CellPress

Current Biology
Review

Progress towards the Tree of Eukaryotes

Patrick J. Keeling[1,*] and Fabien Burki[2,3]
[1]Department of Botany, University of British Columbia, Vancouver V6T 1Z4, British Columbia, Canada
[2]Department of Organismal Biology, Program in Systematic Biology, Uppsala University, Uppsala, Sweden
[3]Science for Life Laboratory, Uppsala University, Uppsala, Sweden
*Correspondence: pkeeling@mail.ubc.ca
https://doi.org/10.1016/j.cub.2019.07.031

Developing a detailed understanding of how all known forms of life are related to one another in the tree of life has been a major preoccupation of biology since the idea of tree-like evolution first took hold. Since most life is microbial, our intuitive use of morphological comparisons to infer relatedness only goes so far, and molecular sequence data, most recently from genomes and transcriptomes, has been the primary means to infer these relationships. For prokaryotes this presented new challenges, since the degree of horizontal gene transfer led some to question the tree-like depiction of evolution altogether. Most eukaryotes are also microbial, but in contrast to prokaryotic life, the application of large-scale molecular data to the tree of eukaryotes has largely been a constructive process, leading to a small number of very diverse lineages, or 'supergroups'. The tree is not completely resolved, and contentious problems remain, but many well-established supergroups now encompass much more diversity than the traditional kingdoms. Some of the most exciting recent developments come from the discovery of branches in the tree that we previously had no inkling even existed, many of which are of great ecological or evolutionary interest. These new branches highlight the need for more exploration, by high-throughput molecular surveys, but also more traditional means of observations and cultivation.

Introduction: The Tree of Life Concept

The 'tree of life' is an important concept that helps us think more clearly about the distant, sometimes murky past of early evolution. Phylogenetic trees in general are a powerful visual aid to depict evolutionary relationships along a range of time scales in a simple branching diagram. They are not family trees in the sense that they show the births, deaths, and unions of contemporaneous organisms, but rather a foundation on which to understand long-term evolutionary processes that happened at the level of populations that are now mostly extinct. As useful cartoons, they can help organize complex evolutionary events; but this strength can also be a weakness, since trees are open to a range of misinterpretations [1,2]. One of the most common is the notion of 'lower' and 'higher' forms of life depicted as early-diverging or late-diverging branches, which are easily misconstrued as steps in a *scala naturae*, or a great chain of being, from simple to complex [3,4]. Of course, all extant organisms have been evolving for the same period of time and none are lower or higher. Molecular phylogenies depicting relationships between species also often make the important assumption that the evolutionary history of the gene(s) used to make the tree somehow represents the species tree [5,6]. Complicating processes like horizontal gene transfer (HGT), gene gain/loss, hybridization, or endosymbiosis can be revealed or obscured depending on the data, a problem amplified in an era where hundreds or thousands of genes are routinely used to produce species trees.

All these complexities became practical problems with the democratization and maturation of genomic methods. For prokaryotes, it was quickly proposed that networks, rather than bifurcating phylogenies, might better describe the evolution of species due to a generally high frequency of HGT [7,8]. The more philosophically minded made the case that the prevalence of HGT was high enough to even question the whole concept of tree-like evolution in prokaryotes, or indeed the existence of a tree of life [9,10]. If HGT and other complicating factors are so common, then potentially every gene in a genome has a different history and perhaps no gene fits the assumption that the evolution of species can be accurately described by bifurcating trees. Today, with vast quantities of bacterial and archaeal genomic data, the important impact of HGT on prokaryote evolution is universally accepted, but the debate over the notion of the tree-like branching of species remains active [11].

What about eukaryotes? When we think of eukaryotes, it is the animals, plants, and fungi that quickly jump to mind, and in these well-studied, macroscopic, and often obligately sexual lineages an overall tree-like evolutionary history has hardly been questioned. However, these organisms represent the minority of the lineage diversity: most eukaryotes are single-celled organisms, and for most of eukaryotic history only microbial species have existed [12–14]. Microbial eukaryotes, or protists, represent dozens of 'kingdom-equivalents', or major lineages containing sometimes astonishingly high numbers of species that share both ancestry and overall cellular and genomic characteristics with their better-known multicellular cousins. But as microbes, they also share many ecological, cell biological, and evolutionary characteristics with prokaryotes [15]. Whether genomics would resolve the tree of eukaryotes or complicate it, as it did with prokaryotes, was an important question that was seldom asked while it was being answered.

Here, we review the current state of knowledge about the tree of eukaryotes, look back on how it took shape and what that history tells us about its reliability, but also look towards some of the future challenges that we know, or suspect, might dominate the

Figure 1. Examples of eukaryotic supergroups.
This plate shows an organism from each of the major supergroups identified in Figure 2, with coloured boxes corresponding to the colour for that group in Figure 2. From left to right, the top row shows the ciliate *Euplotes* (Alveolates) and the diatom *Pinnularia* (Stramenopiles). The second row shows star sand foraminiferans (Rhizarians), a cryptomonad (Cryptista) above a centrohelid (Haptista), and the charophyte green alga *Micrasterias* (Archaeplastids). The third row shows the euglenid *Euglena* (Discobids) and the tubalinid *Rhizamoeba* (Amoebozoans). The bottom row shows the parabasalian *Trichonympha* (Metamonad), the cnidrian animal *Acropora* (Opisthokonts) and a dikaryan fungus (Opisthokonts). All photos by P. Keeling.

field of eukaryote evolution in the coming years. We also compare these advances with parallel discoveries in the diversity and evolution of bacteria and archaea. These characteristics highlight fundamental differences between the evolution of prokaryotes and eukaryotes, as well as the central place of protists in understanding these differences.

A Brief History of Eukaryotic Phylogeny

Before looking at the current tree of eukaryotes, let's take a short historical detour to better appreciate the winding path of discovery. The first trees of eukaryotes were obviously not based on molecular data, but used morphological or trophic characteristics to decipher evolutionary relationships [16–18]. This was relatively straightforward until modern microscopy began to reveal how incredibly diverse these seemingly simple single-celled organisms really were (Figure 1). Early attempts to organize protist diversity based on morphology were successful in identifying a large number of lineages that we still acknowledge today, but the morphological diversity was so great, and characters so often conflicting, that it was impossible to see the details of the relationships between most of these lineages. Attempts to synthesize this information were forced to leave them unresolved [17], or lump them into a 'blob' of 'lower' forms of life [18]. Early molecular phylogenies based on the small subunit ribosomal DNA quickly painted a different picture, revealing that the majority of phylogenetic diversity in eukaryotes was actually in the protists, rather than in the macroscopic organisms [19]. These molecular trees also consistently showed a distinct structure, with a mix of protists and multicellular organisms in the so-called 'crown groups', and a basal ladder of protists with seemingly simpler, more 'primitive' characteristics closer to some conception of the ancestral state [19–22].

As sequencing technologies evolved, the increasing use of protein-coding genes for phylogenies slowly began to challenge this basal/crown dichotomy [23–27]. This led to a period where the tree was more torn down than built up, because no combination of genes or ultrastructural features offered an improved resolution for the entire tree. Although important insights emerged from these relatively small-scale analyses based on both rRNA and proteins (e.g. the union of many small lineages into large ones, like alveolates, stramenopiles, cercozoans, rhizarians, and opisthokonts [24,28–31]), the big picture only began to change fundamentally when the ever-greater access to genomic data produced trees based on many genes concatenated into larger and larger data sets. This 'phylogenomic' approach has now grown to employ hundreds or even thousands of genes to infer phylogenies [32–35]. These phylogenomic analyses have converged on a fundamentally different picture of eukaryotic relationships, in which the diversity is distributed into a small number of very large assemblages [12]. These assemblages included recognized kingdoms (e.g., animals, plants, fungi) embedded as one of many subgroups (see Box 1 for a glossary of these groups), and were informally called

Box 1. A brief description of the major groups of eukaryotes.

