Predatory colponemids are the sister group to all other alveolates

Denis V. Tikhonenkov[a,b,*,1], Jürgen F.H. Strassert[c,d,*,1], Jan Janouškovec[e],
Alexander P. Mylnikov[a,2], Vladimir V. Aleoshin[f,g], Fabien Burki[h], Patrick J. Keeling[b]

[a] Papanin Institute for Biology of Inland Waters, Russian Academy of Sciences, Borok 152742, Russia
[b] Department of Botany, University of British Columbia, 6270 University Boulevard, Vancouver V6T1Z4, British Columbia, Canada
[c] Department of Organismal Biology, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden
[d] Institute of Biology, Free University of Berlin, Königin-Luise-Straße 1–3, 14195 Berlin, Germany
[e] Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, United Kingdom
[f] Belozersky Institute for Physicochemical Biology, Lomonosov Moscow State University, Leninskye Gory, House 1, Building 40, Moscow 119991, Russia
[g] Institute for Information Transmission Problems, Russian Academy of Sciences, Bolshoy Karetny per. 19, build.1, Moscow 127051, Russia
[h] Science for Life Laboratory and Department of Organismal Biology, Uppsala University, Norbyvägen 18D, Uppsala 75236, Sweden

ARTICLE INFO

ABSTRACT

Alveolates are a major supergroup of eukaryotes encompassing more than ten thousand free-living and parasitic species, including medically, ecologically, and economically important apicomplexans, dinoflagellates, and ciliates. These three groups are among the most widespread eukaryotes on Earth, and their environmental success can be linked to unique innovations that emerged early in each group. Understanding the emergence of these well-studied and diverse groups and their innovations has relied heavily on the discovery and characterization of early-branching relatives, which allow ancestral states to be inferred with much greater confidence. Here we report the phylogenomic analyses of 313 eukaryote protein-coding genes from transcriptomes of three members of one such group, the colponemids (Colponemidia), which support their monophyly and position as the sister lineage to all other known alveolates. Colponemid-related sequences from environmental surveys and our microscopical observations show that colponemids are not common in nature, but they are diverse and widespread in freshwater habitats around the world. Studied colponemids possess two types of extrusive organelles (trichocysts or toxicysts) for active hunting of other unicellular eukaryotes and potentially play an important role in microbial food webs. Colponemids have generally plesiomorphic morphology and illustrate the ancestral state of Alveolata. We further discuss their importance in understanding the evolution of alveolates and the origin of myzocytosis and plastids.

1. Introduction

Alveolates are one of the largest major groups of eukaryotes with over ten thousand described species and much larger yet undescribed species diversity based on environmental sequence data (Chambouvet et al., 2008; López-García et al., 2001; de Vargas et al., 2015). These organisms have a special type of cell coverings consisting of a plasmalemma and alveoli — abutting single-membrane flattened sacs, which probably derived from the endomembrane system and always subtend the plasma membrane (Gould et al., 2008). Most alveolates belong to one of three major groups: apicomplexans (e.g., the malaria

parasite *Plasmodium*), dinoflagellates (e.g., the coral endosymbiont *Symbiodinium*), and ciliates (e.g., the model organisms *Tetrahymena* and *Paramecium*). These three groups are among the most ecologically successful eukaryotes on Earth, and are also of evolutionary importance since each lineage also displays a number of distinctive innovations. For example, the apicomplexans infect virtually all known animals and this process is mediated by a complex suite of structures called the apical complex (Katris et al., 2014). Ciliates have evolved a large cell size and structures that make them effective predators, as well as a separation of germ and soma within a single cell. Dinoflagellates have evolved a variety of unique features, including several aspects of genomic

---

organization and function, as well as a wide range of trophic strategies including photosynthetic algae, predators, and parasites. The alveolates have also served as models to understand organelle evolution; mitochondrial genomes and respiratory chains are modified in each of the three groups in different ways (Smith et al., 2007; Nash et al., 2008), and apicomplexans and dinoflagellates (together with their closest relatives collectively known as myzozoans) contain plastids that have adopted diverse functional and genomic states (Janouškovec et al., 2017a).

Reconstructing the deep evolutionary transitions that gave rise to these three well-studied and diverse groups has relied heavily on the discovery and characterization of early-branching lineages, which allow ancestral states to be inferred with much greater confidence. Such deep-branching lineages have been particularly useful in analyses of both main myzozoan groups, apicomplexans and dinoflagellates: sister groups such as perkinsids, chrompodellids, *Digyalum*, and *Platyproteum,* have all shed light on the evolution of major characters (Goggin and Barker, 1993; Moore et al., 2008; Cavalier-Smith, 2014; Janouškovec et al., 2015, 2019; Mathur et al., 2019). But no equivalent sister group to ciliates is known, and most importantly, no group has unambiguously been shown to be sister to the alveolates as a whole (the next neighbour group, the stramenopiles, is distantly related and shares little in common with alveolates).

A handful of predatory protists have been discussed as possible candidates for deep-branching alveolates, based on sequences from one or a small number of genes (Janouškovec et al., 2013; Tikhonenkov et al., 2014; Park and Simpson, 2015). A six gene phylogeny suggested that *Acavomonas* is the sister taxon to myzozoans and *Colponema* is even deeper branching, perhaps as a sister lineage to all other alveolates (Janouškovec et al., 2013). The phylogeny of the small subunit ribosomal RNA gene (SSU rDNA) further showed that *Palustrimonas* and five clades of environmental sequences represent additional deep alveolate lineages (Janouškovec et al., 2013; Park and Simpson, 2015). Their relationships with *Acavomonas*, *Colponema*, ciliates, myzozoans are unresolved on SSU rDNA trees, but they all appear evolutionarily distinct (Tikhonenkov et al., 2014; Park and Simpson, 2015). Such diversity of deep-branching taxa could be instrumental in interpreting the alveolate origin and early evolution. In particular, an early-branching position for *Colponema* could be consistent with it possessing structures and behaviors thought to be plesiomorphic in the group (Mignot and Brugerolle, 1975; O'Kelly, 1993; Tikhonenkov et al., 2014).

Here we report the phylogenomic analyses of transcriptomes of three colponemid strains, which support their monophyly in the phylum Colponemidia, and indicate their position as the sister lineage to all other known alveolates. We further analyse the environmental diversity of colponemids and discuss their role in the ecosystems and importance in understanding the evolution of alveolates.

2. Material and methods

2.1. Collection, culture establishment and microscopy

Clones Colp-22, Colp-26, and Colp-tractor were obtained from samples collected from freshwater bodies in Ukraine (Seret river plankton; Dniester Canyon; 48°40′01.2″N, 25°51′04.2″E), Con Dao Island (Quang Trung lake; bottom detritus within lotuses; 8°41′31.2″N, 106°36′24.1″E), and British Columbia (City of Delta, Burns Bog; water with plant debris and detritus near a sunken tractor; 49°08′41.8″N 122°55′51.3″W), respectively. Colponemid clone Colp-10 was isolated from desert soil in Morocco (same location and sample as for *Moramonas marocensis* in Strassert et al., 2016); Colponemid clone Colp-15 was isolated from Suoi Da Reservoir in South Vietnam (same location and sample as for *Aquavolon hoantrani* in Bass et al., 2018). Other colponemid strains were obtained as described previously (Janouškovec et al., 2013, Tikhonenkov et al., 2014). Following isolation by a glass micropipette, all novel strains were propagated on the bodonid

Parabodo caudatus strain BAS-1, which was grown in spring water (Aqua Minerale, PepsiCo or PC Natural Spring Water, President's Choice) using the bacterium *Pseudomonas fluorescens* as food. Microscopy observations were made using a Zeiss AxioScope A.1 light microscope and JEM-1011 (JEOL) transmission electron microscope as described previously (Tikhonenkov et al., 2014).

