			○		—	—	—	+	—	—
2003	EF1α	[132]	●	●			●				○		+			+	—	—
2006	Multi	[26]	●	●	●	●	●	●	●	●	○		+	+	—	+		
2000	Multi	[115]	●	●									+	+	—	+	+	
2002	SSU	[33]	●	●	●	●	●	●			○		+	—	+	+	—	—
2004	SSU	[49]	●	●	●	●	●	●			○		—	—	—	+	—	+

Figure 7. Support for Membership and Supergroup Monophyly from “Opisthokonta”-Targeted Molecular Genealogies

Member taxa: Mt, Metazoa; Fu, Fungi; Cf, Choanomonada; Cy, chytrids; Ic, Ichthyospora (DRIPs); Cl, *Corallochytrium*; Nu, Nucleariida; Mi, *Ministeria*; Ap, apusomonads. The position of apusomonads, Ap, was not considered when scoring the monophyly of “Opisthokonta” as this organism is highly variable, and it has been removed from recent classifications (see text). See Figure 4 for further notes. References cited in this figure are [24–27,33,49,115,122,129–132].

doi:10.1371/journal.pgen.0020220.g007

Varying Support for Membership within and Monophyly of Targeted Supergroups

Several supergroups are generally well supported when targeted in molecular systematic studies. Strikingly, the monophyly of both the original and expanded “Opisthokonta” members is strongly supported in all investigations targeting the group (ten of ten, Figure 7). Two other supergroups are also well supported: “Rhizaria” monophyly is recovered in 11 of 14 studies focusing on this supergroup (Figure 9) and “Amoebozoa” retained in five of seven topologies (Figure 4). However, support for these groups is

expected, given that they were recognized from molecular gene trees [11,33].

“Excavata” rarely form a monophyletic group in molecular systematic studies targeting this supergroup (two of nine; Figure 6). Moreover, the position of putative members, jakobids, *Malawimonas*, parabasalids, and *Diphyllleia* vary by analysis (Figure 6). Three distinct subclades, all of which are supported by ultrastructural characters [40], are generally recovered (Fornicata [six of six], Preaxostyla [six of six], and Discicristata [five of eight]; Figure 6).

Support for two supergroups varies depending on the type of character used: plastid or nuclear. The monophyly of “Plantae” and “Chromalveolata” are well supported by plastid characters: four of four plastid analyses (Figure 8) and six of nine (Figure 5), respectively. The “Plantae” clade is monophyletic in only three of six analyses using nuclear genes, including Elongation Factor 2 [53] and a 100+ gene analysis that included very limited taxon sample [54]. Nuclear loci never support “Chromalveolata” (zero of six; Figure 5), though alveolates and stramenopiles often form a clade to the exclusion of haptophytes and cryptophytes (e.g., [24,97]; Figures 4 and 7).

Decreased Support for Monophyly of Supergroups as Out-Groups in Other Studies

For each genealogy we also assessed the monophyly of the supergroups when included as out-groups. Overall, we find that support for the monophyly of a given supergroup is stronger when targeted and support decreases when the same supergroup is included as an out-group in other studies.

This trend is particularly unexpected given our less stringent requirements for monophyly of out-groups: a minimum of only two members need be included, while targeted groups had broader taxon sampling (see Methods). A priori, it would seem that the lower stringency could allow a limited sample of supergroup members to substitute for

Year	Gene	Loc	Ref	Membership			Supergroup Monophyly							
				Gr	Rh	Gl	A	C	E	O	P	R		
2002	Genome Pla	[133]	●	●	●		—					+		
2004	Multi	Pla	[134]	●	●	●		—				+		
2004	Genome Pla	[135]	●	●	●			—				+		
2005	Multi	Pla	[54]	●	●	●			+			+		
1995	SSU	Nuc	[136]	○	○	○		—	—			+	—	+
1997	Actin	Nuc	[137]	●	●	●		+	—			—	+	
2000	EF2	Nuc	[53]	●	●			—	—	—		+	+	
2001	RPB1	Nuc	[138]	○	○			—		—	+	+	—	
2005	Multi	Nuc	[139]	○	○	○		—	+	+	+	—	—	
2005	Multi	Nuc	[54]	●	●	●			+			+	+	
2000	Multi	Nuc	[115]	●	●	●		+	+	—		+	+	
2002	SSU	Nuc	[33]	○	○	○		+	—	+	+	—	—	
2004	SSU	Nuc	[49]	○	○	○		—	—	—		+	—	+

Figure 8. Support for Membership and Supergroup Monophyly from “Plantae”-Targeted Molecular Genealogies

Member taxa: Gr, Chloroplastida = Viridiplantae (Green algae, including land plants); Rd, Rhodophyceae (Red algae); Gl, Glaucophyta. See Figure 4 for general notes and Figure 5 for plastid-specific notes. References cited in this figure are [33,49,53,54,115,133–139].

doi:10.1371/journal.pgen.0020220.g008

is support for lower-level groups, such as the “Excavata” subgroups discussed above and perhaps also alveolates plus stramenopiles. This suggests that either there are no higher-level groupings to be found, or there is as yet inadequate data to resolve these clades. We believe that lack of taxon sampling is the key to resolution.

Further evidence against the six-supergroup view of eukaryotic diversity is the existence of “nomadic” taxa—lineages that do not have a consistent sister group, but instead wander between various weakly supported positions. Some nomadic taxa are acknowledged *incertae sedis* (of unknown taxonomic position) such as *Ancyromonas*, *Breviata*, and *Apusomonadidae* [7,8]. Other taxa that have been assigned to supergroups also appear to be nomadic, including Haptophyta (putative member of “Chromalveolata”) and *Malawimonas* (putative member of “Excavata”). For example, the haptophytes variously branch with Centrohelida and red algae [45], sister to a clade of “Rhizaria” and Heterolobosea [48], sister to cryptophytes [56], and in a basal polytomy [61]. These nomadic taxa may either represent independent, early diverging lineages or their phylogenetic position cannot yet be resolved with the data available. Again, we feel that taxon sampling is the key in order to distinguish between these possibilities.