Stramenopiles comprise well-known microbial algae (e.g., diatoms) but also macroscopic multicellular seaweeds (e.g., kelps) and a vast diversity of free-living heterotrophic or mixotrophic protists and important pathogens of animals and plants (e.g., *Blastocystis* or oomycetes). Stramenopiles also contain some of the most abundant environmental taxa in the sea (called MASTs for *MA*rine *ST*ramenopiles), most of which have not been characterized at the cellular level.

Alveolates comprise three well-studied protist groups (ciliates, dinoflagellates, and apicomplexans), in addition to several smaller groups of parasites and flagellates such as perkinsids, chrompodellids, and colponemids. Ciliates are a major group of microbial predators and grazers in all known environments. Dinoflagellates are also extremely abundant in nature, about half as photosynthetic algae and the rest are predators and parasites (including probably the most abundant marine eukaryotes, the MALVs). Apicomplexans in turn are all obligately associated with animals, most commonly as intracellular parasites (e.g. the causative agent of malaria).

Rhizarians includes a wide diversity of predominantly amoeboid protists with thin, filose pseudopodia used more in feeding than in locomotion. The group also includes amoeboflagellates and flagellates, parasites of crop plants and invertebrates (e.g. *Plasmodiophora* and *Bonamia*, respectively), and even amoeboid algae (chlorarachniophytes).

SAR is the conglomerate of Stramenopiles, Alveolates, and Rhizaria, which together make an assemblage encompassing perhaps half of all eukaryote diversity. Recently, the phylum *Telonema* was proposed as the sister lineage to SAR, making the even larger group TSAR. Telonemids contain only two described species, and are thought to represent a widespread, although not very abundant nor highly diverse, lineage of heterotrophic flagellates, and as sister to SAR would represent a good example of a 'lopsided tree' discussed in the text.

Haptista contains two main lineages: the haptophytes and centrohelids. Haptophytes are all photosynthetic, mostly marine species where they can bloom to high density (e.g. *Emiliania*). Centrohelids are free-living heterotrophs characterized by distinctive radiating pseudopodia called axopodia, and are most often found in freshwater environments.

Cryptista contains the cryptomonads, algae best known for their red algal-derived plastids that retain a relict endosymbiont nucleus (i.e. nucleomorph). The group also contains heterotrophic flagellates katablepharids and the lone genus *Palpitomonas*.

Archaeplastids are defined by the presence of primary plastids directly derived from endosymbiosis with a cyanobacterium. This includes green algae (from which land plants emerge), the red algae, and glaucophytes.

Amoebozoans are the second primarily amoeboid group, including groups with classical loose pseudopodia for feeding and locomotion. The group also includes slime molds and flagellates, as well as some important pathogens (e.g. *Entamoeba*).

Opisthokonts includes several protists, but also more familiar multicellular animals and fungi. Opisthokonts are defined by a single posterior flagellum, and protist opisthokonts include a variety of heterotrophic flagellates, amoebae, and parasites. Opisthokonts, Amoebozoans, and a few other small lineages are grouped together as **Amorphea**.

CRuMS is an amalgamation of several 'orphan' taxa: the Collodictyonids, Rigifilida, and Mantamonas. These are all free-living protists without much in common morphologically (i.e. swimming flagellates, filose amoeboid cells, and gliding cells), but which branch together in molecular phylogenies.

Discobids and **Metamonads** are two groups referred to as 'excavate' lineages, previously (with malawimonads) classified as the Excavate supergroup based on a distinctive morphology. Molecular phylogenies have mostly failed to support this, so they are treated as two separate and possibly related groups. Discobids includes photosynthetic euglenids (e.g. *Euglena*) parasites (e.g. *Trypanosoma*) and many free-living heterotrophic flagellates. Matamonads contains anaerobic protists, including several pathogens (e.g. *Giardia*, *Trichomonas*) and symbionts in many animal guts (e.g. *Trichonympha*).

'supergroups' to reflect their 'larger than kingdom' nature [36,37]. In Figure 2, we synthesize many phylogenomic studies into one schematic tree representing (we hope) a relatively uncontroversial consensus of the current state of eukaryotic relationships. The idea of supergroups remains the currency of eukaryotic relationships at their highest levels. Unlike the old Kingdoms of life, supergroups are mostly known by esoteric names unheard of outside a small group of experts. Their composition can be opaque, even if they include lineages we all know well. For example, we know humans are animals, but at various higher levels they are also holozoans, opisthokonts, or obazoans (Figure 3).

Considering how many revisions have been made to the tree of eukaryotes in the past, we certainly should ask whether the current, phylogenomic-based tree is also just a blip, soon-to-be-replaced by yet a different view. Anyone recalling the history of this field would be reluctant to rule this out, but the recent history also contains cause for optimism that has not always been evident. Phylogenetic methods, evolutionary models, the quantity and quality of data, and computational power available at a given time all influence the reconstruction of the tree, so change over time is expected. The question is what kind of changes have we observed as the field evolved? To answer this question, we can look back on how the tree was described 15 years ago, when the supergroup concept took hold, to see what has changed.

Changes to the tree fall under two general themes. The first and most obvious theme is that the tree grew through the addition of more taxa [38–45]. This takes the form of re-sampled groups that we already knew about, but also newly discovered groups (usually consisting of just a few species or genera, see below for more details). Sometimes these new taxa proved difficult to resolve into the bigger picture of supergroups, but in other cases they fell immediately into place. The second major theme

