The transcriptome of *Paraphelidium tribonemae* illuminates the ancestry of Fungi and Opisthosporidia

Guifré Torruella[1], Xavier Grau-Bové[2], David Moreira[1], Sergey A. Karpov[1,3], John A. Burns[4], Arnau Sebé-Pedrós[5], Eckhard Völcker[6], Purificación López-García[1,7,*]

[1] Unité d'Ecologie, Systématique et Evolution, CNRS, Université Paris-Sud, Université Paris-Saclay, AgroParisTech, Orsay, France

[2] Institut de Biologia Evolutiva, Universitat Pompeu Fabra-CSIC, Barcelona, Catalonia

[3] Zoological Institute, Russian Academy of Sciences and St. Petersburg State University, St. Petersburg, Russian Federation

[4] Sackler Institute for Comparative Genomics and Division of Invertebrate Zoology, American Museum of Natural History, New York, USA

[5] Weizmann Institute of Science, Rehovot, Israel

[6] Penard Laboratory, Cape Town 7708, South Africa

[7] Lead contact

*Correspondence: puri.lopez@u-psud.fr, guifftc@gmail.com

SUMMARY

Aphelids constitute a group of diverse, yet poorly known, parasites of algae [1, 2]. Their life cycle and morphology resemble those of zoosporic fungi (chytrids) and rozellids (Cryptomycota/Rozellosporidia), another specious group of parasites of fungi and oomycetes [3, 4]. Unlike fungi, which are osmotrophs, aphelids and rozellids are phagotrophs, feeding on the host's cytoplasm. Combined RNA polymerase and rRNA gene trees [5] suggested that aphelids and rozellids relate to Microsporidia, extremely reduced parasites with remnant mitochondria [6]. Accordingly, aphelids, rozellids and Microsporidia were proposed to form a monophyletic clade sister to Fungi, called Opisthosporidia [1]. Microsporidia would have subsequently lost the ancestral opisthosporidian phagotrophy. However, the limited phylogenetic signal of those genes combined with microsporidian fast-evolving sequences have resulted in incongruent tree topologies, showing either rozellids [5, 7] or aphelids [8] as the earliest-branching lineages of Opisthosporidia, and challenging their monophyly. We have generated the first transcriptome data for one aphelid species, *Paraphelidium tribonemae* [2]. Multi-gene phylogenomic analyses clearly confirm the monophyly of Opisthosporidia, placing aphelids as the earliest-branching opisthosporidian lineage. This is consistent with the rich proteome inferred for *P. tribonemae*, which includes cellulases likely involved in algal cell-wall penetration, enzymes involved in chitin biosynthesis and several metabolic pathways that were lost in the comparatively reduced *Rozella allomycis* genome [9]. Contrary to recent claims suggesting a parasitic root for Fungi, our results suggest that Fungi and Opisthosporidia evolved from a free-living phagotrophic ancestor that became osmotrophic at the fungal root and evolved towards phagotrophic parasitism in the opisthosporidian line.

Keywords: aphelid; opisthokont; protist evolution; fungi; cell wall; chitin; phagotrophy; osmotrophy; parasitism; phylogenomics

RESULTS AND DISCUSSION

Aphelids are the earliest-branching lineage of the monophyletic Opisthosporidia clade

To improve the phylogenetic signal for an accurate placement of aphelids in the Opisthokont branch containing fungi, nucleariids, rozellids and Microsporidia (usually referred to as Holomycota), we generated transcriptome data for the aphelid species *Paraphelidium tribonemae* [2]. Because *P. tribonemae* has a complex life cycle (Figure 1, inlet), we maximized transcript recovery by constructing two cDNA libraries corresponding to young and old enrichment cultures of its host, the yellow-green alga *Tribonema gayanum*, infected with the aphelid. Accordingly, transcripts corresponding to zoospores, infective cysts, trophonts, plasmodia, sporangia and sporocysts were represented [2]. After paired-end Illumina HiSeq2500 sequencing, we assembled a metatranscriptome of 68,130 contigs corresponding to the aphelid, its host and bacterial contaminants. After a process of supervised cleaning (see Supplementary Methods) we obtained an initial set of 10,669 annotated peptides that were considered free of contaminants. The predicted proteome (*Paraphelidium tribonemae* version 1; Table S1) was 98% complete according to BUSCO [10]. We found no stop codons within transcripts. Therefore, in contrast to *Amoeboaphelidium protococcarum*, for which the TAG and TAA stop triplets appear to encode glutamine [5], *P. tribonemae* has a canonical genetic code.