Why Is Eukaryotic Taxonomy So Difficult?

The variable support for relationships is in part attributable to the inherent difficulty of deep phylogeny, the chimeric nature of eukaryotes, misidentified organisms, and conflicting approaches to taxonomy. Here we elaborate on these destabilizing trends and provide illustrative examples.

Challenges of deep phylogeny. Reconstructing the history of eukaryotic lineages requires extraction of phylogenetic signal from the noise that has accumulated over many hundreds of millions of years of divergent evolutionary histories. There is doubt whether resolution of divergences this deep can be resolved with molecular data [62]. Additionally, the nature of the relationships may also pose a significant challenge. For example, a rapid radiation of major eukaryotic lineages has been proposed [63] and is the most difficult scenario to resolve because of the lack of time to accumulate synapomorphies at deep nodes.

Further, phylogenetic relationships can be obscured by heterogeneous rates of evolution and divergent selection pressures. For example, genes in many parasitic lineages of eukaryotes experience elevated rates of evolution. If not properly accounted for, these fast lineages will group together due to long-branch attraction [64,65]. This was the case for Microsporidia, intracellular parasites of animals; early small subunit rDNA (SSU) genealogies placed the Microsporidia at the base of the tree with other amitochondriate taxa, including *Giardia* and *Entamoeba* [66]. These parasites were united under the “Archezoa” hypothesis [67]. More recent analyses with appropriate models of evolution [68] and those using protein-coding genes [69] place the Microsporidia within fungi and falsify “Archezoa.” This example demonstrates the importance of phylogenetic methods in the interpretation of eukaryotic diversity. In our analysis we find no clear correlation between method of tree building and group stability. Arguments about phylogenetic inference have been discussed extensively [62,70–76],

and increasingly sophisticated algorithms are being developed to compensate for the difficulties [77–79].

The chimeric nature of eukaryotes. Reconstructing the history of eukaryotic lineages is complicated by the horizontal transfer of genes and organelles [74,80–83]. For example, “Chromalveolata” plastid genes tell one story, consistent with a single transfer from red algae, which is not currently supported by available nuclear genes (Figure 5). There is also a growing body of evidence for aberrant lateral gene transfers in eukaryotes (reviewed in [80,82]).

Instability due to misidentification. Misidentification destabilizes taxonomy because all efforts to classify a misidentified organism reach erroneous conclusions. Cases of misidentification lead to inaccurate conclusions and require considerable effort to remedy. There is a rigorous standard for identifying microbial eukaryotes, but this standard is not always upheld. For example, the putative “Amoebozoa” species “*Mastigamoeba invertens*” that always branched outside the “Amoebozoa” clade [45,49,84] was misidentified [85]; it has now been properly described as *Breviata anathema* and is not yet placed within any of the supergroups [85].

Inaccurate conclusions about organismal relationships can also result from contamination (e.g., from symbionts and parasites). The results of subsequent molecular genealogies are therefore wrong and misleading. For example, opalinids, multinucleated flagellates that inhabit the lower digestive track of Anurans, were placed in the stramenopiles (Slopalinida: “Chromalveolata”) based on ultrastructural data [86]. However, the first molecular sequences for this group placed them within fungi (*Opalina ranarum* and *Cepedea virguloidea* [87,88]). These sequences were later shown to belong to zygomycete fungal contaminants, not to the opalinids. Subsequent isolates (*Protoopalina intestinalis*) yielded genealogies congruent with the ultrastructural data, placing *P. intestinalis* within the stramenopiles [89]. To avoid setbacks and confusion due to misidentification, we propose that all analyses of eukaryotic diversity include a vouchersing system for strains, images, and DNAs.

Conflicting approaches to taxonomy. Our evaluation of the stability of taxonomy for supergroups reveals a rapidly changing landscape (Figures 1 and 3). The instability in higher-level classifications of eukaryotes reflects the diversity of philosophical approaches, the exploratory state of eukaryotic taxonomy, and premature taxon naming. Many researchers seek schemes based on monophyletic groupings so that their taxonomies reflect evolutionary relationships [7,8,90,91]. In contrast, others employ a taxonomic philosophy in which evolutionary relatedness and monophyly are just one criterion from a set of group characteristics [33]. Paraphyly—a taxon defined without all descendants—is tolerated in these systems, and paraphyletic taxa are designated as such (see [6] p. 210–215 for explanation of such a philosophy).

In many cases, classification schemes that are separated by two years or less vary substantially from one another (e.g., Figure 3A and 3B). New groups and fluctuating group composition result in numerous cases of homonymy (two concepts linked to one name), synonymy (one concept linked to two names), and redefinition of existing terms. For example, at the highest level the terms “Amoebozoa,” “Opisthokonta,” and “Plantae” were all introduced under different definitions [4,44,52] before being applied to super-

groups. The term “Plantae” is an extreme case of homonymy having referred to numerous groups of photosynthetic organisms over the past century and a half (Figure 3C). The rapidly changing taxonomic landscape makes it difficult for non-specialists as well as specialists to follow the current debate over supergroups.