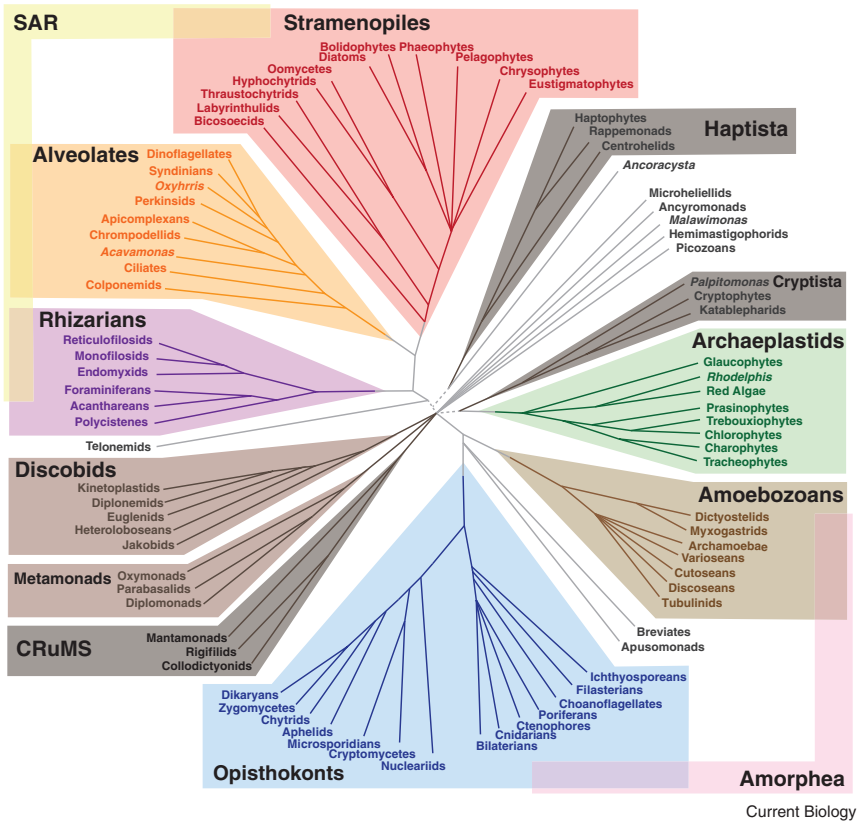

Figure 2. The eukaryotic tree of life.
A schematic tree of eukaryotes based on a consensus of phylogenomic studies together with morphological and cell biological information. The more uncontentious supergroups are boxed in colour, while more contentious ones are boxed in grey. Completely unresolved relationships are left unresolved, and some particularly contentious proposals for deep-branches are indicated by dashed lines. Wherever possible, informal names are used.

is the coalescence or shuffling of lineages; more data and additional analyses resulting in two or more major groups branching together, forming a new, larger group [46–52]. Neither of these two kinds of change are mutually exclusive, nor necessarily contradicts the original structure of the tree. Thus, much of the overall structure of the trees as conceived 15 years ago is consistent with today's view, but today's tree is better resolved, better fleshed out, and contains many groups that were not part of the original supergroup schemes. For example, where we once had alveolates and stramenopiles together (as part of a larger group called chromalveolates, discussed below), the addition of rhizarians led to the SAR supergroup [46,47], and the subsequent addition of telonemids led to the TSAR supergroup [52], all through processes of growth and coalescence. The same processes have been repeated in other parts of the tree: obazoans grew from the addition of breviates and apusomonads to opisthokonts [40]; CRuMS is an unforeseen collection of orphan flagellated or filose protists that have recently coalesced [51]; haptistans represent the merging of long-known haptophytes and centrohelids [50]. These changes will have to withstand repeated tests, ideally involving independent datasets and the addition of more taxa, to become fully accepted in mainstream evolutionary schemes. But the way the tree has changed over time suggests that a clearly describable tree underlays the evolution of eukaryotes over long time spans.

Controversial Branches in the Tree of Eukaryotes

The changes to the tree outlined above collectively represent numerous augmentations and amalgamations. In contrast, changes in the form of reversals from what we thought were well-resolved branches have been relatively rare. This is not to say that there has been no controversy, dead ends, or misleading results. Indeed, a number of longstanding questions remain in a fog even after repeated direct analyses using very large data sets and the best available methods. In general, we can break down the remaining outstanding questions about the eukaryotic tree into two general categories: the parts of the tree that remain uncertain despite close examination (the controversial branches discussed below), and the parts of the tree we know are uncertain because we lack the data to address them (the missing branches discussed in the next section).

One very controversial branch on the tree, but which is not always recognized as such, is the archaeplastids. This group has been included in broad evolutionary schemes of eukaryotic diversity for decades, but continues to lack comprehensive support from most molecular trees. Unlike any other current supergroup, the archaeplastids are anchored in an evolutionary hypothesis supported by solid morphological, molecular, and biochemical commonalities of their primary plastids [53,54]. Molecular phylogenies based on plastid data and other molecular features common to these plastids have from the beginning strongly supported their single origin [53–56]. However, to this day no phylogeny using nuclear data for a sufficiently broad taxon-sampling has provided unambiguous strong support for the monophyly of the group, or any consistent alternative (they are most often interrupted by cryptists, but not always and not in a consistent position [49,50,52]). This puts archaeplastids in a peculiar position: it has remained a supergroup throughout all recent iterations of the tree, but the consistency of evidence supporting it are still crucially missing.

Another hotspot for debate is the excavates, a supergroup that has been included in most schematics of eukaryotic diversity for many years, but is absent from Figure 2. The idea that 'excavate' organisms formed a supergroup was originally based on their morphology, or specifically a particular feeding groove found in many enigmatic protists [57]. However, a monophyletic excavate lineage was never really fully supported in single gene trees or phylogenomics, or only after some tweaking [48,58,59], and instead recent analyses split

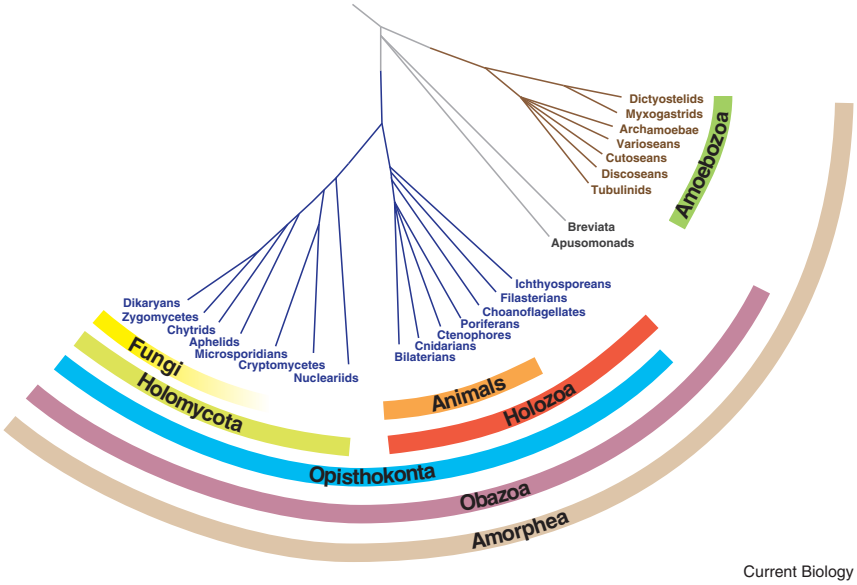

Figure 3. Naming supergroups.
Supergroup names have subsumed traditional kingdoms of life. This is a portion of the tree surrounding the familiar kingdoms, animals and fungi (formally Metazoa and Fungi). Under the current hierarchical names above the kingdom level, the familiar animals are also known by their less familiar, larger formal group names, Holozoa, Holomycota, Opisthokonta, Obazoa, or Amorphea.

them into three subgroups; the discobids, metamonads, and the small orphan group malawimonads [51,60]. At the same time, a similar feeding groove has been found in protists clearly branching in other supergroups [61,62], suggesting that the excavate condition might be ancestral to most or even all eukaryotes and not an indication of relatedness in one subgroup. Ultimately, the lack of comprehensive phylogenetic support left the excavate hypothesis mired in debate surrounding a few poorly studied taxa and the technical problems of reconstructing trees from the fast-evolving genes common to many excavate taxa.

A similar story unfolded for another former supergroup that we already mentioned, the chromalveolates. This supergroup was also proposed based on a uniting biological feature of plastids, in this case the common origin of a red algal-derived secondary plastid [63,64]. As with archaeplastids, data for the monophyly of the plastids soon appeared [65–67], but concurrent support from nuclear data never materialized. Instead, the main chromalveolate subgroups were split by the addition of previously unrelated lineages: alveolates and stramenopiles branched with rhizarians, altogether forming the SAR group (see above); and cryptomonads and haptophytes saw a collection of relatively small lineages coalescing around them, forming cryptists (*Palpitomonas*, katablepharids, cryptomonads) [42,68] and haptists (haptophytes, centrohelids) [43,50,69], respectively. Although these respective lineages are all robust, their position in the tree proved to be among the most controversial branches, and remain so today. Most frequently, haptists are related to the TSAR group, and cryptists to archaeplastids, but in both cases further evidence is required.