3. DNA and RNA extraction, sequencing, and assembling

Cells of strains Colp-7a, Colp-10, Colp-15, Colp-22, Colp-26, and Colp-tractor were grown in clonal laboratory cultures and were harvested following peak abundance after eating most of the prey. Cells were collected by centrifugation (2000g, room temperature) on the 0.8 μm membrane of Vivaclear Mini columns (Sartorium Stedim Biotech Gmng, Germany, Cat. No. VK01P042). Total RNA of strains Colp-7a, Colp-10, Colp-15 was extracted using the RNAqueous®-Micro Kit (ThermoFisher Scientific, cat. # AM1931) and was converted into cDNA prior to sequencing using the SMARTer technology (SMARTer Pico PCR cDNA Synthesis Kit, Takara, cat. # 634928). Sequencing libraries were prepared using the Illumina TruSeq protocol and sequenced on an Illumina HiSeq using 150 bp paired end reads. Transcriptomes were assembled in Inchworm (Trinity v2.0.3) by using the default settings. Raw RNA-seq data are available at the NCBI Sequence Read Archive (accession numbers SRX8009001–SRX8009003). The Transcriptome Shotgun Assembly projects have been deposited at DDBJ/EMBL/GenBank under the accessions GILI00000000, GILJ00000000, GILK00000000. The versions described in this paper are the first versions, GILI01000000, GILJ01000000, GILK01000000.

Genomic DNA was extracted from freshly harvested cells using the Epicentre DNA extraction kit (Cat. No. MC85200). SSU rDNA genes of clones Colp-10, Colp-15, Colp-22, Colp-26, and Colp-tractor (GenBank accession numbers: MT254068, MT254081, MT252660, MT252669, MT250996) were amplified by PCR using general eukaryotes primers: 18SFU-FAD4 (Colp-10), PF1-FAD4 (Colp-15), EukA-EukB (Colp-22, Colp-26, Colp-tractor) (Medlin et al., 1988; Keeling, 2002; Tikhonenkov et al., 2016). PCR products were subsequently cloned and sequenced (Colp-15) or sequenced directly (Colp-10, Colp-22, Colp-26, Colp-tractor) by Sanger dideoxy sequencing. SSU rDNA genes of other colponemid strains were obtained as described previously (Janouškovec et al., 2013; Tikhonenkov et al., 2014).

3.1. SSU rDNA gene phylogeny

The SSU RNA gene phylogeny was based on an earlier dataset (Janouškovec et al., 2013). Sampling was modified to include the new colponemid sequences and a more representative selection of other eukaryotes and to exclude several longer-branching sequences. Sequences were aligned by the localpair algorithm in MAFFT v. 7.402 (Katoh and Standley, 2013) and trimmed by using the -h 0.4 and -g 0.35 settings in BMGE v. 1.12 (Criscuolo and Gribaldo, 2010). Maximum likelihood phylogeny was computed in IQ-TREE v. 1.6.9 (Nguyen et al., 2015) by using the best fit TN + F + R4 model (as determined by the in-built ModelFinder) and 300 non-parametric bootstraps. MrBayes v. 3.2.6 (Ronquist and Huelsenbeck, 2003) was run with default sampling scheme around the GTR + I + GAMMA4 model space until reaching chain convergence diagnostic (stopval) of 0.01, after which a majority consensus of 2576 trees in the posterior probability distribution was calculated (17,175,000 total generations; burnin 25%).

3.2. Phylogenomic dataset construction

The assembled sequences of *Colponema vietnamica* Colp-7a, and colponemid strains Colp-10 and Colp-15 were added to 320 publicly available protein-coding genes (Brown et al., 2018; Strassert et al., 2019) from a broad range of taxa as follows: 1) Protein sequences of the

new taxa were clustered with CD-HIT (Fu et al., 2012) using an identity threshold of 85%; 2) Candidates for homologous copies were retrieved by BLASTP searches (Altschul et al., 1990) using the 320 genes as queries (e-value: 1e-20; coverage cutoff: 0.5); 3) In three rounds, phylogenetic trees were constructed and carefully inspected in order to detect and remove paralogs and contaminants and select orthologous copies. For that, sequences were aligned using MAFFT v. 7.310 (Katoh and Standley, 2013) with the -auto flag employed (first round) and MAFFT L-INS-i with default settings (second and third round). Alignments were filtered with trimAL v. 1.4 (Capella-Gutiérrez et al., 2009) using a gap threshold of 0.8 (all three rounds) and single-gene Maximum Likelihood (ML) trees were inferred using FastTree v. 2.1.10 (Price et al., 2010) with the -lg -gamma setting (first round) and RAxML v. 8.2.10 (Stamatakis, 2014) with the PROTGAMMALGF model and 100 rapid bootstrap searches (second and third round).

Seven genes were removed from further analyses because no representative colponemid sequences could be retrieved. Moreover, to allow computationally demanding analyses in later steps, the number of taxa was reduced by excluding several fast-evolving taxa, taxa with more missing data, and taxa outside the 'TSAR' clade (Strassert et al., 2019; with exception of a few haptists that were retained as the outgroup). The resulting dataset comprised 313 genes and 71 taxa. Non-homologous characters were removed from individual unaligned sequences with PREQUAL v. 1.01 (Whelan et al., 2018) using a filter threshold of 0.95. Sequences were then aligned with MAFFT G-INS-i using the VSM option (–unalignlevel 0.6) and trimmed with BMGE v. 1.12 (Criscuolo and Gribaldo, 2010) employing -g 0.2, -b 5, -m, and BLOSUM75 parameters. Partial sequences belonging to the same taxon that did not show evidence for paralogy or contamination on the gene trees were merged. The genes (Supplemental Material) were concatenated to a supermatrix with SCaFos v. 1.25 (Roure et al., 2007) and an initial ML tree was calculated using IQ-TREE v. 1.6.9 (Nguyen et al., 2015) with the site-homogeneous model LG + G + F and ultrafast bootstrap approximation (UFBoot, Hoang et al., 2018; 1,000 replicates) employing the -bb and -bnni options. To reduce missing data, based on this tree, some monophyletic strains or species/genera complexes were combined to chimera that were used as operational taxonomic unit (OTU), and a new supermatrix (51 OTUs; 84,451 amino acid positions) was built as described above (see Supplemental Material for the corresponding files).

3.3. Phylogenomic analyses

The final supermatrix was used to infer a ML tree using IQ-TREE with the best-fitting site-heterogeneous model LG + C60 + G + F with the relatively fast PMSF approach (Wang et al., 2018) to obtain nonparametric bootstrap support (100 replicates). Copies of the obtained tree were manually edited to test alternative topologies using the approximately unbiased test (AU test; Shimodaira, 2002) in IQ-TREE: 1) colponemids sister to all other alveolates (original tree) *versus* colponemids sister to ciliates, 2) colponemids sister to all other alveolates *versus* colponemids sister to myzozoa. Also, the gene alignments that have been used for generating the final supermatrix were subjected to the Multi Species Coalescent (MSC) model in ASTRAL-III (Zhang et al., 2018) to incorporate gene tree uncertainties. ML single-gene trees (Supplemental Material) were inferred using IQ-TREE under the best-fitting substitution model, as determined by the BIC criterion. Branches with bootstrap support < 10 were collapsed in the input trees, as recommended by the developers. Quadripartition supports were calculated from quartet frequencies among the set of ML input trees (Sayyari and Mirarab, 2016). Bayesian analyses were conducted using PHYLO-BAYES-MPI v. 1.8 (Lartillot et al., 2013) with the CAT + GTR + Γ4 model employing the -dc flag to remove constant sites. Two independent Markov Chain Monte Carlo (MCMC) chains were run for 4,330 generations (all sampled). For each chain, the burnin period was estimated by monitoring evolution of the log-likelihood (Lnl) at each

sampled point, and 1000 generations were removed from both chains. A consensus tree of the two chains was built with the bpcomp command and global convergence between chains was assessed by the maxdiff statistics measuring the discrepancy in posterior probabilities (PP). As frequently recognized in PHYLOBAYES, global convergence was not received (for details, see Supplemental Material).