Toward a Robust Scaffold to the Eukaryotic Tree of Life

Taxonomic sampling. Perhaps the most critical aspect of the current state of eukaryotic systematics is the very limited taxonomic sampling to date. This is particularly problematic as the supergroup literature is often derived from a resampled pool of genes and taxa. More than 60 lineages of microbial eukaryotes have been identified by ultrastructure [8], yet only about one-half of these have been included in molecular analyses. Furthermore, even when these lineages are included, they are generally represented by a single species. Such sparse sampling increases the risk of long-branch attraction as discussed above, such as occurred for *Giardia*, and may cause artifactual relationships [92]. Further, analyses of sequences from newly sampled lineages have altered or expanded supergroup definitions (e.g., nucleariids in “Opisthokonta” [93] and Phaeodarea in “Rhizaria” [94]). Thus, statements of monophyly may be premature when taxonomic sampling is low.

There is tension between increasing the number of taxa versus the numbers of genes. Several theoretical works have demonstrated the diminishing returns of increased number of genes relative to increased taxon sampling [95–97], but see [98]. In addition, increasing taxon sampling can lead to shifts in molecular tree topology [99–101]. These results provide incentive to concentrate sequencing efforts on obtaining more taxa and a moderate number of genes. We recommend increasing the lineages sampled and the number of diverse taxa within lineages. We are optimistic that as data become available from a greater diversity of taxa, eukaryotic phylogeny will become increasingly more resolved.

Multiple character sets. We further anticipate that support for clades will increase as additional character sets are incorporated. Phylogenies based on single characters, whether genes, morphology, or ultrastructure, are subject to biases in the data and are not reliable by themselves. Hence, multiple character sets should be used to corroborate results. Ultrastructural apomorphies combined with molecular genealogies have proven to be good indicators of phylogeny at the level below supergroups [40,102]. This approach has bolstered support for “Fornicata” and “Preaxostyla,” which are consistently recovered in molecular genealogies and have defining ultrastructural characters. As we move forward with multiple character sets, we must shift from searching for characters to support hypotheses to evaluating hypotheses in light of all available data.

Well-sampled multigene and genome scale molecular systematics provide another powerful tool for resolving ancient splits in the tree of life. The National Science Foundation initiative “Assembling the Tree of Life” provides evidence of this shift in systematics research, whereby all proposals involve multigene or genome (organellar) sequencing to establish robust phylogenetic hypotheses (see <http://www.nsf.gov/pubs/2005/nsf05523/nsf05523.htm>; [54,97]). The EuTree consortium (<http://www.eutree.org>) aims to increase

substantially the sampled diversity of eukaryotes by focusing on understudied lineages in our multigene project to assemble the tree of life.

An example of multigene study is analysis of genes involved in clade-specific functions. This approach has been employed in testing “Plantae” and “Chromalveolata” (e.g., [103]). A single endosymbiosis (of a cyanobacterium in “Plantae” and red alga in “Chromalveolata”) predicts that the systems that facilitate controlled exchange of metabolic intermediates between the symbiotic partners be shared by putative members of these two supergroups [104]. This prediction has been supported by analyses of the plastid import machinery [105] and antiporters that transport fixed carbons across the plastid membranes [106]. However, taxon sampling has been limited in these studies. Currently, increased sampling of genomes from diverse photosynthetic eukaryotes is yielding additional genes for clade-specific predictions [107,108].

A conservative approach to taxonomy. Because taxonomy is the foundation for much of the dialog and research in evolutionary biology, there must be an unambiguous taxonomic system in which one term is linked to one concept. In contrast to this ideal, homonymy and redefinition are prevalent in the taxonomy of eukaryotes, often as the result of premature introduction or redefinition of taxa (see above; Figure 3). Emerging hypotheses benefit the community by sparking new research to test the hypothesis, but they also introduce ambiguity. To alleviate the confusion, we suggest introducing hypotheses as informal groups and using inverted commas to indicate the existence of a caveat, as done with the uncertain groups in this manuscript. These steps will inform the community that group composition is likely to change, alleviate quick taxon turnover, and promote stable taxa that are more resistant to compositional change.

As increasing amounts of data become available, well-supported nodes emerge and classifications tend to stabilize, such as is occurring for the ordinal framework for angiosperms [109,110]. Similarly, we expect that this conservative approach, combined with increased sampling of taxa and genes, will promote the future stabilization of eukaryotic classification.

Conclusion

Although the level of support varies among groups, the current classification of eukaryotes into six supergroups is being adopted broadly by the biological community (i.e., evidenced by its appearance in biology textbooks). The supergroup “Opisthokonta” and a number of nested clades within supergroups are supported by most studies. However, support for “Amoebozoa,” “Chromalveolata,” “Excavata,” “Plantae,” and “Rhizaria” is less consistent. The supergroups, and eukaryotic taxonomy in general, are further destabilized by considerable fluidity of taxa, taxon membership, and ambiguous nomenclature as revealed by comparison of classification schemes.

The accurate reconstruction of the eukaryotic tree of life requires: (1) a more inclusive sample of microbial eukaryotes; (2) distinguishing emerging hypotheses from taxa corroborated by multiple datasets; and (3) a conservative, mutually agreed upon approach to establishing taxonomies. Analyses of these types of data from a broad, inclusive sampling of

eukaryotes are likely to lead to a robust scaffold for the eukaryotic tree of life.

Methods

Stability of taxonomy. To assess the stability of supergroup taxonomies over time, we selected three classification schemes for each supergroup and tracked both the stability of taxa membership (solid and dashed lines; Figures 1 and 3) and the fate of newly created taxon names (asterisk; Figure 3). In sampling representative taxonomies, we aimed to capture a diversity of authors and opinions. In the case of “Opisthokonta” and “Chromalveolata” we are aware of only one formal, peer-reviewed classification scheme [7]. Given the lack of equivalency in ranks between taxonomies, we have chosen to display three levels with the intention of listing equivalent levels clearly.