Finally, and perhaps the most vexing unsolved issue, is the root of the eukaryotic tree. The root is a particularly thorny question since it demands that we look even further back in time to the relationships between eukaryotes, archaea, and bacteria. This problem has been tackled by traditional phylogenetic reconstruction based on single or genome-level sampling, simply analyzing genes shared between eukaryotes and prokaryotes, or more recently mitochondria and their proteobacterial sisters [70–72]. Other phylogenetic tricks have been used, for instance trying to identify deep splits in the eukaryotes indicated by gene fusions, gene duplications, new gene families, or insertions and deletions in proteins, but none of these has converged towards a widely-accepted position for the root. The absence of a clear root also affects our understanding of the earliest splits in the tree, so until we have a better idea of where the root lies it will be impossible to fully reconstruct how the supergroups are related to one another.

Missing Branches in the Tree of Eukaryotes

The controversies listed above can be thought of as problems that we have identified and affect specific parts of the tree. But we also know there are problems that are more difficult to enumerate with current data, and in some cases we can't even predict exactly what part of the tree they affect. These issues come from taxa that are currently missing in the tree.

More than a century of classical microscopy studies has provided us with numerous accounts, sometimes very detailed, of protists that need to be re-visited to add to phylogenomic trees [73]. In some cases, we have solid expectations of where they will go, but there are also many 'orphan' taxa for which it is anyone's guess where they will eventually fall. Recent examples show that these orphans offer an exciting pool of untapped information. Adding centrohelids, telonemids, colponemids, and colpodellids to the tree added complexity to our understanding of the evolution of complex plastids [50,52,68,74,75]. Aphelids were found to hold important clues into the contentious early evolution of fungi and their close relatives, the microsporidian parasites and the still-mysterious rozellids/cryptomycetes [76,77]. Other lineages, like hemimastigophorids, or diphyllateans, rigifilids, and mantamonads have also recently been added to phylogenomic trees, possibly forming completely new branches [45,51]. Simply skimming reference volumes of protist diversity yields a long list of other potential candidates worthy of re-discovery [73]. Many of these orphans are likely isolated genera with unique features, rather than large and diverse lineages, but it is still possible that they form the tip of a diversity iceberg (see below section on lopsided trees). Either way, each new lineage added to the tree is an important piece of the puzzle to reconstruct how eukaryotic diversity evolved that will hopefully spark someone's curiosity.

The above lineages were all known from microscopy and their importance really emerged with molecular phylogenetic analyses, but there is also huge potential for discovery from the opposite direction — discovering lineages by molecular means. Environmental DNA or RNA sequencing has been a major source of discovery for bacterial and archaeal diversity, revealing dozens of environmental phyla [11]. Environmental surveys of eukaryotes have yielded somewhat different results, probably because the morphological record had allowed more eukaryotic diversity to be recognized prior to molecular surveys. Indeed, environmental DNA sequencing of eukaryotes has yielded massive expansions of diversity (sometimes major unknown groups) but generally within or related to already established supergroups [15,78–80]. These environmental lineages are still very important, especially as some represent very diverse and ecologically significant groups that we had no inkling even existed, or at least had little appreciation of their ecological relevance. Marine samples in particular have revealed several widespread and highly abundant lineages like MALVs, MASTs, MAOPs (marine alveolates, stramenopiles, and opisthokonts, respectively), eupelagonemids [78,79], and dozens of cases of smaller new subgroups related to virtually every major protist lineage known [81,82]. In soils, several groups of gregarine apicomplexan parasites of invertebrates were shown to dominate the diversity of protists [80], and most likely similar observations will be made everywhere else we look.

Lastly, an unanticipated challenge to reconstructing a complete tree of eukaryotes has emerged from recent discoveries of high-ranking lineages that were presaged by neither morphology nor environmental DNA surveys. These are new taxa that were discovered by the doggedly traditional means of exploring the environment and characterizing new organisms at the morphological and molecular levels, either through establishing cultures or by single cell microscopy and genome-wide sequencing. Some of these new taxa are so different from known lineages as to prove challenging to position in the tree, even with phylogenomic data, and remain without any obvious closest relatives (e.g. *Picozoa*, *Ancoracysta*) [44,83–85]. Others occupy crucial phylogenetic positions close to established groups (e.g., *Chromera*, *Vitrella*, *Acavamonas*, *Rhodelphis*, *Palpitomonas*) [42,74,86–88]. Each of these taxa has provided new perspectives on how major lineages or evolutionary transitions originated or unfolded, affecting how we interpret the evolution of mitochondria, plastids, the cytoskeleton, the endomembrane system, as well as processes as diverse as parasitism, photosynthesis, or mixotrophy. But they also shed equally important light on the depth of methodological biases in molecular diversity surveys, and the importance of multiple approaches, which will be outlined in the next section.

Biased Both Ways

A lot has been said about the extreme biases inherent in microbial cultivation methods. Indeed, the famous ‘great plate count anomaly’, which argued that only a small fraction of natural bacterial diversity can be cultivated [89], ultimately led to the widespread use of environmental DNA and metagenomics to circumvent this bias and measure microbial diversity directly from the environment. Cultivating protists comes with at least as many problems as bacteria, and so shares many of the same biases: we are better at cultivating algae, so they are over-represented in culture collections, whereas free-living heterotrophs are badly under-represented [90]. Since the widespread adoption of environmental DNA for surveying both prokaryotic and eukaryotic microbes, the inherent biases of this approach have been examined, but most of this work has taken the form of comparing one environmental method against others. For example, issues surrounding the choice of PCR primers, PCR-independent methods such as miTags and metagenomics, or using DNA versus RNA as template have all been tested [91]. But testing the biases of the environmental approach against entirely molecular-independent methods, like culturing and microscopy, is much more difficult. This requires a strong body of morphological information and culturing expertise to cross reference with molecular surveys. Without such comparisons, it is hard to say whether molecular surveys in general have systematic biases leading us to overlook whole lineages, as we know it to be the case with cultivation.

A growing body of evidence suggests that environmental DNA, when applied globally, is indeed missing large chunks of the diversity. Most obviously, some groups known to be relatively abundant in some environments are systematically poorly represented in environmental DNA surveys of those same environments. For example, certain lifestyles, like some forms of parasitism or symbiosis in general, are likely not represented as much as they should be. Indeed, if you filter all the animals out of an environment and survey the remaining protist diversity, it is not surprising that parasites or other symbionts that spend little or no time outside of their host will be under-represented. This sounds trivial, but a major fraction of protist diversity corresponds to parasitic taxa, and many diverse and common parasitic lineages have escaped detection in exactly this way. Rhizarian microcell parasites are an excellent example. One species of these tiny intracellular parasites, *Mikrocytos mackini*, has long been known from histopathology to be common in food oysters [92], but corresponding sequences were consistently absent from environmental surveys, even oyster habitats. However, a survey of diverse marine invertebrates with a combination of specific primers eventually revealed a widespread lineage of diverse parasites (the mikrocytids), none of which appears in environmental DNA surveys of environmental samples with taxonomically broad primers [81]. Their invisibility in molecular surveys was the result of a bias against their divergent rRNA sequence (and this is a common feature of many parasites), compounded by a bias against their obligate intracellular lifestyle. Thus, to adequately reveal the phylogenetic diversity of such a group by environmental DNA methods, one must target not only the right habitat (inside marine invertebrates), but also use lineage-specific PCR primers, all of which is predicated on a substantial body of prior knowledge. Organisms like these are not going to be detected and characterized in broad surveys of environmental diversity; they must be deliberately looked for.