4. Results

4.1. Phylogenomic analyses

To determine the evolutionary position of colponemids, we generated transcriptomes of *Colponema vietnamica* strain Colp-7a and two newly-discovered relatives, and incorporated their sequences into a phylogenomic dataset (Materials and Methods). All genes were individually inspected for orthology by computing single-gene phylogenies and removing paralogs and contaminant sequences. The final phylogenomic matrix contained 51 OTUs and 313 protein-coding genes (84,451 amino acid positions). The colponemid data sets showed deep coverage of the analysed genes: 99% for the colponemid strain Colp-10, 97% for the colponemid strain Colp-15, and 80% for *C. vietnamica* (see Supplemental Material).

Maximum Likelihood analysis of this dataset recovered a monophyly of Colponemidia (although represented by three strains only), which was placed as a sister lineage to all other alveolates with maximal bootstrap support (Fig. 1), in agreement with Bayesian analysis (Figs. 1 and S1). The AU test supports this topology (p-AU = > 0.95) by rejecting the alternative hypothesis of a colponemids/ciliates sister relationship at the p-AU = < 0.05 significance level. The AU test also rejects the position of colponemids sister to Myzozoa (p-AU = 0; with p-AU = 1 for colponemids sister to all other alveolates). In agreement with previous studies and current systematics, the Maximum Likelihood analysis also recovered monophyletic alveolates, ciliates, myzozoans, core dinoflagellates, and apicomplexans, as well as the sister relationship between alveolates and stramenopiles, each fully supported (Strassert et al., 2019). Bayesian analysis showed the same results (Figs. 1 and S1): two Markov chains recovered monophyletic colponemids as a sister group to other alveolates with maximal posterior probability support, as well as monophyletic ciliates, myzozoans, alveolates, stramenopiles, rhizarians, and 'SAR'.

The result of the multi species coalescent analysis is in an overall agreement with the Maximum Likelihood topology (Fig. 1). It also indicates a sister relationship of colponemids to all other alveolates although only with modest local posterior probability support.

4.2. Diversity, distribution, and ecological role of Colponemidia

We previously isolated and described several strains of *C. edaphicum* and *C. vietnamica* from different freshwater and soil environments of Chukotka, Caucasus, and Vietnam (Mylnikov and Tikhonenkov, 2007; Janouškovec et al., 2013; Tikhonenkov et al., 2014). The two new colponemid strains isolated here from Morocco (Colp-10) and Vietnam (Colp-15) are considerably divergent from both species (Figs. 1 and 2a). Additionally, we have isolated three new colponemid strains (Colp-22, Colp-26, and Colp-tractor) from a river in Ukraine, lake in South Vietnam, and bog in British Columbia, Canada. All the three strains share the overall *C. vietnamica* morphotype (e.g., they have elongated-oval not flattened cells with comparatively short ventral groove and anterior flagellum as well as long posterior flagellum, which sometimes undulates in the longitudinal ventral groove), and a phylogenetic analysis of their SSU rRNA genes confirms that they belong to the *C. vietnamica/edaphicum* clade (Fig. 2a and 3).

We examined SSU rDNA sequences from colponemid-related lineages (CRLs) from environmental surveys retrieved from GenBank (Fig. 2a). Only twelve sequences have been retrieved, which confirmed colponemids are not common in natural surveys, but they are diverse
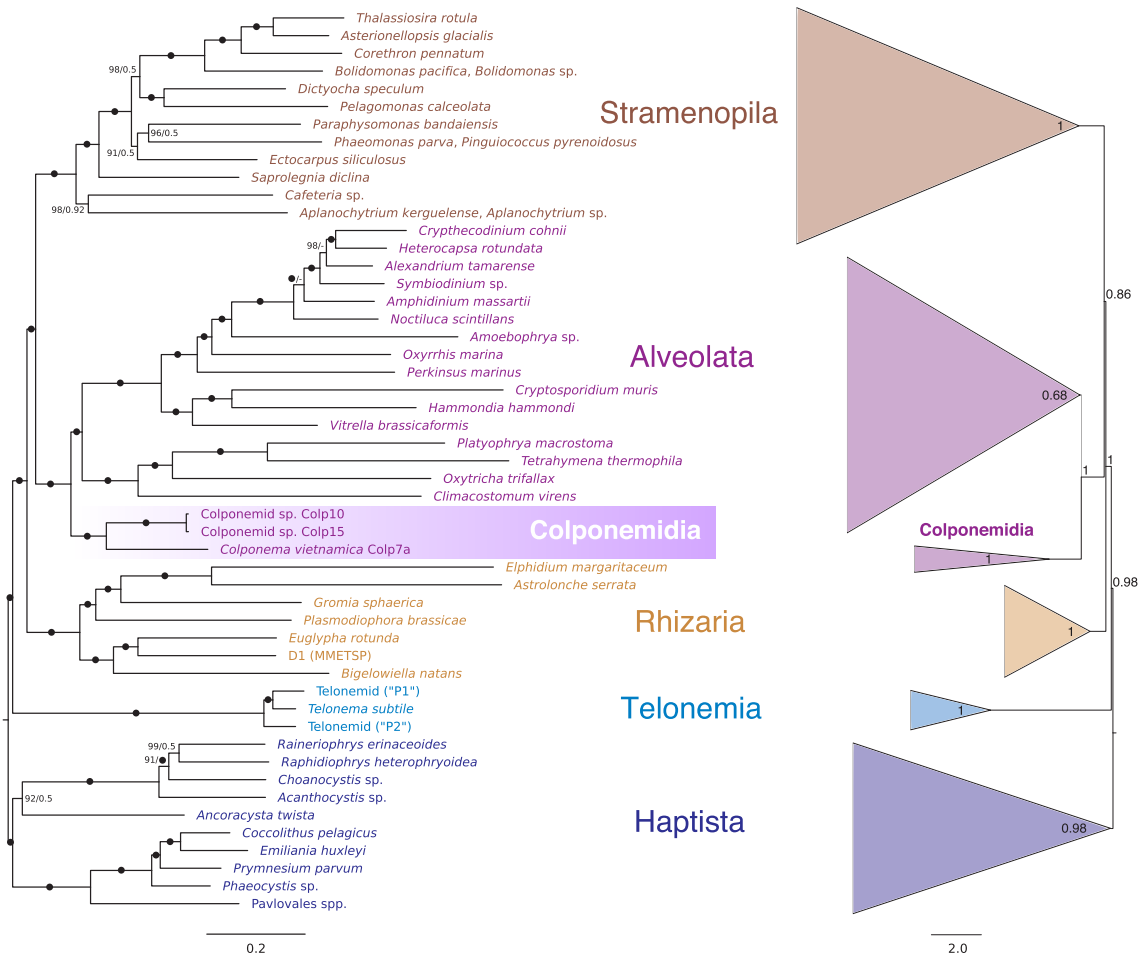
Fig. 1. Phylogenomic trees showing the sister group relationship between colponemids and all other alveolates. The tree on the left side represents the best Maximum Likelihood (ML) tree inferred from a concatenated alignment of 313 protein-coding genes using the LG + C60 + G + F-PMSF model. Numbers at nodes indicate bootstrap support (left) and Bayesian posterior probability (right). Full support in both analyses is shown by black circles. Maximum support in only one of the two analyses is indicated by a black circle appended to a number. Dashes at nodes indicate topological incongruence with the Phylobayes tree (see Fig. S1). The tree on the right is based on summary-coalescent analyses (Astral-III) of 313 single-gene ML trees (Supplemental Material) using the same taxa as shown on the left (collapsed to polygons). The numbers at nodes are local posterior probabilities.

and widespread, and also potentially form multiple subgroups basal to alveolates. Within the main colponemid clade, two environmental sequences are closely related to *C. edaphicum*, *C. vietnamica,* and three new strains (Colp-22, Colp-26, and Colp-Tractor). Five other environmental sequences previously referred to as CRL-II, III, IV (Janouškovec et al., 2013) form a weakly-supported subgroup with Colp-10 and Colp-15. Maximum Likelihood and Bayesian trees also show two groups of environmental sequences, CRL-V and CRL-I, to be independent of other colponemids, but again without statistical support (Fig. 2a).