Membership support. Within each supergroup, we assess the support for each member taxon by documenting its inclusion in molecular genealogies (Figures 4–9). Member taxa were chosen because they are historically a well-supported group, usually with an ultrastructural identity. The haptophytes are such a group, and share a haptonema [8]. We included members that represent a broad interpretation of the supergroup. For example, “Rhizaria” member taxa include groups (e.g., apusomonads) originally placed in “Rhizaria” but later removed. We considered a taxon to be a supported member of its supergroup (filled circles; Figures 4–9) when it falls within a monophyletic clade containing a majority of the supergroup members. A taxon that falls outside of its supergroup clade, or on the occasion that a majority of members do not form a monophyletic clade, is considered unsupported in that genealogy (open circles; Figures 4–9).

The inclusion of a genealogy requires that it be found in a paper that specifically addresses one of the supergroups or analyzes broad eukaryotic diversity. The genealogies must also include adequate sampling—two-member taxa per supergroup—from at least two of the six supergroups to allow for the comparison of supergroup monophyly. In cases where multiple gene trees are presented we display the authors’ findings as multiple entries when the trees are not congruent or as a single entry when the trees are concordant. Due to the lack of monophyly in virtually all analyses, we have evaluated the support for several hypothesized subgroups within the “Excavata” (geometric shapes; Figure 6).

Supergroup monophyly. To assess monophyly of supergroups, we used the set of genealogies described above to evaluate the molecular

support for the supergroups as interpreted by Adl et al. 2005 ([7]; Figures 4–9). We analyzed the monophyly [43] of each supergroup in trees having at least two member taxa present (+/– Figures 4–9). We do not indicate the method of tree construction. Although the algorithm used is important, we did not find a clear correlation between supported groups and algorithm used. We were also liberal in accepting any level of support (e.g., bootstrap values and posterior probabilities ranged from 4%–100%) when determining monophyly, in part because there is debate over acceptable cutoff values [111–113].

Supporting Information

Accession Numbers

Information about commonly used genes for phylogenesis of microbial eukaryotes discussed in this paper can be found in the Homologene database at NCBI (<http://www.ncbi.nlm.nih.gov/Genbank>): actin (88645), α -tubulin (81745), β -tubulin (69099), Elongation Factor 1 α gene (68181), small subunit rDNA (6629), and ubiquitin (39626). Accession numbers for genes from misidentified organisms can be found at NCBI in GenBank (<http://www.ncbi.nlm.nih.gov/Genbank>). Misidentified opalinids: *Opalina ranarum* (AF141969) and *Cepedea virguloidea* (AF141970); correctly identified *Protoopalina intestinalis* (AY576544–AY576546) and *Breviata anathema* (AF153206). Sequences for *Encephalitozoon cuniculi* can be found at NCBI under genome project number 9545.

Acknowledgments

Many thanks to Giselle Walker for discussions about supergroups and for graciously sharing a manuscript of her own. The authors also thank John Logsdon and Toby Kiers for comments and Jan Pawlowski for helpful discussions on nomadic lineages.

Author contributions. LWP, EB, DJP, and LAK conceived and designed the experiments. LWP, EB, EL, and LAK analyzed the data. LWP, EB, MD, DB, DJP, and LAK wrote the paper.

Funding. This work is supported by the National Science Foundation Assembling the Tree of Life grant (043115) to DB, DJP, and LAK.

Competing interests. The authors have declared that no competing interests exist.

References

- Javaux EJ, Knoll AH, Walter MR (2001) Morphological and ecological complexity in early eukaryotic ecosystems. *Nature* 412: 66–69.
- Whittaker RH (1969) New concepts of kingdoms of organisms. *Science* 163: 150–160.
- Dobell CC (1911) The principles of protistology. *Arch Protist* 23: 269–310.
- Copeland HF (1956) Classification of lower organisms. Palo Alto (California): Pacific Books. 302 p.
- Margulis L, Schwartz K (1988) Five kingdoms: An illustrated guide to the phyla of life on earth. 2nd edition. New York: W. H. Freeman and Company. 520 p.
- Cavalier-Smith T (1998) A revised six-kingdom system of life. *Biol Rev Cambridge Philosophic Soc* 73: 203–266.
- Adl SM, Simpson AGB, Farmer MA, Andersen RA, Anderson OR, et al. (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Euk Microbiol* 52: 399–451.
- Patterson DJ (1999) The diversity of eukaryotes. *Am Nat* 154: 96–124.
- Keeling PJ, Burger G, Durnford DG, Lang BF, Lee RW, et al. (2005) The tree of eukaryotes. *Trends Ecol Evol* 20: 670–676.
- Baldauf SL (2003) The deep roots of eukaryotes. *Science* 300: 1703–1706.
- Cavalier-Smith T (1996/97) Amoeboflagellates and mitochondrial cristae in eukaryote evolution: Megasytematics of the new protozoan subkingdoms Eozoa and Neozoa. *Arch Protistenkd* 147: 237–258.
- Williams JG, Noegel AA, Eichinger L (2005) Manifestations of multicellularity: Dictyostelium reports in. *Trends Genet* 21: 392–398.
- Stauffer W, Ravdin JI (2003) *Entamoeba histolytica*: An update. *Curr Opin Infect Dis* 16: 479–485.
- Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: Euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *J Euk Microbiol* 46: 347–366.
- Delwiche CF (1999) Tracing the tread of plastid diversity through the tapestry of life. *Am Nat* 154: 164–177.
- Keeling PJ (2004) Diversity and evolutionary history of plastids and their hosts. *Am J Bot* 91: 1481–1493.
- Cavalier-Smith T (1982) The origins of plastids. *Biol J Linn Soc* 17: 289–306.
- Prescott DM (1994) The DNA of ciliated protozoa. *Microbiol Rev* 58: 233–267.
- May KJ, Ristaino JB (2004) Identity of the mtDNA haplotype(s) of *Phytophthora infestans* in historical specimens from the Irish Potato Famine. *Mycol Res* 108: 471–479.
- Sherman IW (1998) Malaria. Washington (D. C.): ASM Press. 575 p.
- Simpson AGB, Patterson DJ (1999) The ultrastructure of *Carpodemonas membranifera* (Eukaryota) with reference to the “Excavate hypothesis.” *Eur J Protistol* 35: 353–370.
- Kreier J (1978) Parasitic Protozoa Volume II: Intestinal flagellates, histomonads, trichomonads, amoeba, opalinids, and ciliates. New York: Academic Press. 730 p.
- Smith P (1996) Molecular biology of parasitic protozoa. Oxford: Oxford University Press. 233 p.
- Wainright PO, Hinkle G, Sogin ML, Stickel SK (1993) Monophyletic origins of the Metazoa: An evolutionary link with fungi. *Science* 260: 340–342.
- Cavalier-Smith T, Chao EE (1995) The opalozoan *Apusomonas* is related to the common ancestor of animals, fungi, and choanoflagellates. *Proc Roy Soc Lond B-Biol Sci* 261: 1–6.
- Steenkamp ET, Wright J, Baldauf SL (2006) The protistan origins of animals and fungi. *Mol Biol Evol* 23: 93–106.
- Baldauf SL, Palmer JD (1993) Animals and fungi are each other’s closest relatives: Congruent evidence from multiple proteins. *Proc Natl Acad Sci, U S A* 90: 11558–11562.
- Keeling PJ, Slamovits CH (2004) Simplicity and complexity of microsporidian genomes. *Eukaryot Cell* 3: 1363–1369.
- King N (2004) The unicellular ancestry of animal development. *Dev Cell* 7: 313–325.
- Cavalier-Smith T (1981) Eukaryote kingdoms—seven or nine. *Biosystems* 14: 461–481.
- McFadden GI (2001) Primary and secondary endosymbiosis and the origin of plastids. *J Phycol* 37: 951–959.
- Michod RE, Viostat Y, Solari CA, Hurand M, Nedelcu AM (2006) Life-history evolution and the origin of multicellularity. *J Theor Bio* 239: 257–272.