While the current trend is towards molecular methods and their high-throughput nature (for good reasons), there is also increasing evidence that what’s currently missed by these methods is highly relevant. Imagine for a moment that you are visiting the Earth from another planet to investigate biodiversity, and randomly select 1000 animals from a patch of the Serengeti.

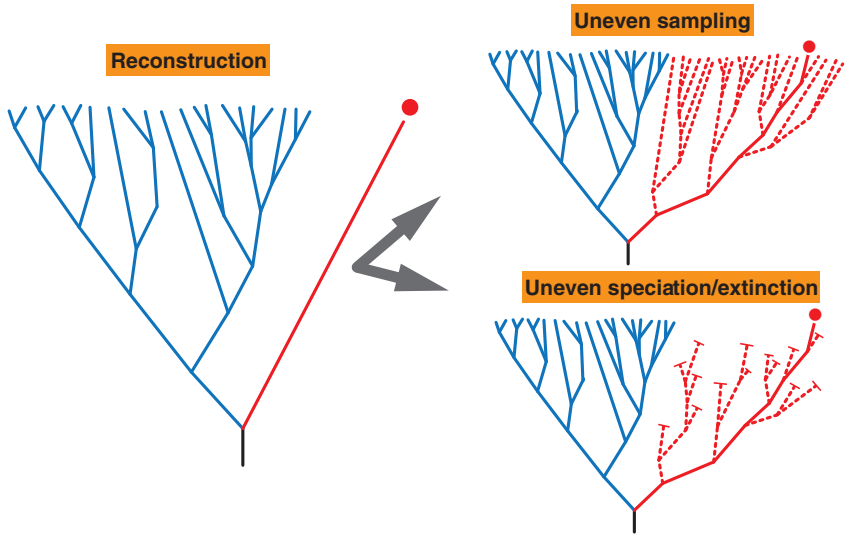

Figure 4. Lopsided trees.
One result of a current wave of discovery is that more nodes of the tree are becoming lopsided — on one side of a node is a long-studied branch of great diversity with thousands of known species, but on the other side is a single genus or species (left side). There are two main explanations for this observation. The sampling may be uneven (right side, top), such that the second lineage is also large and diverse but has mostly escaped our notice. In this case, continued exploration will yield a great number of potentially interesting new taxa. The other possibility is that there has been uneven evolutionary processes of speciation and/or extinction (right side, bottom), such that only a few relict branches remain to be discovered. In this case the few remaining taxa are potentially very important for understanding the evolution of their well-studied sisters.

You may come to the correct conclusion that wildebeest and gazelles are common, but you may not realize that lions and leopards even exist, or you may write them off as insignificant 'rare biosphere' taxa. But, of course, they play a crucial role in the ecosystem and represent a significant branch in the tree of animals. Global surveys of protist diversity have thus far concentrated mainly on what's abundant, although the rare microbial biosphere is increasingly being scrutinized in more detail [93–95]. It is possible that because such rare species are lying low in all environments, some of the most surprising recent results are coming from a re-appreciation of 'old' culture- and microscopy-based biodiversity surveys [44,45,74]. These low-throughput approaches are not free of bias, but they have *different* biases and so capture diversity that is overlooked by high-throughput molecular methods. As a result, some of the most interesting protist lineages described in recent years have been found through the often-undervalued work of 'classical' protistology, or just by chance, and are absent, rare, or disregarded in even very large environmental DNA databases. Biases against such lineages may be in part due to their status as microbial predators, expected to be relatively rare in nature, but this no more means they are unimportant than it does of lions or leopards.

Increasingly Lopsided Trees

An interesting implication of each new discovery of a protist species that falls outside of a known phylum, kingdom, or supergroup is that large parts of the tree of eukaryotes are becoming lopsided. By this we mean that on one side of a node is a large, diverse, and well-studied group that contains hundreds or thousands of described species, but on the other side of the same node is a single species or genus. They diverged from a common ancestor at the same time, so why is the diversity not more uniformly distributed? One answer may be that the diversity is in fact uniform, but we have only scratched the surface of this new lineage and a lot more is out there waiting to be discovered (Figure 4). This argument has been borne out for several such branches on the tree of eukaryotes, sometimes with far-reaching evolutionary implications. For example, the genus *Rozella* was formerly the lone sister to all 'classical' fungi [96], but we now know it to be just the first identified representative of several paraphyletic and diverse lineages (rozellids/cryptomycetes, aphelids, microsporidians) that have reshaped our views of early fungal evolution [76,77,97]. Similarly, the only known sisters to the diverse phylum of apicomplexan parasites were a curious pair, the photosynthetic chromerids and the predatory colpodellids [86,98], but these two are now seen instead as the few cultured representatives of a diverse 'chrompodellid' branch on the tree encompassing substantial ecological and evolutionary diversity [75].

The alternative to the 'more will come' situation described above is that the tree really is lopsided: one descendent of this ancestor diversified a lot more successfully than the other. This explanation might intuitively seem far-fetched, but there are examples known from multicellular lineages where it seems unlikely that vast fractions of the diversity have been overlooked. *Amborella*, for example, is a mono-specific genus found only on the island of New Caledonia, which appears to be sister to all other flowering plants [99]. This species is widely recognized as something of a treasure since it alone holds many clues into plant evolution. But the tree of eukaryotes quite possibly includes many such treasures scattered throughout the less-studied protist branches. If their 'lone-branch' status holds up, organisms like *Telonema*, *Rhodelphis*, *Ancoracysta*, *Acavomonas*, or *Palpitomonas* will be every bit as informative about the evolution of major lineages as *Amborella* is for plants.

Concluding Remarks

Probably the most significant conclusion to emphasize is that there are many reasons for optimism that a tree of eukaryotes exists to be found, and that the phylogenomic approach is progressing in the right direction. This is significant, because it suggests that different parts of the tree of life have been shaped by various evolutionary processes in different ways. Over the last decade and more, the primary changes we have seen in the tree of eukaryotes have been what we might call 'constructive'; new branches have been added, and most often found a well-supported place in the tree, and old branches have merged into

fewer, larger assemblages. Notable remaining problems include a few key taxa, the relationships between the largest assemblages, and the position of the root. Some of these may be beyond our ability to resolve, only time will tell, but based on the general trends we are optimistic that progress will continue to be made.

Less obvious, but equally important to emphasize, is the importance of exploration and the discovery of new lineages. We need to look beyond molecular data and recognize the major advances coming from simply looking at what's out there in the environment, and getting new species into culture. It is illuminating that the biggest surprises from this direction come from new organisms that are at first glance perhaps not especially interesting looking. The world is full of small flagellates eating bacteria and other flagellates that rarely catch our eye. But these are what the ancestor of all modern eukaryotes looked like, and arguably represent the greatest fraction of modern eukaryotic diversity. They are hard to identify, hard to isolate, hard to culture, but worth the effort.