We incorporated *P. tribonemae* orthologs to a dataset of 93 Single-Copy Protein Domains previously used for phylogenomic analysis of basal-branching opisthokonts [11], updating the dataset with more microsporidian species including the early-branching *Mitosporidium daphniae* [12] and the metchnikovellid *Amphiamblys* sp. [8]. The concatenated matrix contained 22,207 conserved amino acid sites for 48 species including opisthokonts and close outgroups (see Supplementary Methods for details). Bayesian inference (BI) trees reconstructed under the CAT-Poisson mixture model [13] and Maximum Likelihood (ML) trees reconstructed under C60, the closest mixture model to CAT [14], yielded the same topology (Figure 1 and S1). We recovered the monophyly of all opisthokont lineages previously reported [11, 15]. In addition, we retrieved the monophyly of Opisthosporidia, as a sister group to Fungi in both, BI and ML analyses. However, while Bayesian posterior probability provided full statistical support for the monophyly of Opisthosporidia, the support dropped in ML trees to 79% (ultra-fast bootstrapping [16]) and 80% (non-parametric bootstrap under the Posterior Mean Site Frequency model [17]). Within Opisthosporidia, the aphelid was always the earliest-branching lineage, *Rozella* being the sister group to Microsporidia with full statistical support. The moderate ML bootstrap support for the opisthosporidian monophyly appears to reflect the attraction of *Paraphelidium* towards the fungal lineage Blastocladiomycota, as can be noticed in the SplitsTree that depicts the bootstrap support for all bipartitions (Figure S1D). Globally, our results confirm the

monophyly of Opisthosporidia and show the basal position of Aphelida with respect to Rozellosporidia and Microsporidia, in agreement with morphological and life cycle observations [1].

Enzymes Involved in Cell-Wall Synthesis and Degradation

Despite secondary losses in some fungi, the presence of chitin in cell walls was for long considered a fungal specific trait. However, chitin is also present in the cell wall of many other protists across the eukaryotic tree, implying that the machinery for chitin synthesis/remodeling originated prior to the radiation of Fungi and other eukaryotic lineages [18]. Microsporidia and *Rozella* are also able to synthesize chitin [9, 19] but, unlike fungi, which possess chitin cell walls during the vegetative stage, they produce chitin only in cysts [20]. Staining with fluorescently-labeled Wheat Germ Agglutinin (WGA) showed that *Paraphelidium* also possesses chitin in the wall of infecting cysts, but not in zoospores or trophonts (Figure 2A-B). In agreement with this observation, we identified homologs of chitin synthases, chitin deacetylases, chitinases and 1-3 beta glucan-synthases in the *Paraphelidum* transcriptome (Figure S2A-D). Specifically, we detected seven homologs of division II chitin synthases [11, 19, 21, 22] in *Paraphelidium* (*Rozella* contains only four [9]). Three of them clustered with class IV chitin synthases, including Microsporidia and *Rozella* homologs (Figure S2A). The remaining four branched within class V/VII enzymes [22], two of them forming a deep-branching group with fungal, mostly chytrid, sequences (Figure 2A). Class V enzymes include a myosin motor thought to intervene in polarized fungal hyphal growth that has been hypothesized to take part in the formation of the germ tube in aphelids and rozellids [12]. Class V chitin synthases were lost in Microsporidia (with the exception of *Mitosporidium*, still retaining, like *Rozella*, one homolog), endowed instead with highly specialized polar tube extrusion mechanisms [12]. Neither spore wall nor polar tube proteins specific to Microsporidia [23] occurred in the *Paraphelidium* transcriptome. Therefore, our data (Figure S21, Table S1-S3C) lend credit to the hypothesis that class V chitin synthases are involved in germ tube polar growth.