33. Cavalier-Smith T (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. *Int J Syst Evol Microbiol* 52: 297–354.
34. Cavalier-Smith T, Chao EE (1997) Sarcomonad ribosomal RNA sequences, rhizopod phylogeny, and the origin of euglyphid amoebae. *Arch Protistenkd* 147: 227–236.
35. Bhattacharya D, Helmchen T, Melkonian M (1995) Molecular evolutionary analyses of nuclear-encoded small subunit ribosomal RNA identify an independent rhizopod lineage containing the euglyphina and the chlorarachniophyta. *J Euk Microbiol* 42: 65–69.
36. Archibald JM, Longet D, Pawlowski J, Keeling PJ (2003) A novel polyubiquitin structure in *Cercozoa* and Foraminifera: Evidence for a new eukaryotic supergroup. *Mol Biol Evol* 20: 62–66.
37. Yoon HS, Reyes-Prieto A, Melkonian M, Bhattacharya D (2006) Minimal plastid genome evolution in the *Paulinella* endosymbiont. *Curr Biol* 16: R670–R672.
38. Marin B, Nowack ECM, Melkonian M (2005) A plastid in the making: Primary endosymbiosis. *Protist* 156: 425–432.
39. Loeblich AR Jr, Tappan H (1987) Foraminiferal genera and their classification. New York: Van Nostrand Reinhold Company. 970 p.
40. Simpson AGB (2003) Cytoskeletal organization, phylogenetic affinities, and systematics in the contentious taxon Excavata (Eukaryota). *Int J Syst Evol Micro* 53: 1759–1777.
41. Doyle JJ (1992) Gene trees and species trees: Molecular systematics as one-character taxonomy. *Syst Bot* 17: 144–163.
42. Maddison WP (1997) Gene trees in species trees. *Syst Biol* 46: 523–536.
43. Hennig W (1966) Phylogenetic systematics. Urbana: University of Illinois Press. 263 p.
44. Lühe M (1913) Erstes Urreich der Tiere. In: Lang A, editor. *Handbuch der morphologie der wirbellosen Tiere*. Jena: G. Fischer. 416 p.
45. Cavalier-Smith T, Chao EEY, Oates B (2004) Molecular phylogeny of Amoebozoa and the evolutionary significance of the unikont *Phalansterium*. *Eur J Protistol* 40: 21–48.
46. Kudryavtsev A, Bernhard D, Schlegel M, Chao EEY, Cavalier-Smith T (2005) 18S ribosomal RNA gene sequences of *Cochliopodium* (Himatismenida) and the phylogeny of Amoebozoa. *Protist* 156: 215–224.
47. Smirnov A, Nassonova E, Berney C, Fahrni J, Bolivar I, et al. (2005) Molecular phylogeny and classification of the lobose amoebae. *Protist* 156: 129–142.
48. Fahrni JF, Bolivar I, Berney C, Nassonova E, Smirnov A, et al. (2003) Phylogeny of lobose amoebae based on actin and small-subunit ribosomal RNA genes. *Mol Biol Evol* 20: 1881–1886.
49. Cavalier-Smith T (2004) Only six kingdoms of life. *Proc Biol Sci* 271: 1251–1262.
50. Cavalier-Smith T, Chao EEY (2003) Phylogeny and classification of phylum Cercozoa (Protozoa). *Protist* 154: 341–358.
51. Philippe H, Adoutte A (1998) The molecular phylogeny of Eukaryota: Solid facts and uncertainties. In: Coombs GH, Vickerman K, Sleight MA, Warren A, editors. *Evolutionary relationships among protozoa*. Dordrecht: Kluwer Academic Publishers. pp. 25–56.
52. Haeckel EHPA (1866) *Generelle morphologie der organismen. Allgemeine grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenztheorie*. Berlin: G. Reimer.
53. Moreira D, Le Guyader H, Philippe H (2000) The origin of red algae and the evolution of chloroplasts. *Nature* 405: 69–72.
54. Rodriguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, et al. (2005) Monophyly of primary photosynthetic eukaryotes: Green plants, red algae, and glaucophytes. *Curr Biol* 15: 1325–1330.
55. Nishi A, Ishida K, Endoh H (2005) Reevaluation of the evolutionary position of opalinids based on 18S rDNA, and alpha- and beta-tubulin gene phylogenies. *J Mol Evol* 60: 695–705.
56. Harper JT, Waanders E, Keeling PJ (2005) On the monophyly of chromalveolates using a six-protein phylogeny of eukaryotes. *Int J Syst Evol Micro* 55: 487–496.
57. Longet D, Burki F, Flakowski J, Berney C, Polet S, et al. (2004) Multigene evidence for close evolutionary relations between *Gromia* and Foraminifera. *Acta Protozool* 43: 303–311.
58. Keeling PJ, Deane JA, McFadden GI (1998) The phylogenetic position of alpha- and beta-tubulins from the *Chlorarachnion* host and *Cercomonas* (Cercozoa). *J Eukaryot Microbiol* 45: 561–570.
59. Reece KS, Siddall ME, Stokes NA, Burren EM (2004) Molecular phylogeny of the Haplosporidia based on two independent gene sequences. *J Parasitol* 90: 1111–1122.
60. Cavalier-Smith T (2003) 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 Micro* 53: 1741–1758.
61. Nikolaev SI, Berney C, Fahrni JF, Bolivar I, Polet S, et al. (2004) The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. *Proc Natl Acad Sci U S A* 101: 8066–8071.
62. Penny D, McComish BJ, Charleston MA, Hendy MD (2001) Mathematical elegance with biochemical realism: The covarion model of molecular evolution. *J Mol Evol* 53: 711–723.