ACKNOWLEDGEMENTS

P.J.K.'s research is supported by grants from the Natural Sciences and Engineering Research Council of Canada, the Hakai Institute, and the Gordon and Betty Moore Foundation. F.B.'s research is supported by a Fellowship grant from SciLifeLab, a research project grant from the Swedish Research Council (VR-2017-04563), and a Future Research Leader grant from Formas (2017-01197).

REFERENCES

1. Baum, D.A. (2005). The tree-thinking challenge. Science *310*, 979–980.
2. Omland, K.E., Cook, L.G., and Crisp, M.D. (2008). Tree thinking for all biology: the problem with reading phylogenies as ladders of progress. Bioessays *30*, 854–867.
3. Lovejoy, A.O. (1936). The Great Chain of Being: A Study of the History of an Idea, (Cambridge: Harvard University Press).
4. Dayrat, B. (2003). The roots of phylogeny: how did Haeckel build his trees? Syst. Biol. *52*, 515–527.
5. Pamilo, P., and Nei, M. (1988). Relationships between gene trees and species trees. Mol. Biol. Evol. *5*, 568–583.
6. Nichols, R. (2001). Gene trees and species trees are not the same. Trends Ecol. Evol. *16*, 358–364.
7. Bapteste, E. (2014). The origins of microbial adaptations: how introgressive descent, egalitarian evolutionary transitions and expanded kin selection shape the network of life. Front. Microbiol. *5*, 83.
8. Corel, E., Lopez, P., Méheust, R., and Bapteste, E. (2016). Network-thinking: graphs to analyze microbial complexity and evolution. Trends Microbiol. *24*, 224–237.
9. Doolittle, W.F. (1999). Phylogenetic classification and the universal tree. Science *284*, 2124–2129.
10. Doolittle, W.F., and Bapteste, E. (2007). Pattern pluralism and the Tree of Life hypothesis. Proc. Natl. Acad. Sci. USA *104*, 2043–2049.
11. Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., Butterfield, C.N., Hernsdorf, A.W., Amano, Y., Ise, K., *et al*. (2016). A new view of the tree of life. Nat. Microbiol. *1*, 16048.
12. O'Malley, M.A., Simpson, A.G.B., and Roger, A.J. (2012). The other eukaryotes in light of evolutionary protistology. Biol. Philos. *28*, 299–330.
13. Betts, H.C., Puttick, M.N., Clark, J.W., Williams, T.A., Donoghue, P.C.J., and Pisani, D. (2018). Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. Nat. Ecol. Evol. *17*, 1–10.
14. Keeling, P.J. (2019). Combining morphology, behaviour, and genomics to understand the evolution and ecology of microbial eukaryotes. Philos. Trans. R. Soc. Lond. B, *in press*.
15. Keeling, P.J., and Campo, J.D. (2017). Marine protists are not just big bacteria. Curr. Biol. *27*, R541–R549.
16. Haeckel, E.H.P.A. (1866). Generelle Morphologie der Organismen : allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von C. Darwin reformirte Decendenz-Theorie (Berlin: Druck und Verlag von Georg Reimer).
17. Copeland, H.F. (1938). The kingdoms of organisms. Quart. Rev. Biol. *13*, 383–420.
18. Whittaker, R.H. (1969). New concepts of kingdoms of organisims. Science *163*, 150–163.
19. Sogin, M.L. (1989). Evolution of eukaryotic microorganisms and their small subunit ribosomal RNAs. Am. Zool. *29*, 487–499.
20. Sogin, M.L., Elwood, H.J., and Gunderson, J.H. (1986). Evolutionary diversity of eukaryotic small-subunit rRNA genes. Proc. Natl. Acad. Sci. USA *83*, 1383–1387.
21. Vossbrinck, C.R., Maddox, J.V., Friedman, S., Debrunner-Vossbrinck, B.A., and Woese, C.R. (1987). Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature *326*, 411–414.
22. Sogin, M.L., Gunderson, J.H., Elwood, H.J., Alonso, R.A., and Peattie, D.A. (1989). Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from*Giardia lamblia*. Science *243*, 75–77.
23. Keeling, P.J., and Doolittle, W.F. (1996). Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. Mol. Bio. Evol. *13*, 1297–1305.
24. Baldauf, S.L. (1999). A search for the origins of animals and fungi: comparing and combining molecular data. Am. Nat. *154*, S178–S188.
25. Roger, A.J., Sandblom, O., Doolittle, W.F., and Philippe, H. (1999). An evaluation of elongation factor 1 alpha as a phylogenetic marker for eukaryotes. Mol. Biol. Evol. *16*, 218–233.
26. Roger, A.J. (1999). Reconstructing early events in eukaryotic evolution. Am. Nat. *154*, S146–S163.
27. Baldauf, S.L., Roger, A.J., Wenk-Siefert, I., and Doolittle, W. (2000). A kingdom-level phylogeny of eukaryotes based on combined protein data. Science *290*, 972–977.
28. Gajadhar, A.A., Marquardt, W.C., Hall, R., Gunderson, J., Ariztia-Carmona, E.V., and Sogin, M.L. (1991). Ribosomal RNA sequences of *Sarcocystis muris, Theileria annulata* and *Crypthecodinium cohnii* reveal evolutionary relationships among apicomplexans, dinoflagellates, and ciliates. Mol. Biochem. Parasitol. *45*, 147–154.
29. Bhattacharya, D., Helmchen, T., and Melkonian, M. (1995). Molecular evolutionary analyses of nuclear-encoded small subunit ribosomal RNA identify an independent rhizopod lineage containing the Euglyphidae and the Chlorarachniophyta. J. Eukaryot. Microbiol. *42*, 64–68.
30. Keeling, P.J. (2001). Foraminifera and Cercozoa are related in actin phylogeny: two orphans find a home? Mol. Biol. Evol. *18*, 1551–1557.
31. Nikolaev, S.I., Berney, C., Fahrni, J.F., Bolivar, I., Polet, S., Mylnikov, A.P., Aleshin, V.V., Petrov, N.B., and 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.
32. Delsuc, F., Brinkmann, H., and Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. *6*, 361–375.
33. Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., and Baurain, D. (2011). Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. *9*, e1000602.
34. Chan, C.X., and Ragan, M.A. (2013). Next-generation phylogenomics. Biol. Direct *8*, 3.

35. Burki, F. (2014). The Eukaryotic Tree of Life from a global phylogenomic perspective. Cold Spring Harb. Perspect. Biol. *6*, a016147–a016147.

36. Simpson, A.G.B., and Roger, A.J. (2004). The real "kingdoms" of eukaryotes. Curr. Biol. *14*, R693–696.

37. Keeling, P.J., Burger, G., Durnford, D.G., Lang, B.F., Lee, R.W., Pearlman, R.E., Roger, A.J., and Gray, M.W. (2005). The tree of eukaryotes. Trends Ecol. Evol. *20*, 670–676.

38. Minge, M.A., Silberman, J.D., Orr, R.J., Cavalier-Smith, T., Shalchian-Tabrizi, K., Burki, F., Skjaeveland, A., and Jakobsen, K.S. (2009). Evolutionary position of breviate amoebae and the primary eukaryote divergence. Proc. Biol. Sci. *276*, 597–604.

39. Zhao, S., Burki, F., Brate, J., Keeling, P.J., klaveness, D., and Shalchian-Tabrizi, K. (2012). Collodictyon–an ancient lineage in the Tree of Eukaryotes. Mol. Biol. Evol. *29*, 1557–1568.