Among the rest of chitin-related enzymes, we identified at least 5 different homologs of chitin deacetylase [24, 25] (Figure S2B). We detected at least three class II chitinase homologs (8 sequences), which are ancestral in opisthokonts, containing the Glyco_hydro_18 (PF00704) domain, and a class I chitinase (CTSI) with a Glyco_hydro_19 (PF00182) domain [26]. The latter included the catalytic site, an N-terminal predicted signal peptide and a transmembrane region, suggesting an extracellular chitinase activity. CTSI has a peculiar phylogenetic distribution in eukaryotes, occurring only in Viridiplantae, Fungi, Opisthosporidia and Ecdysozoa (Figure S2C). *Rozella* contains two homologs and Microsporidia at least one; they have N-terminal signal peptides and are predicted to localize extracellularly but lack transmembrane domains. In our phylogenetic tree, opisthosporidian sequences appeared scattered within metazoan and fungal sequences. This might be the result of

hidden paralogy and/or horizontal gene transfer (HGT) (Figure S2C). Regardless its evolutionary origin, aphelid CTSI might be involved in the self-degradation of cyst wall chitin. This might happen both, during their release from chitin-containing resting spores or at the tip of the germ tube during infection, as previously suggested for *Rozella* [9] (Figure 1).

Although not found in chytrids, ß(1,3)-glucan is also considered an idiosyncratic fungal cell-wall component. Surprisingly, we identified two ß(1,3)-glucan synthase (FKS1) homologs, with split Glucan_synthase (PF02364) and FKS1_dom1 (PF14288) domains (both are fused in fungi) (Figure S2D). The presence of FKS1 in aphelids, while being absent in *Rozella* and Microsporidia, traces its origin back to the ancestor of Fungi and Opisthosporidia, suggesting that aphelids retain some ancestral characters.

To feed on the algal cytoplasm, aphelids need to traverse the algal cell wall but, so far, the specific penetration mechanism, whether mechanical (germ-tube penetration at algal cell-wall junctures) or enzymatic (digestion of algal cell-wall components) was uncertain [1, 2]. Scanning electron microscopy (SEM) observations showed both, clear cases of mechanical penetration via junctures but also infecting cysts scattered on the algal cell-wall surface far from junctures (Figure 2C). SEM images additionally confirmed WGA-epifluorescence observations of multiple parasitoid cysts co-infecting the same host cell (Figure 2A-C). Multiple infections open the intriguing possibility for aphelids to have sex within the algal host. *Tribonema* cell walls contain cellulose II based on 1,6-linked glucan (alkali soluble cellulose), 1,3 and 1,4-linked xylose, 1,3-linked rhamnose and, mostly, 1,3, 1,4 and 1,6-linked glucose [27]. We performed sequence similarity searches of known fungal cellulases [28] using the database mycoCLAP [29] , which contains functionally characterized proteins, to look for these enzymes in aphelids, followed by phylogenetic analyses. In support of an enzymatic algal cell-wall penetration, we identified various cellulases in the *Paraphelidium* transcriptome belonging to glycoside-hydrolase families GH3 and GH5. We detected three homologs of the GH3 cellulase ß-glucosidase/xylosidase [30], which is not found in other opisthosporidians but is present in fungi, amoebozoans, several opisthokonts and other protists, as well as bacteria. Our phylogenetic analysis shows that the three aphelid sequences are most closely related to deep-branching opisthokont protists (respectively, *Capsaspora*, choanoflagellates, nucleariids) (Figure S2E). Additionally, we identified at least five GH5 cellulase [31] homologs in *P. tribonemae* (*Rozella* has only one), which were related to three GH5 subfamilies widespread in fungi, GH5_11, GH5_12 and GH5_24 (Figure S2F).

In summary, *Paraphelidium* has a complex life cycle with chitin-containing infective cysts and a diversified set of enzymes involved in chitin metabolism and cellulose hydrolysis.