63. Philippe H, Lopez P, Brinkmann H, Budin K, Germot A, et al. (2000) Early branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. *Proc Biol Sci* 267: 1213–1221.
64. Felsenstein J (1988) Phylogenies from molecular sequences: Inference and reliability. *Annu Rev Genet* 22: 521–565.
65. Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* 27: 401–410.
66. Sogin ML (1991) Early evolution and the origin of eukaryotes. *Curr Opin Genet Dev* 1: 457–463.
67. Cavalier-Smith T (1983) A 6-kingdom classification and a unified phylogeny. In: Schenk HEA, Schwemmler W, editors. *Endocytobiology II: Intracellular space as oligogenetic ecosystem*. Berlin: Walter de Gruyter. pp. 1027–1034.
68. Van de Peer Y, Ben Ali A, Meyer A (2000) Microsporidia: Accumulating molecular evidence that a group of amitochondriate and suspectedly primitive eukaryotes are just curious fungi. *Gene* 246: 1–8.
69. Keeling PJ, Luker MA, Palmer JD (2000) Evidence from Beta-tubulin phylogeny that Microsporidia evolved within the fungi. *Mol Biol Evol* 17: 23–31.
70. Embley TM, van der Giezen M, Horner DS, Dyal PL, Bell S, et al. (2003) Hydrogenosomes, mitochondria, and early eukaryotic evolution. *IUBMB Life* 55: 387–395.
71. Susko E, Inagaki Y, Field C, Holder ME, Roger AJ (2002) Testing for differences in rates-across-sites distributions in phylogenetic subtrees. *Mol Biol Evol* 19: 1514–1523.
72. Gribaldo S, Philippe H (2002) Ancient phylogenetic relationships. *Theor Popul Biol* 61: 391–408.
73. Philippe H, Germot A (2000) Phylogeny of eukaryotes based on ribosomal RNA: Long-branch attraction and models of sequence evolution. *Mol Biol Evol* 17: 830–834.
74. Roger AJ (1999) Reconstructing early events in eukaryotic evolution. *Am Nat* 154: 146–163.
75. Philippe H, Forterre P (1999) The rooting of the universal tree of life is not reliable. *J Mol Evol* 49: 509–523.
76. Inagaki Y, Susko E, Fast NM, Roger AJ (2004) Covarion shifts cause a long-branch attraction artifact that unites Microsporidia and Archaeobacteria in EF-1 alpha phylogenies. *Mol Biol Evol* 21: 1340–1349.
77. Yang ZH (1997) PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput Appl Bio* 13: 555–556.
78. Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Evolution—Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294: 2310–2314.
79. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
80. Katz LA (2002) Lateral gene transfers and the evolution of eukaryotes: Theories and data. *Int J Syst Evol Micro* 52: 1893–1900.
81. Katz LA (1999) The tangled web: Gene genealogies and the origin of eukaryotes. *Am Nat* 154: 137–145.
82. Andersson JO (2005) Lateral gene transfer in eukaryotes. *Cell Mol Life Sci* 62: 1182–1197.
83. Huang JL, Mullapudi N, Sicheritz-Ponten T, Kissinger JC (2004) A first glimpse into the pattern and scale of gene transfer in the Apicomplexa. *Int J Parasitol* 34: 265–274.
84. Bolivar I, Fahrni JF, Smirnov A, Pawlowski J (2001) SSU rRNA-based phylogenetic position of the genera *Amoeba* and *Chaos* (Lobosea, Gymnamoebia): The origin of gymnamoebae revisited. *Mol Biol Evol* 18: 2306–2314.
85. Walker G, Dacks JB, Embley TM (2006) Ultrastructural description of *Breviata anathema*, n. Gen., n. Sp., the organism previously studied as “Mastigamoeba invertens.” *J Euk Microbiol* 53: 65–78.
86. Patterson DJ (1985) The fine structure of *Opalina ranarum* (family Opalinidae): Opalinid phylogeny and classification. *Protistologica* 4: 413–428.
87. Karpov S, Sogin ML, Silberman JD (2001) Rootlet homology, taxonomy, and phylogeny of bicosoecids based on 18S rRNA gene sequences. *Protistology* 2: 34–47.
88. Guillou L, Chretiennot-Dinet MJ, Boulben S, Moon-Van der Staay SY, Vault D (1999) *Symbiomonas scintillans* gen. et sp. nov. *Picophagus flagellatus* gen. et sp. nov. (Heterokonta): Two new heterotrophic flagellates of picoplanktonic size. *Protist* 150: 383–398.
89. Kostka M, Hampl V, Cepicka I, Flegr J (2004) Phylogenetic position of *Protoopalina intestinalis* based on SSU rRNA gene sequence. *Mol Phy Evol* 33: 220–224.
90. Lipscomb DL (1984) Methods of systematic analysis: The relative superiority of phylogenetic systematics. *Orig Life* 13: 235–240.
91. Simpson AGB (1997) The identity and composition of the Euglenozoa. *Archiv Fur Protistenkunde* 148: 318–328.
92. Hendy MD, Penny D (1989) A framework for the quantitative study of evolutionary trees. *Syst Zool* 38: 297–309.
93. Amaral-Zettler LA, Nerad TA, O’Kelly CJ, Sogin ML (2001) The nuclearioid amoebae: More protists at the animal-fungal boundary. *J Euk Microbiol* 48: 293–297.
94. 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.