Geographical distribution analysis has shown that colponemids are cosmopolitan at the species level (*C. vietnamica*) (Fig. 2b). They have been observed in all hemispheres, and in different geographic zones including the tropical, subtropical, temperate, and both Arctic and Antarctic polar zones. Interestingly, all molecularly identified colponemids and CRLs sequences have never been found in marine samples, and are only known to inhabit freshwater ecosystems and soil (i.e., thin freshwater films surrounding soil particles). They are present in different types of biotopes including plankton, bottom detritus, plant debris, peatlands, aquatic mosses, and wastewater from treatment facilities, but they are not known as an abundant component of microbial communities.

All newly studied colponemids are eukaryovores and bacteria alone are not sufficient for their survival in laboratory conditions, i.e., cultures die in the absence of eukaryotic prey. Colponemids fed on other

small protists in all our microscopical observations from natural samples and in laboratory cultures (e.g., on bodonids and chrysophytes), which is consistent with the presence of prey-immobilizing toxicysts (membrane surrounded capsules with long tubes inside) in microscopically investigated species *Colponema loxodes*, *C.* aff. *loxodes*, *C. edaphicum*, and *C. vietnamica* (Mignot and Brugerolle, 1975; Myl'nikova and Myl'nikov, 2010; Tikhonenkov et al., 2012, 2014). However, many dinoflagellates, colpodellids, and *Psammosa* possess trichocysts (spindle-shaped or rhomboid paracrystalline bodies with protein filaments inside, sometimes furnished with a specially constructed tip) but not toxicysts, while ciliates have both (Rosati and Modeo, 2003; Okamoto et al., 2012). Interestingly, the cells of colponemid strain Colp-10 have elongated skittle-shaped trichocysts (~0.9 × 0.15 μm) with an extended tip, roundish in a cross-section (Fig. 3). The presence of both types of extrusive organelles in Colponemidia is notable and potentially represents the state of early alveolate morphology.

5. Discussion

5.1. Phylogeny and diversity of Colponemidia

Previous phylogenetic analyses of *Colponema* based on rDNA and a set of six genes failed to conclusively resolve its evolutionary position. SSU and SSU/LSU rDNA phylogenies placed it as a sister to myzozoans,
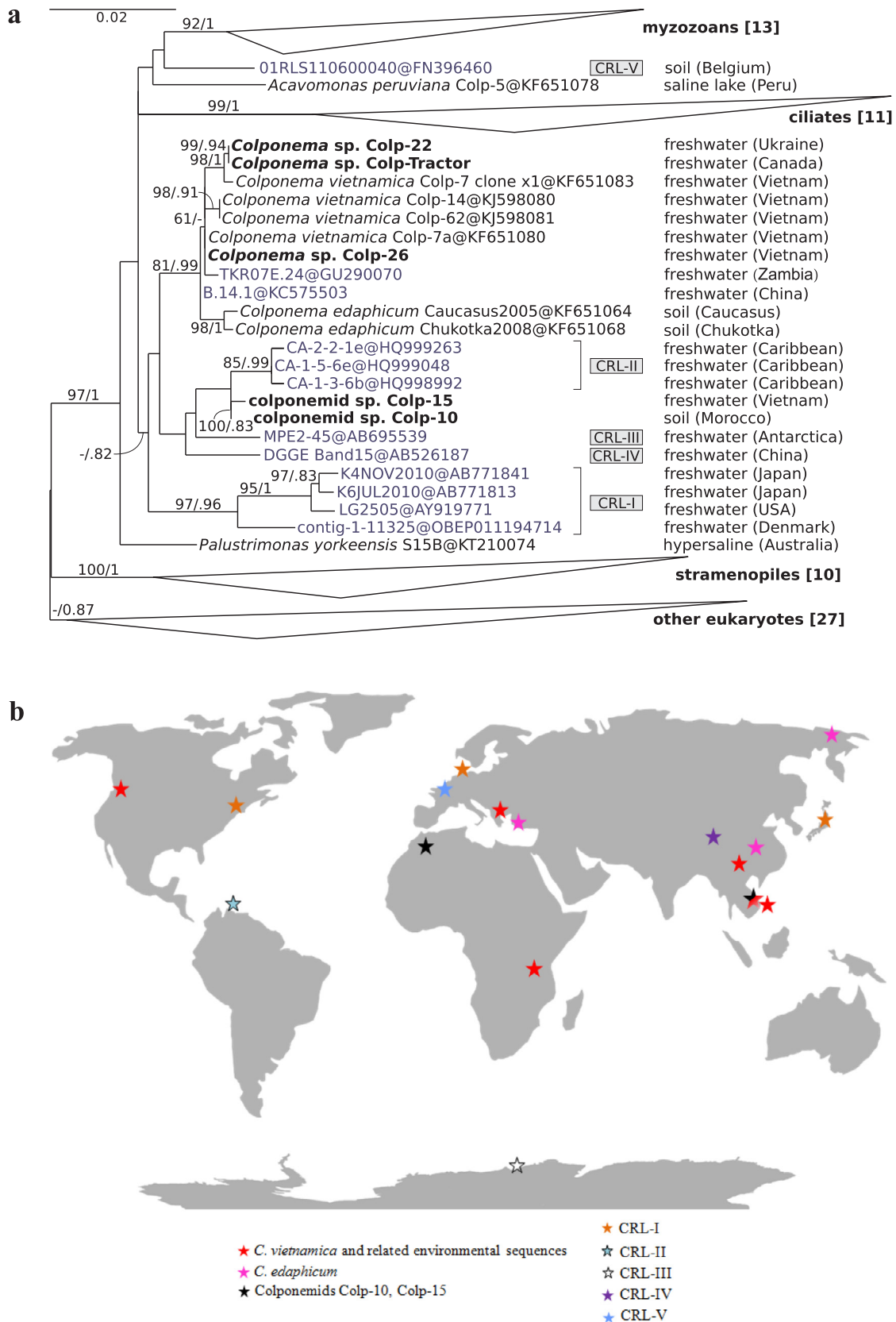
a