40. Brown, M.W., Sharpe, S.C., Silberman, J.D., Heiss, A.A., Lang, B.F., Simpson, A.G.B., and Roger, A.J. (2013). Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. Proc. R. Soc. B *280*, 20131755.

41. Burki, F., Corradi, N., Sierra, R., Pawlowski, J., Meyer, G.R., Abbott, C.L., and Keeling, P.J. (2013). Phylogenomics of the intracellular parasite Mikrocytos mackini reveals evidence for a mitosome in rhizaria. Curr. Biol. *23*, 1541–1547.

42. Yabuki, A., Kamikawa, R., Ishikawa, S.A., Kolisko, M., Kim, E., Tanabe, A.S., Kume, K., Ishida, K.-I., and Inagaki, Y. (2014). *Palpitomonas bilix* represents a basal cryptist lineage: insight into the character evolution in Cryptista. Sci. Rep. *4*, 4641.

43. Cavalier-Smith, T., Chao, E.E., and Lewis, R. (2015). Multiple origins of Heliozoa from flagellate ancestors: New cryptist subphylum Corbihelia, superclass Corbistoma, and monophyly of Haptista, Cryptista, Hacrobia and Chromista. Mol. Phylogenet. Evol. *93*, 331–362.

44. Janouskovec, J., Tikhonenkov, D.V., Burki, F., Howe, A.T., Rohwer, F.L., Mylnikov, A.P., and Keeling, P.J. (2017). A new lineage of eukaryotes illuminates early mitochondrial genome reduction. Curr. Biol. *27*, 3717–3724.e3715.

45. Lax, G., Eglit, Y., Eme, L., Bertrand, E.M., Roger, A.J., and Simpson, A.G.B. (2018). Hemimastigophora is a novel supra-kingdom-level lineage of eukaryotes. Nature *564*, 410–414.

46. Burki, F., Shalchian-Tabrizi, K., Minge, M., Skjaeveland, A., Nikolaev, S.I., Jakobsen, K.S., and Pawlowski, J. (2007). Phylogenomics reshuffles the eukaryotic supergroups. PLoS One *2*, e790.

47. Hackett, J.D., Yoon, H.S., Li, S., Reyes-Prieto, A., Rummele, S.E., and Bhattacharya, D. (2007). Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. Mol. Biol. Evol. *24*, 1702–1713.

48. Rodriguez-Ezpeleta, N., Brinkmann, H., Burger, G., Roger, A.J., Gray, M.W., Philippe, H., and Lang, B.F. (2007). Toward resolving the eukaryotic tree: the phylogenetic positions of jakobids and cercozoans. Curr. Biol. *17*, 1420–1425.

49. Baurain, D., Brinkmann, H., Petersen, J., Rodriguez-Ezpeleta, N., Stechmann, A., Demoulin, V., Roger, A.J., Burger, G., Lang, B.F., and Philippe, H. (2010). Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. Mol. Biol. Evol. *27*, 1698–1709.

50. Burki, F., Kaplan, M., Tikhonenkov, D.V., Zlatogursky, V., Minh, B.Q., Radaykina, L.V., Smirnov, A., Mylnikov, A.P., and Keeling, P.J. (2016). Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. Proc. R. Soc. B *283*, 20152802.

51. Brown, M.W., Heiss, A.A., Kamikawa, R., Inagaki, Y., Yabuki, A., Tice, A.K., Shiratori, T., Ishida, K.-I., Hashimoto, T., Simpson, A.G.B., *et al*. (2018). Phylogenomics places orphan protistan lineages in a novel eukaryotic super-group. Gen. Biol. Evol. *10*, 427–433.

52. Strassert, J.F.H., Jamy, M., Mylnikov, A.P., Tikhonenkov, D.V., and Burki, F. (2019). New phylogenomic analysis of the enigmatic phylum Telonemia further resolves the Eukaryote Tree of Life. Mol. Biol. Evol. *36*, 757–765.

53. Palmer, J.D. (2003). The symbiotic birth and spread of plastids: how many times and whodunit? J. Phycol. *39*, 1–9.

54. Keeling, P.J. (2010). The endosymbiotic origin, diversification and fate of plastids. Philos. Trans. R. Soc. Lond. B *365*, 729–748.

55. Hagopian, J.C., Reis, M., Kitajima, J.P., Bhattacharya, D., and de Oliveira, M.C. (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.

56. Rodriguez-Ezpeleta, N., Brinkmann, H., Burey, S.C., Roure, B., Burger, G., Loffelhardt, W., Bohnert, H.J., Philippe, H., and Lang, B.F. (2005). Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr. Biol. *15*, 1325–1330.

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

58. Hampl, V., Hug, L., Leigh, J.W., Dacks, J.B., Lang, B.F., Simpson, A.G.B., and Roger, A.J. (2009). Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". Proc. Natl. Acad. Sci. USA *106*, 3859–3864.

59. Takishita, K., Kolisko, M., Komatsuzaki, H., Yabuki, A., Inagaki, Y., Cepicka, I., Smejkalová, P., Silberman, J.D., Hashimoto, T., Roger, A.J., *et al*. (2012). Multigene phylogenies of diverse Carpediemonas-like organisms identify the closest relatives of "Amitochondriate" diplomonads and retortamonads. Protist *163*, 344–355.

60. Heiss, A.A., Kolisko, M., Ekelund, F., Brown, M.W., Roger, A.J., and Simpson, A.G.B. (2018). Combined morphological and phylogenomic re-examination of malawimonads, a critical taxon for inferring the evolutionary history of eukaryotes. R. Soc. Open Sci. *5*, 171707.

61. Yubuki, N., and Leander, B.S. (2013). Evolution of microtubule organizing centers across the tree of eukaryotes. Plant J. *75*, 230–244.

62. Tikhonenkov, D.V., Janouskovec, J., Mylnikov, A.P., Mikhailov, K.V., Simdyanov, T.G., Aleoshin, V.V., and Keeling, P.J. (2014). Description of Colponema vietnamica sp.n. and Acavomonas peruviana n. gen. n. sp., two new alveolate phyla (Colponemidia nom. nov. and Acavomonidia nom. nov.) and their contributions to reconstructing the ancestral state of alveolates and eukaryotes. PLoS One *9*, e95467.

63. 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. Eukaryot. Microbiol. *46*, 347–366.

64. Keeling, P.J. (2009). Chromalveolates and the evolution of plastids by secondary endosymbiosis. J. Eukaryot Microbiol. *56*, 1–8.

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

66. Yoon, H.S., Hackett, J.D., Pinto, G., and Bhattacharya, D. (2002). A single, ancient origin of the plastid in the Chromista. Proc. Natl. Acad. Sci. USA *99*, 15507–15512.

67. Janouskovec, J., Horak, A., Obornik, M., Lukes, J., and Keeling, P.J. (2010). A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc. Natl. Acad. Sci. USA *107*, 10949–10954.

68. Burki, F., Okamoto, N., Pombert, J.F., and Keeling, P.J. (2012). The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. Proc. Biol. Sci. *279*, 22406–2254.

69. Cavalier-Smith, T., and von der Heyden, S. (2007). Molecular phylogeny, scale evolution and taxonomy of centrohelid heliozoa. Mol. Phylogenet. Evol. *44*, 1186–1203.