95. Cummings MP, Meyer A (2005) Magic bullets and golden rules: Data sampling in molecular phylogenetics. *Zoology* 108: 329–336.
96. Cummings MP, Otto SP, Wakeley J (1995) Sampling properties of DNA-sequence data in phylogenetic analysis. *Mol Biol Evol* 12: 814–822.
97. Bapteste E, Brinkmann H, Lee JA, Moore DV, Sensen CW, et al. (2002) The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. *Proc Natl Acad Sci U S A* 99: 1414–1419.
98. Rokas A, Carroll SB (2005) More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. *Mol Biol Evol* 22: 1337–1344.
99. DeBry RW (2005) The systematic component of phylogenetic error as a function of taxonomic sampling under parsimony. *Syst Biol* 54: 432–440.
100. Hillis DM (1998) Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Syst Biol* 47: 3–8.
101. Hillis DM, Pollock DD, McGuire JA, Zwickl DJ (2003) Is sparse taxon sampling a problem for phylogenetic inference? *Syst Biol* 52: 124–126.
102. Taylor FJR (1999) Ultrastructure as a control for protistan molecular phylogeny. *Am Nat* 154: 125–136.
103. Nosenko T, Lidie KL, Van Dolah FM (2006) Chimeric plastid proteome in the Florida “red tide” dinoflagellate *Karenia brevis*. *Mol Biol Evol* 23: 2026–2038.
104. Bhattacharya D, Yoon HS, Hackett JD (2004) Photosynthetic eukaryotes unite: Endosymbiosis connects the dots. *Bioessays* 26: 50–60.
105. McFadden GI, van Dooren GG (2004) Evolution: Red algal genome affirms a common origin of all plastids. *Curr Biol* 14: R514–R516.
106. Weber A, Linka M, Bhattacharya D (2006) Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembranes-derived ancestor. *Eukaryot Cell* 5: 609–612.
107. Yoon HS, Hackett JD, Van Dolah FM, Nosenko T, Lidie L, et al. (2005) Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. *Mol Biol Evol* 22: 1299–1308.
108. Sanchez-Puerta MV, Bachvaroff TR, Delwiche CF (2005) The complete plastid genome sequence of the haptophyte *Emiliania huxleyi*: A comparison to other plastid genomes. *DNA Res* 12: 151–156.
109. APG II (2003) An update of the Angiosperm Phylogeny group classification for the orders and families of flowering plants. *Bot J Linnean Soc*: 399–436.
110. APG I (1998) An ordinal classification for the families of flowering plants. *Ann Missouri Bot Gard*: 531–553.
111. Felsenstein J (2004) Inferring phylogenies. Sunderland, Massachusetts: Sinaur Associates. 580 p.
112. Efron B, Halloran E, Holmes S (1996) Bootstrap confidence levels for phylogenetic trees. *Proc Natl Acad Sci, U S A* 93: 13429–13434.
113. Lewis PO, Holder MT, Holsinger KE (2005) Polytomies and Bayesian phylogenetic inference. *Syst Biol* 54: 241–253.
114. Cavalier-Smith T (1997) Amoeboflagellates and mitochondrial cristae in eukaryote evolution: Megasystematics of the new protozoan subkingdoms Eozoa and Neozoa. *Arch Protistenkd* 147: 237–258.
115. Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittle WF (2000) A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* 290: 972–977.
116. Zhang Z, Green BR, Cavalier-Smith T (1999) Single-gene circles in dinoflagellate chloroplast genomes. *Nature* 400: 155–160.
117. Yoon HS, Hackett JD, Pinto G, Bhattacharya D (2002) The single, ancient origin of chromist plastids. *Proc Natl Acad Sci U S A* 99: 15507–15512.
118. Harper JT, Keeling PJ (2003) Nucleus-encoded, plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) indicates a single origin for chromalveolate plastids. *Mol Biol Evol* 20: 1730–1735.
119. Inagaki Y, Simpson AGB, Dacks JB, Roger AJ (2004) Phylogenetic artifacts can be caused by leucine, serine, and arginine codon usage heterogeneity: Dinoflagellate plastid origins as a case study. *Syst Biol* 53: 582–593.
120. Patron NJ, Rogers MB, Keeling PJ (2004) Gene replacement of fructose-1,6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates. *Eukaryot Cell* 3: 1169–1175.
121. Bachvaroff TR, Puerta MVS, Delwiche CF (2005) Chlorophyll c-containing plastid relationships based on analyses of a multigene dataset with all four chromalveolate lineages. *Mol Biol Evol* 22: 1772–1782.
122. Stechmann A, Cavalier-Smith T (2003) Phylogenetic analysis of eukaryotes using heat-shock protein Hsp90. *J Mol Evol* 57: 408–419.