b

Fig. 2. The small subunit ribosomal RNA gene phylogeny (**a**) and geographic distribution (**b**) of colponemids and environmental colponemid-related lineages (CRLs). A maximum likelihood tree of 86 sequences (best fit TN + F + R4 model) is shown with non-parametric bootstrap supports and Bayesian posterior probabilities at branches (values > 50/ > 0.7 are shown). Environmental sequences are highlighted in dark blue and include five previously identified colponemid-related lineages (CRLs; grey boxes). Sequence accession numbers in GenBank are shown after the @ sign. Numbers of species in compacted clades are shown in square brackets (for interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).
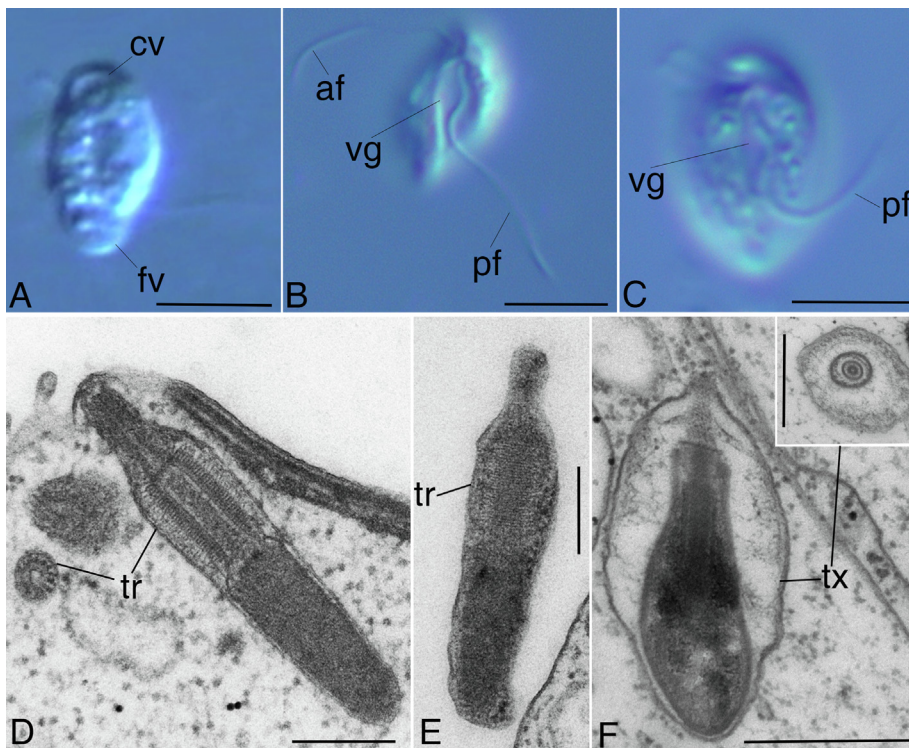
Fig. 3. External morphology and extrusive organelles of colponemids. **A–C** – general view of *C. vietnamica* (differential interference contrast) showing heterodynamic anterior (af) and posterior (pf) flagella, pronounced ventral groove (vg) with undulating posterior flagellum, apical contractile vacuole (cv), and posterior food vacuole (fv). **D, E** – Ultrathin sections (TEM) of skittle-shaped trichocysts (tr) with an extended tip, roundish in a cross-section of colponemid clone Colp-10. **F** – Longitudinal and cross (inset) sections (TEM) of bottle-shaped toxicyst (tx) enclosed inside a vesicle of *Colponema vietnamica*. Scale bars: A–C – 5 μm, D, E – 200 nm, F – 500 nm.

or myzozoans + *Acavomonas* (Janouškovec et al., 2013, 2017b; Mikhailov et al., 2014; Tikhonenkov et al., 2014; Cavalier-Smith, 2018), or sister to myzozoans + *Acavomonas* + *Palustrimonas* (Park and Simpson, 2015). Hsp90 and alpha tubulin phylogenies placed *Colponema* as a sister to ciliates, whereas concatenations of SSU and LSU with combinations of HSP90, actin, and alpha- and beta-tubulin placed it basal to other alveolates without statistical support (Janouškovec et al., 2013). Our phylogenomic analyses of 313 eukaryote protein-coding genes from several colponemids now resolves this position, placing them as a sister to all other studied alveolates with strong support in Maximum Likelihood site-heterogeneous mixture model and Phylobayes estimations.

Colponemids are widespread in nature and found in a broad range of habitats, including plankton, bottom sediments, boggy sites, and soil, and in extreme environments such as Arctic permafrost (*C. edaphicum*) and desert soil (colponemid Colp-10). Their widespread, or perhaps cosmopolitan nature is common in small heterotrophic flagellates in general (Lee and Patterson, 1998; Azovsky et al., 2016), which are also the primary food resources for colponemids. Exactly how they survive hostile conditions remains to be determined; only Colp-10 possesses cysts in its life cycle, and the survivability of *C. edaphicum* in 28–32 millennial permafrost (Shatilovich et al., 2009, 2010) remains a mystery. Despite this, however, all molecularly-identified colponemids to date originate from freshwaters and freshwater soil microhabitats — never from marine habitats.

5.2. Colponemids and the evolution of alveolates

Taking the structural features of colponemids into consideration together with their basal position in the tree has a major impact on how we reconstruct the common ancestor of alveolates and the major subgroups. In particular, the overall picture of the ancestral alveolate that emerges is somewhat different from the characteristics that typify the three major alveolate groups individually: it was a free-living predatory protist with submembrane alveolar vesicles, two heterodynamic flagella, tubular mitochondrial cristae, a posterior digestive vacuole, and extrusive organelles (trichocysts and/or toxicysts) for active hunting by

phagocytosis. This description is a close match with modern colponemids, and indeed they have been identified in the past as being potentially closely representative of early alveolate morphology (Leander and Keeling, 2003), even being described as "living fossils" (Cavalier-Smith, 2018). Interestingly, however, the morphostasis of colponemids may go even deeper, because they also resemble excavate protists (O'Kelly, 1993; Tikhonenkov et al., 2014), with ventral groove, tubular mitochondrial cristae, and vane on the posterior flagellum (although the vane is ventral in *Colponema*, ventral or both ventral and dorsal in malawimonads and plesiomorphic metamonads, and dorsal in jakobids). Other new lineages of alveolates will doubtless be characterized in the future (Fig. 2a), and it is even possible that some will be found to branch more deeply than Colponemidia. Fine structure of the colponemid cytoskeleton needs further investigation, as even here colponemids appear to retain ancient characteristics (Tikhonenkov et al., 2014). In the flagellar root system, the so-called R2 fibre splits into oR2 and iR2 that support the feeding apparatus, and this split likely represents an ancient characteristic of the last common ancestor of excavates, stramenopiles, apusozoans, amoebozoans, collodictyonids, haptophytes, cryptophytes, and probably alveolates (Yubuki and Leander, 2013). The cytoskeleton configuration in *Colponema* would therefore be ancestral not just to alveolates but to all eukaryotes (Tikhonenkov et al., 2014).

Cytoskeletal elements for feeding and locomotion are connected functionally and historically. In bacteriotrophs (most heterotrophic protists), flagellar beating creates a current directed to the cytostome by the feeding groove, moving suspended bacteria to the 'mouth'. Microtubular bands originating at the basal body provide mechanical rigidity to the groove margins. Colponemids retain this ancestral organization, although they are specialized to phagocytize eukaryotes and not bacteria (Myl'nikova and Myl'nikov, 2010), and use active hunting of mobile prey rather than passive suspension feeding. Subsequently, the ancestor of myzozoans and *Acavomonas* probably lost the link between flagellar apparatus and feeding groove as it transitioned from bacteriotrophy and phagotrophy to suction feeding on larger cells. For example, *Acavomonas* has no pronounced ventral groove (Tikhonenkov et al., 2014). But the newly-developed feeding system based on the

apical complex retained its link to the basal bodies because the microtubular band formerly supporting the feeding groove was most likely converted into the conoid that strengthens the rostrum and the edge of the cytostome in extant myzozoans. This link is obscured by the complex life cycles of most myzozoans, but can be clearly observed in the direct link between basal bodies and conoid in *Psammosa* (Okamoto and Keeling, 2014).