70. Derelle, R., and Lang, B.F. (2012). Rooting the Eukaryotic Tree with mitochondrial and bacterial proteins. Mol. Biol.Evol. *29*, 1277–1289.

71. He, D., Fiz-Palacios, O., Fu, C.-J., Fehling, J., Tsai, C.-C., and Baldauf, S.L. (2014). An alternative root for the eukaryote tree of life. Curr. Biol. *24*, 465–470.

72. Derelle, R., Torruella, G., Klimeš, V., Brinkmann, H., Kim, E., Vlček, Č., Lang, B.F., and Elias, M. (2015). Bacterial proteins pinpoint a single eukaryotic root. Proc. Natl. Acad. Sci. USA *112*, E693–699.

73. Lee, J., Leedale, G., and Bradbury, P. (2000). An Illustrated Guide to the Protozoa (Lawrence, Kansas: Allen Press Inc).

74. Janouskovec, J., Tikhonenkov, D.V., Mikhailov, K.V., Simdyanov, T.G., Aleoshin, V.V., Mylnikov, A.P., and Keeling, P.J. (2013). Colponemids represent multiple ancient alveolate lineages. Curr. Biol. *23*, 2546–2552.

75. Janouskovec, J., Tikhonenkov, D.V., Burki, F., Howe, A.T., Kolisko, M., Mylnikov, A.P., and Keeling, P.J. (2015). Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. Proc. Natl. Acad. Sci. USA *112*, 10200–10207.

76. James, T.Y., Pelin, A., Bonen, L., Ahrendt, S., Sain, D., Corradi, N., and Stajich, J.E. (2013). Shared signatures of parasitism and phylogenomics unite Cryptomycota and microsporidia. Curr. Biol. *23*, 1548–1553.

77. Torruella, G., Grau-Bové, X., Moreira, D., Karpov, S.A., Burns, J.A., Sebé-Pedrós, A., Völcker, E., and López-García, P. (2018). Global transcriptome analysis of the aphelid Paraphelidium tribonemae supports the phagotrophic origin of fungi. Commun. Biol. *1*, 231.

78. Pernice, M.C., Forn, I., Gomes, A., Lara, E., Alonso-Saez, L., Arrieta, J.M., del Carmen Garcia, F., Hernando-Morales, V., MacKenzie, R., Mestre, M., *et al*. (2015). Global abundance of planktonic heterotrophic protists in the deep ocean. ISME J. *9*, 782–792.

79. de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., *et al*. (2015). Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. Science *348*, 1261605.

80. Mahé, F., de Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., Singer, D., Mayor, J., Bunge, J., Sernaker, S., *et al*. (2017). Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. Nat. Ecol. Evol. *1*, 91.

81. Hartikainen, H., Stentiford, G.D., Bateman, K.S., Berney, C., Feist, S.W., Longshaw, M., Okamura, B., Stone, D., Ward, G., Wood, C., *et al*. (2014). Mikrocytids are a broadly distributed and divergent radiation of parasites in aquatic invertebrates. Curr. Biol. *24*, 807–812.

82. Freeman, M.A., Fuss, J., Kristmundsson, A., Bjorbaekmo, M.F.M., Mangot, J.F., Del Campo, J., Keeling, P.J., Shalchian-Tabrizi, K., and Bass, D. (2017). X-cells are globally distributed, genetically divergent fish parasites related to perkinsids and dinoflagellates. Curr. Biol. *27*, 1645–1651.e3.

83. Not, F., Valentin, K., Romari, K., Lovejoy, C., Massana, R., Tobe, K., Vaulot, D., and Medlin, L.K. (2007). Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. Science *315*, 253–255.

84. Yoon, H.S., Price, D.C., Stepanauskas, R., Rajah, V.D., Sieracki, M.E., Wilson, W.H., Yang, E.C., Duffy, S., and Bhattacharya, D. (2011). Single-cell genomics reveals organismal interactions in uncultivated marine protists. Science *332*, 714–717.

85. Seenivasan, R., Sausen, N., Medlin, L.K., and Melkonian, M. (2013). Picomonas judraskeda Gen. Et Sp. Nov.: The first identified member of the Picozoa Phylum Nov., a widespread group of picoeukaryotes, formerly known as 'Picobiliphytes'. PLoS One *8*, e59565.

86. Moore, R.B., Obornik, M., Janouskovec, J., Chrudimsky, T., Vancova, M., Green, D.H., Wright, S.W., Davies, N.W., Bolch, C.J., Heimann, K., *et al*. (2008). A photosynthetic alveolate closely related to apicomplexan parasites. Nature *451*, 959–963.

87. Obornik, M., Modry, D., Lukes, M., Cernotikova-Stribrna, E., Cihlar, J., Tesarova, M., Kotabova, E., Vancova, M., Prasil, O., and Lukes, J. (2012). Morphology, ultrastructure and life cycle of Vitrella brassicaformis n. sp., n. gen., a novel chromerid from the Great Barrier Reef. Protist *163*, 306–323.

88. Gawryluk, R.M.R., Tikhonenkov, D.V., Hehenberger, E., Husnik, F., Mylnikov, A.P., and Keeling, P.J. (2019). Non-photosynthetic predators are sisters to red algae. Nature. In Press.

89. Staley, J.T., and Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annu. Rev. Microbiol. *39*, 321–346.

90. del Campo, J., Sieracki, M.E., Molestina, R., Keeling, P., Massana, R., and Ruiz-Trillo, I. (2014). The others: our biased perspective of eukaryotic genomes. Trends Ecol. Evol. *29*, 252–259.

91. Leray, M., and Knowlton, N. (2016). Censusing marine eukaryotic diversity in the twenty-first century. Philos. Trans. R. Soc. Lond. B, 371.

92. Abbott, C.L., and Meyer, G.R. (2014). Review of Mikrocytos microcell parasites at the dawn of a new age of scientific discovery. Dis. Aquat. Organ *110*, 25–32.

93. Sogin, M.L., Morrison, H.G., Huber, J.A., Mark Welch, D., Huse, S.M., Neal, P.R., Arrieta, J.M., and Herndl, G.J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". Proc. Natl. Acad. Sci. USA *103*, 12115–12120.

94. Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., Claverie, J.-M., Decelle, J., Dolan, J.R., Dunthorn, M., *et al*. (2014). Patterns of rare and abundant marine microbial eukaryotes. Curr. Biol. *24*, 813–821.

95. Logares, R., Mangot, J.-F., and Massana, R. (2015). Rarity in aquatic microbes: placing protists on the map. Res. Microbiol. *166*, 831–841.

96. James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., *et al*. (2006). Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature *443*, 818–822.

97. Jones, M.D., Forn, I., Gadelha, C., Egan, M.J., Bass, D., Massana, R., and Richards, T.A. (2011). Discovery of novel intermediate forms redefines the fungal tree of life. Nature *474*, 200–203.

98. Kuvardina, O.N., Leander, B.S., Aleshin, V.V., Myl'nikov, A.P., Keeling, P.J., and Simdyanov, T.G. (2002). The phylogeny of colpodellids (Alveolata) using small subunit rRNA gene sequences suggests they are the free-living sister group to apicomplexans. J. Eukaryot. Microbiol. *49*, 498–504.

99. Drew, B.T., Ruhfel, B.R., Smith, S.A., Moore, M.J., Briggs, B.G., Gitzendanner, M.A., Soltis, P.S., and Soltis, D.E. (2014). Another look at the root of the angiosperms reveals a familiar tale. Syst. Biol. *63*, 368–382.