123. Dacks JB, Silberman JD, Simpson AGB, Moriya S, Kudo T, et al. (2001) Oxymonads are closely related to the excavate taxon *Trimastix*. *Mol Biol Evol* 18: 1034–1044.
124. Silberman JD, Simpson AGB, Kulda J, Cepicka I, Hampl V, et al. (2002) Retortamonad flagellates are closely related to diplomonads—Implications for the history of mitochondrial function in eukaryote evolution. *Mol Biol Evol* 19: 777–786.
125. Simpson AGB, Roger AJ, Silberman JD, Leipe DD, Edgcomb VP, et al. (2002) Evolutionary history of “early-diverging” eukaryotes: The excavate taxon *Carpediemonas* is a close relative of *Giardia*. *Mol Biol Evol* 19: 1782–1791.
126. Moriya S, Dacks JB, Takagi A, Noda S, Ohkuma M, et al. (2003) Molecular phylogeny of three oxymonad genera: *Pyronympha*, *Dinenympha*, and *Oxymonas*. *J Eukaryot Microbiol* 50: 190–197.
127. Simpson AGB, Inagaki Y, Roger AJ (2006) Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of “primitive” eukaryotes. *Mol Biol Evol* 23: 615–625.
128. Hampl V, Horner DS, Dyal P, Kulda J, Flegr J, et al. (2005) Inference of the phylogenetic position of oxymonads based on nine genes: Support for Metamonada and Excavata. *Mol Biol Evol* 22: 2508–2518.
129. Cavalier-Smith T, Allsopp M (1996) *Corallochytrium*, an enigmatic non-flagellate protozoan related to choanoflagellates. *Eur J Protistol* 32: 306–310.
130. Ragan MA, Goggin CL, Cawthorn RJ, Cerenius L, Jamieson AVC, et al. (1996) A novel clade of protistan parasites near the animal-fungal divergence. *Proc Natl Acad Sci U S A* 93: 11907–11912.
131. Cavalier-Smith T, Chao EEEY (2003) Phylogeny of choanozoa, apusozoa, and other protozoa and early eukaryote megaevolution. *J Mol Evol* 56: 540–563.
132. Ragan MA, Murphy CA, Rand TG (2003) Are Ichthyosporea animals or fungi? Bayesian phylogenetic analysis of elongation factor 1 alpha of *Ichthyophonus irregularis*. *Mol Phylogenet Evol* 29: 550–562.
133. Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, et al. (2002) Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci U S A* 99: 12246–12251.
134. Hagopian JC, Reis M, Kitajima JP, Bhattacharya D, de Oliveira MC (2004) Comparative analysis of the complete plastid genome sequence of the red alga *Gracilaria tenuistipitata* var. liui provides insights into the evolution of rhodoplasts and their relationship to other plastids. *J Mol Evol* 59: 464–477.
135. Chu KH, Qi J, Yu ZG, Anh V (2004) Origin and phylogeny of chloroplasts revealed by a simple correlation analysis of complete genomes. *Mol Biol Evol* 21: 200–206.
136. Bhattacharya D, Helmchen T, Bibeau C, Melkonian M (1995) Comparisons of nuclear-encoded small-subunit ribosomal-RNAs reveal the evolutionary position of the Glaucocystophyta. *Mol Biol Evol* 12: 415–420.
137. Bhattacharya D, Weber K (1997) The actin gene of the glaucocystophyte *Cyanophora paradoxa*: Analysis of the coding region and introns, and an actin phylogeny of eukaryotes. *Curr Genet* 31: 439–446.
138. Stiller JW, Riley J, Hall BD (2001) Are red algae plants? A critical evaluation of three key molecular datasets. *J Mol Evol* 52: 527–539.
139. Nozaki H (2005) A new scenario of plastid evolution: Plastid primary endosymbiosis before the divergence of the “Plantae,” emended. *J Plant Res* 118: 247–255.
140. Cavalier-Smith T, Chao EE (1997) Sarcomonad ribosomal RNA sequences, rhizopod phylogeny, and the origin of euglyphid amoebae. *Archiv Fur Protistenkd* 147: 227–236.
141. Keeling PJ (2001) Foraminifera and Cercozoa are related in actin phylogeny: Two orphans find a home? *Mol Biol Evol* 18: 1551–1557.
142. 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.
143. Nikolaev SI, Berney C, Fahrni J, Mylnikov AP, Aleshin VV, et al. (2003) *Gymnophrys cometa* and *Lecythium* sp are core Cercozoa: Evolutionary implications. *Acta Protozool* 42: 183–190.
144. 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 Micr* 53: 1735–1739.
145. Archibald JM, Keeling PJ (2004) Actin and ubiquitin protein sequences support a cercozoan/foraminiferan ancestry for the plasmodiophorid plant pathogens. *J Eukaryot Microbiol* 51: 113–118.