The origin of myzocytosis and radiation of myzocytosis-feeding taxa may have been linked to the increase of eukaryote (prey) cells sizes. Macrophagous predatory protists search and select prey by size that is suitable for feeding, at that larger cells of potential prey is more difficult to swallow (Bengtson, 2002). Unicellular eukaryotes reached their maximum size before the Cambrian explosion, although an increase in the average cell size through the Proterozoic was not a continuous trend (Huntley et al., 2006). The observed size ratios between modern predatory protists (dinoflagellates, chrysophytes, bicosoecids, pedinellids, ciliates) and their prey range from 0.4:1 to 30:1 because of different mechanisms of prey capture (Hansen et al., 1994). Myzocytosis allows a predator to consume biomass accumulated in prey of any size, for example, *Colpodella gonderi* attacks large ciliates (Olmo et al., 2011). The ancestors of dinoflagellates and apicomplexans were likely capable of myzocytosis and possessed an apical complex-like structure including an open conoid, rhoptries, and micronemes. These characteristics are mostly lost in autotrophic dinoflagellates and their relatives, but parasitic apicomplexans and dinoflagellates have adapted their myzocytotic feeding to multicellular organisms. Although parasitic dinoflagellates and perkinseids possess flagellar apparatus to search their multicellular or unicellular prey, apicomplexan parasites have lost flagellated zoospores. This loss could be associated with the events in the Middle Cambrian, known as the "Agronomical revolution", or the "Cambrian substrate revolution" (Seilacher and Pflüger, 1994). At this time, organic material began to accumulate on the seabed and burrowing organisms increased in number and diversity. This would in turn increase the likelihood of immobile resting oocysts of apicomplexans being acquired by feeding of burrowing host. Further evolution without mobile infectious stages has led to complex apicomplexan life cycles exploiting the behavior of animal vectors.

Another well-studied character in alveolate evolution is the plastid organelle. Photosynthetic and non-photosynthetic plastids derived from a secondary enslavement of a red alga are found throughout the myzozoans (Janoušovec et al., 2015), but no evidence for plastids or plastid genes has been found in ciliates or colponemids to date (Janoušovec et al., 2013). The origin of this plastid, has been widely debated: on one side it is argued to be the result of a single ancient plastid origin predating alveolates, stramenopiles and perhaps several other lineages (Cavalier-Smith, 1999, 2013), while on the other side it has been argued to have originated more recently within alveolates by more hierarchical endosymbioses (Petersen et al., 2014; Stiller et al., 2014). If the former is true, then the ancestor of alveolates would have been mixotrophic, and both colponemids and ciliates would have lost plastids and plastid genes secondarily (as have parasitic myzozoans *Cryptosporidium* and *Hematodinium*: Gornik et al., 2015; Xu et al., 2004; Zhu et al., 2000). In contrast, the latter possibility suggests colponemids never had plastids and are more similar to the ancestral host of myzozoan plastids.

No plastid genes have been identified in our transcriptomic surveys of colponemids, and while complete genome sequencing of colponemids may reveal their presence, it seems very unlikely given existing data that they retain a plastid or any substantial number of relict plastid genes. Complete loss of the plastid and plastid genes is rare and difficult to prove. The only uncontested cases are observed in parasitic myzozoans with completely sequenced genomes (Abrahamsen et al., 2004; Gornik et al., 2015; Zhu et al., 2000) and a robust phylogenetic position within plastid-bearing lineages. And, even though plastids in myzozoan parasites were probably lost at least four times (Janoušovec et al., 2019), metabolic dependence on them is apparently more difficult to overcome in metabolically active free-living protists where no definite cases of plastid loss exist (Gawryluk et al., 2019; Janoušovec et al., 2015; Janoušovec et al., 2017a). Distinguishing whether colponemids are ancestrally or secondarily free of plastids is therefore difficult. At that, a detailed reconstruction of the metabolic pathways in colponemids and analysis of whether they are decoupled from known plastid functions to show functionally how they are independent of plastid pathways would be essential. Acquiring data on the biological nature of colponemid-related lineages could be similarly informative. Overall, however, the morphostasis in colponemids strengthens the assumption that they lacked the plastid ancestrally, and their predatory, eukaryovorous nature and a large feeding groove provide a good model for the pre-symbiotic life style that would have enabled plastid acquisition via an engulfment of a eukaryotic alga.

Distinguishing between the heterotrophic *versus* mixotrophic nature of the ancestral alveolate also has direct ramifications for the origin of plastids in the photosynthetic stramenopiles, haptophytes and cryptophytes, and plastid presence or loss in the related heterotrophs. As such, the origin of alveolate plastids could resolve a major gap in our understanding of the evolution of photosynthesis in a major part of eukaryotic diversity and uncover the general principles governing plastid acquisition and loss. We speculate that a breakthrough in this debate will come from the isolation and characterization of new taxa with combinations of traits we have yet to observe.

CRediT roles

Denis V. Tikhonenkov: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Jürgen F.H. Strassert:** Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Jan Janoušovec:** Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Alexander P. Mylnikov:** Conceptualization, Funding acquisition, Investigation. **Vladimir V. Aleoshin:** Conceptualization, Writing - original draft, Writing - review & editing. **Fabien Burki:** Conceptualization, Data curation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Patrick J. Keeling:** Conceptualization, Funding acquisition, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2020.106839.

References

Abrahamsen, M.S., Templeton, T.J., Enomoto, S., Abrahante, J.E., Zhu, G., Lancto, C.A., Deng, M., Liu, C., Widmer, G., Tzipori, S., Buck, G.A., Xu, P., Bankier, A.T., Dear, P.H., Konfortov, B.A., Spriggs, H.F., Iyer, L., Anantharaman, V., Aravind, L., Kapur, V., 2004. The intestinal parasite *Cryptosporidium parvum* lacks several common biochemical pathways and intracellular organelles and thus must rely on its host for nutrients. Science 304, 441–445. https://doi.org/10.1126/science.1094786.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Azovsky, A.I., Tikhonenkov, D.V., Mazei, Yu.A., 2016. An estimation of the global diversity and distribution of the smallest eukaryotes: biogeography of marine benthic heterotrophic flagellates. Protist 167, 411–424. https://doi.org/10.1016/j.protis.2016.07.001.

Bass, D., Tikhonenkov, D.V., Foster, R., Dyal, P., Janouskovec, J., Keeling, P.J., Gardner, M., Neuhauser, S., Hartikainen, H., Mylnikov, A.P., Berney, C., 2018. Rhizarian 'Novel Clade 10′ revealed as abundant and diverse planktonic and terrestrial flagellates, including Aquavolon n. gen. J. Eukaryot. Microbiol. 65, 828–842. https://doi.org/10.1111/jeu.12524.

Bengtson, S., 2002. Origins and early evolution of predation. In: Kowalewski, M., Kelley, P.H. (Eds.), The fossil record of predation. The Paleontological Society Papers 8. The Paleontological Society, New York, pp. 289–317.

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., Roger, A.J., 2018. Phylogenomics places orphan protistan lineages in a novel eukaryotic super-group. Genome Biol. Evol. 10, 427–433. https://doi.org/10.1093/gbe/evy014.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

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. https://doi.org/10.1111/j.1550-7408.1999.tb04614.x.

Cavalier-Smith, T., 2013. Early evolution of eukaryote feeding modes, cell structural diversity, and classification of the protozoan phyla Loukozoa, Sulcozoa, and Choanozoa. Eur. J. Protistol. 49, 115–178. https://doi.org/10.1016/j.ejop.2012.06.001.

Cavalier-Smith, T., 2014. Gregarine site–heterogeneous 18S rDNA trees, revision of gregarine higher classification, and the evolutionary diversification of Sporozoa. Eur. J. Protistol. 50, 472–495. https://doi.org/10.1016/j.ejop.2014.07.002.

Cavalier-Smith, T., 2018. Kingdom Chromista and its eight phyla: a new synthesis emphasizing periplastid protein targeting, cytoskeletal and periplastid evolution, and ancient divergences. Protoplasma 255, 297–357. https://doi.org/10.1007/s00709-017-1147-3.

Chambouvet, A., Morin, P., Marie, D., Guillou, L., 2008. Control of toxic marine dinoflagellate blooms by serial parasitic killers. Science 322, 1254–1257. https://doi.org/10.1126/science.1164387.

Criscuolo, A., Gribaldo, S., 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol. Biol. 10, 210. https://doi.org/10.1186/1471-2148-10-210.

Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565.

Gawryluk, R.M.R., Tikhonenkov, D.V., Hehenberger, E., Husnik, F., Mylnikov, A.P., Keeling, P.J., 2019. Non-photosynthetic predators are sister to red algae. Nature 572, 240–243. https://doi.org/10.1038/s41586-019-1398-6.

Goggin, C.L., Barker, S.C., 1993. Phylogenetic position of the genus *Perkinsus* (Protista, Apicomplexa) based on small subunit ribosomal RNA. Mol. Biochem. Parasitol. 60, 65–70. https://doi.org/10.1016/0166-6851(93)90029-W.

Gornik, S.G., Febri Marsa Cassin, A.M., MacRae, J.I., Ramaprasad, A., Rchiad, Z., McConville, M.J., Bacic, A., McFadden, G.I., Pain, A., Waller, R.F., 2015. Endosymbiosis undone by stepwise elimination of the plastid in a parasitic dinoflagellate. Proc. Natl. Acad. Sci. USA 112, 5767–5772. https://doi.org/10.1073/pnas.1423400112.

Gould, S.B., Tham, W.-H., Cowman, A.F., McFadden, G.I., Waller, R.F., 2008. Alveolins, a new family of cortical proteins that define the protist infrakingdom Alveolata. Mol. Biol. Evol. 25, 1219–1230. https://doi.org/10.1093/molbev/msn070.

Hansen, B., Bjørnsen, P.K., Hansen, H.J., 1994. The size ratio between planktonic predators and their prey. Limnol. Oceanogr. 39, 395–403. https://doi.org/10.4319/lo.1994.39.2.0395.

Huntley, J., Xiao, S., Kowalewski, M., 2006. 1.3 Billion years of acritarch history: An empirical morphospace approach. Precambrian Res. 144, 52–68. https://doi.org/10.1016/j.precamres.2005.11.003.

Hoang, D.T., Chernomor, O., Haeseler, von, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281.

Janouškovec, J., Tikhonenkov, D.V., Mikhailov, K.V., Simdyanov, T.G., Aleoshin, V.V., Mylnikov, A.P., Keeling, P.J., 2013. Colponemids represent multiple ancient alveolate lineages. Curr. Biol. 23, 2546–2552. https://doi.org/10.1016/j.cub.2013.10.062.

Janouškovec, J., Tikhonenkov, D.V., Burki, F., Howe, A.T., Kolísko, M., Mylnikov, A.P., 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. https://doi.org/10.1073/pnas.1423790112.

Janouškovec, J., Gavelis, G.S., Burki, F., Dinh, D., Bachvaroff, T.R., Gornik, S.G., Bright, K.J., Imanian, B., Strom, S.L., Delwiche, C.F., Waller, R.F., Fensome, R.A., Leander, B.S., Rohwer, F.L., Saldarriaga, J.F., 2017a. Major transitions in dinoflagellate evolution unveiled by phylotranscriptomics. Proc. Natl. Acad. Sci. USA 114, E171–E180. https://doi.org/10.1073/pnas.1614842114.

Janouškovec, J., Tikhonenkov, D.V., Burki, F., Howe, A.T., Rohwer, F.L., Mylnikov, A.P., Keeling, P.J., 2017b. A new lineage of eukaryotes illuminates early mitochondrial genome reduction. Curr. Biol. 27, 3717–3724. https://doi.org/10.1016/j.cub.2017.10.051.

Janouškovec, J., Paskerova, G.G., Miroliubova, T.S., Mikhailov, K.V., Birley, T., Aleoshin, V.V., Simdyanov, T.G., 2019. Apicomplexan-like parasites are polyphyletic and widely but selectively dependent on cryptic plastid organelles. eLife 8, e49662. https://doi.org/10.7554/eLife.49662.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.

Katris, N.J., van Dooren, G.G., McMillan, P.J., Hanssen, E., Tilley, L., Waller, R.F., 2014. The Apical Complex Provides a Regulated Gateway for Secretion of Invasion Factors in *Toxoplasma*. PLoS Pathog. 10, e1004074. https://doi.org/10.1371/journal.ppat.1004074.

Keeling, P.J., 2002. Molecular phylogenetic position of *Trichomitopsis termopsidis* (Parabasalia) and evidence for the Trichomitopsiinae. Eur. J. Protistol. 38, 279–286. https://doi.org/10.1078/0932-4739-00874.

Lartillot, N., Rodrigue, N., Stubbs, D., Richer, J., 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst. Biol. 62, 611–615. https://doi.org/10.1093/sysbio/syt022.

Leander, B.S., Keeling, P.J., 2003. Morphostasis in alveolate evolution. Trends Ecol. Evol. 18, 395–402. https://doi.org/10.1016/S0169-5347(03)00152-6.

Lee, W.J., Patterson, D.J., 1998. Diversity and geographic distribution of free-living heterotrophic flagellates – analysis by PRIMER. Protist 149, 229–243. https://doi.org/10.1016/S1434-4610(98)70031-8.

López-García, P., Rodriguez-Valera, F., Pedrós-Alió, C., Moreira, D., 2001. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature 409, 603–607. https://doi.org/10.1038/35054537.

Mathur, V., Kolisko, M., Hehenberger, E., Irwin, N.A.T., Leander, B.S., Kristmundsson, Á., Freeman, M.A., Keeling, P.J., 2019. Multiple independent origins of apicomplexan-like parasites. Curr. Biol. 29, 2936–2941.e5. https://doi.org/10.1016/j.cub.2019.07.019.

Medlin, L., Elwood, H.J., Stickel, S., Sogin, M.L., 1988. The characterization of enzymatically amplified eukaryotes 16S-like ribosomal RNA coding regions. Gene 71, 491–500. https://doi.org/10.1016/0378-1119(88)90066-2.

Mignot, J.P., Brugerolle, G., 1975. Étude ultrastructurale du flagelle phagotrophe *Colponema loxodes* Stein. Protistologica 11, 429–444.

Mikhailov, K.V., Janouškovec, J., Tikhonenkov, D.V., Mirzaeva, G.S., Diakin, A.Y., Simdyanov, T.G., Mylnikov, A.P., Keeling, P.J., Aleoshin, V.V., 2014. A complex distribution of elongation family GTPases EF1A and EFL in basal alveolate lineages. Genome Biol. Evol. 6, 2361–2367. https://doi.org/10.1093/gbe/evu186.

Moore, R.B., Oborník, M., Janouškovec, J., Chrudimský, T., Vancová, M., Green, D.H., Wright, S.W., Davies, N.W., Bolch, C.J.S., Heimann, K., Šlapeta, J., Hoegh-Guldberg, O., Logsdon, J.M., Carter, D.A., 2008. A photosynthetic alveolate closely related to apicomplexan parasites. Nature 451, 959–963. https://doi.org/10.1038/nature06635.

Mylnikov, A.P., Tikhonenkov, D.V., 2007. A new species of soil predatory flagellate *Colponema edaphicum* sp. n. from Vorontsovskaya Cave, North Caucasus (Protista, Alveolata: Colponemidae). Zoosyst. Rossica 16, 1–4.

Myl'nikova, Z.M., Myl'nikov, A.P., 2010. Biology and morphology of freshwater rapacious flagellate Colponema aff. loxodes Stein (Colponema, Alveolata). Inland Water Biol., 3, 21–26. https://doi.org/10.1134/S1995082910010037.

Nash, E.A., Nisbet, R.E.R., Barbrook, A.C., Howe, C.J., 2008. Dinoflagellates: a mitochondrial genome all at sea. Trends Genet. 24, 328–335. https://doi.org/10.1016/j.tig.2008.04.001.

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300.

Okamoto, N., Horák, A., Keeling, P.J., 2012. Description of two species of early branching dinoflagellates, *Psammosa pacifica* n. g., n. sp. and *P. atlantica* n. sp. PLoS ONE 7, e34900. https://doi.org/10.1371/journal.pone.0034900.

Okamoto, N., Keeling, P.J., 2014. The 3D structure of the apical complex and association with the flagellar apparatus revealed by serial TEM tomography in *Psammosa pacifica*, a distant relative of the apicomplexa. PLoS ONE 9, e84653. https://doi.org/10.1371/journal.pone.0084653.

O'Kelly, C.J., 1993. The jakobid flagellates: structural features of *Jakoba*, *Reclinomonas* and *Histiona* and implications for the early diversification of eukaryotes. J. Eukaryot. Microbiol. 40, 627–636. https://doi.org/10.1111/j.1550-7408.1993.tb06120.x.

Olmo, J.L., Esteban, G.B., Finlay, B.J., 2011. New records of the ectoparasitic flagellate *Colpodella gonderi* on non-*Colpoda* ciliates. Int. Microbiol. 14, 207–211. https://doi.org/10.2436/20.1501.01.150.

Park, J.S., Simpson, A.G.B., 2015. Diversity of Heterotrophic Protists from Extremely Hypersaline Habitats. Protist 166, 422–437. https://doi.org/10.1016/j.protis.2015.06.001.

Petersen, J., Ludewig, A.K., Michael, V., Bunk, B., Jarek, M., Baurain, D., Brinkmann, H., 2014. *Chromera velia*, endosymbioses and the rhodoplex hypothesis plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages). Genome Biol. Evol. 6, 666–684. https://doi.org/10.1093/gbe/evu043.

Price, M.N.M., Dehal, P.S.P., Arkin, A.P.A., 2010. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490. https://doi.org/10.1371/journal.pone.0009490.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574. https://doi.org/10.1093/bioinformatics/btg180.

Rosati, G., Modeo, L., 2003. Extrusomes in ciliates: diversification, distribution, and phylogenetic implications. J. Eukaryot. Microbiol. 50, 383–402. https://doi.org/10.1111/j.1550-7408.2003.tb00260.x.

Roure, B., Rodriguez-Ezpeleta, N., Philippe, H., 2007. SCaFoS: a tool for selection, concatenation and fusion of sequences for phylogenomics. BMC Evol. Biol. 7, S2. https://doi.org/10.1186/1471-2148-7-S1-S2.

Sayyari, S., Mirarab, S., 2016. Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. 33, 1654–1668. https://doi.org/10.1093/molbev/msw079.

Seilacher, A., Pflüger, F., 1994. In: Biostabilization of Sediments. Bibliotheks und Informationssystem der Carl von Ossietzky Universität Odenburg, pp. 97–105.

Shatilovich, A.V., Shmakova, L.A., Mylnikov, A.P., Gilichinsky, D.A., 2009. Ancient Protozoa isolated from Permafrost. In: Margesin, R. (Ed.), Permafrost Soils, Soil Biology 16. Springer-Verlag, Berlin Heidelberg, pp. 97–115.

Shatilovich, A.V., Mylnikov, A.P., Stoopin, D.V., 2010. The fauna and morphology of heterotrophic flagellates and heliozoans from late Pleistocene fossil rodent burrows (Kolyma Lowland). Zool. Zh. 89, 387–397.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508. https://doi.org/10.1080/10635150290069913.

Smith, D.G.S., Gawryluk, R.M.R., Spencer, D.F., Pearlman, R.E., Siu, K.W.M., Gray, M.W., 2007. Exploring the mitochondrial proteome of the ciliate protozoon *Tetrahymena thermophila*: direct analysis by tandem mass spectrometry. J. Mol. Biol. 374, 837–863. https://doi.org/10.1016/j.jmb.2007.09.051.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

Stiller, J.W., Schreiber, J., Yue, J., Guo, H., Ding, Q., Huang, J., 2014. The evolution of photosynthesis in chromist algae through serial endosymbiosis. Nat. Commun. 5, 5764. https://doi.org/10.1038/ncomms6764.

Strassert, J.F.H., Tikhonenkov, D.V., Pombert, J.F., Kolisko, M., Tai, V., Mylnikov, A.P., Keeling, P.J., 2016. *Moramonas marocensis* gen. nov., sp. nov.: a jakobid flagellate isolated from desert soil with a bacteria-like, but bloated mitochondrial genome. Open Biol. 6, 150239. https://doi.org/10.1098/rsob.150239.

Strassert, J.F.H., Jamy, M., Mylnikov, A.P., Tikhonenkov, D.V., Burki, F., 2019. New phylogenomic analysis of the enigmatic phylum Telonemia further resolves the eukaryote tree of life. Mol. Biol. Evol. 36, 757–765. https://doi.org/10.1093/molbev/msz012.

Tikhonenkov, D., Mylnikov, A., Gong, Y., Feng, W., Mazei, Yu., 2012. Heterotrophic flagellates from freshwater and soil habitats in subtropical China (Wuhan area Hubei province). Acta Protozool. 1, 65–79. https://doi.org/10.4467/16890027AP.12.006.0389.

Tikhonenkov, D.V., Janouškovec, J., Mylnikov, A.P., Mikhailov, K.V., Simdyanov, T.G., Aleoshin, V.V., 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. https://doi.org/10.1371/journal.pone.0095467.

Tikhonenkov, D.V., Janouškovec, J., Keeling, P.J., Mylnikov, A.P., 2016. The morphology, ultrastructure and SSU rRNA gene sequence of a new freshwater flagellate, *Neobodo borokensis* n. sp. (Kinetoplastea, Excavata). J. Eukaryot. Microbiol. 63, 220–232. https://doi.org/10.1111/jeu.12271.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Coordinators, Tara Oceans, Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P., Karsenti, E., 2015. Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. Science 348, 1261605. https://doi.org/10.1126/science.1261605.

Wang, H.C., Minh, B.Q., Susko, E., Roger, A.J., 2018. Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. Syst. Biol. 67, 216–235. https://doi.org/10.1093/sysbio/syx068.

Whelan, S., Irisarri, I., Burki, F., 2018. PREQUAL: detecting non-homologous characters in sets of unaligned homologous sequences. Bioinformatics 496, 311–312. https://doi.org/10.1093/bioinformatics/bty448.

Xu, P., Widmer, G., Wang, Y., Ozaki, L.S., Alves, J.M., Serrano, M.G., Puiu, D., Manque, P., Akiyoshi, D., Mackey, A.J., Pearson, W.R., Dear, P.H., Bankier, A.T., Peterson, D.L., Abrahamsen, M.S., Kapur, V., Tzipori, S., Buck, G.A., 2004. The genome of *Cryptosporidium hominis*. Nature 431, 1107–1112. https://doi.org/10.1038/nature02977.

Yubuki, N., Leander, B.S., 2013. Evolution of microtubule organizing centers across the tree of eukaryotes. Plant J. 75, 230–244. https://doi.org/10.1111/tpj.12145.

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinform. 19, 153. https://doi.org/10.1186/s12859-018-2129-y.

Zhu, G., Marchewka, M.J., Keithly, J.S., 2000. *Cryptosporidium parvum* appears to lack a plastid genome. Microbiology 146, 315–321. https://doi.org/10.1099/00221287-146-2-315.