Primary Metabolism Reminiscent of Free-Living Lifestyles

Analysis of the *Rozella allomycis* genome showed that, like microsporidian parasites, it has significantly reduced metabolic capabilities [9]. To comparatively assess the metabolic potential of aphelids, we investigated the presence of orthologous groups (OGs) related to eight primary metabolism categories (Gene Ontology) in the transcriptome of *Paraphelidium*, using eggNog [32]. We thus identified 933 OGs in *Paraphelidium* and a set of 40 eukaryotic species including representatives of major fungal lineages, opisthokont protists and other eukaryotic parasites (Table S2). Based on their OG distribution, we built a dissimilarity matrix that was analyzed by Principal Coordinate Analysis (PCoA; see Supplementary Methods). The first axis clearly separated *Paraphelidium* from Microsporidia, *Mitosporidium* and *Rozella*, the latter two positioned near one another and having an intermediate position similar to other protist parasites (e.g. *Trypanosoma, Leishmania, Toxoplasma*) (Figures 3A and S3A). *Paraphelidium* positioned at the same level as fungi, *Capsaspora, Corallochytrium* and *Parvularia*, along axis 1. However, axis 2 separated *Paraphelidium* and fungi from the rest of eukaryotes. These relationships were further visualized in cluster analysis of pairwise species comparisons (Figure S3B). The PCoA suggested that *Paraphelidium* had a rich metabolic gene complement, which was made evident by the OG presence/absence heatmap showing that aphelids have a metabolic potential fully comparable to that of (especially chytrid) fungi (Figure 3B).

The most distinctive metabolism categories when comparing *Paraphelidium* and other Opisthosporidia were energy conversion followed by amino acid, nucleotide and lipid transport and metabolism. In all metabolic categories, the aphelid clustered with fungi, and more specifically with chytrids, and sometimes with other free-living opisthokonts (e.g. nucleariids, *Capsaspora*). By contrast, *Rozella* always clustered with *Mitosporidium* either together with other Microsporidia or with other parasitic protists (Figures S3C-F).

To check whether these differences between *Paraphelidium* and *Rozella* affected particular metabolic pathways, we compared the annotated proteins in the two organisms based on KEGG annotation [33]. The comparison of the KEGG general metabolic pathway map showed that, even accounting for the possibility that we missed some genes in the *Paraphelidium*'s transcriptome (e.g. mitochondrial-encoded proteins), the aphelid map contained 200 OGs more than *Rozella* (548 vs 348 OGs) (Figure S3G). In agreement with previous observations, major differences were observed in energy conversion, and amino acid, nucleotide and lipid metabolism. In particular, contrary to *Rozella*, which lacks most subunits of the mitochondrial electron transport chain complex I (NADH dehydrogenase; ETC-I) [9], *Paraphelidium* possesses a practically complete ETC-I (Figure S3H). *Paraphelidium* also possesses wide capabilities related to nucleotide (e.g. purine, uridine and inosine biosynthesis) and amino acid (serine, threonine, methionine, lysine, ornithine, histidine, shikimate

pathway) metabolism, which *Rozella* has lost (Figure S3I). Likewise, the aphelid has also retained pathways for phosphatidylcholine, cholesterol and fatty acid biosynthesis that were subsequently lost in *Rozella* [9]. Most carbohydrate metabolic pathways were conserved in the two opisthosporidians, except for the galactose synthesis and degradation, also lost in *Rozella* (Table S2).

By contrast, compared to *Rozella*, and under the assumption that our transcriptome is rather complete, the aphelid seems to lack several enzymes involved in catecholamine biosynthesis (Table S2). However, some of these are also absent from free-living *Capsaspora, Monosiga, Salpingoeca* or *Spizellomyces*. These compounds are likely involved in cell-cell communication in microbes [34], e.g. parasite-host,, suggesting that they might have a role in rozellid parasitism. The aphelid seems to lack other parasite-specific proteins, such as crinkler, nucleoside H^+-symporters or ATP/ADP-antiporters, which occur in *Rozella* and/or Microsporidia [9].

These observations suggest that *Paraphelidium* has a complex metabolic profile, being functionally closer to free-living protists than to parasites and having affinities with fungi and, to a lesser extent, nucleariids and holozoan protists.

Distinct and Ancestral-Like Phagotrophy-Related Machinery

Like rozellids, aphelids are phagotrophs, but their global metabolism resembles that of osmotrophic fungi. What does their phagocytosis-related proteome look like? The core phagocytic molecular machinery, already present in the last eukaryotic common ancestor [35], is involved in multiple dynamic cell processes (endocytosis, exocytosis, autophagy, protein recycling, etc). Structurally, the phagocytic machinery encompasses various endomembrane organelles (phagosomes, lysosomes, peroxisomes, endoplasmic reticulum, Golgi apparatus), and multiple membrane-trafficking components (signaling pathways, transporters, specific enzymes, cytoskeleton elements, motors, etc.) [36]. To look for phagotrophy-related genes in *Paraphelidium* and the 40 additional eukaryotes used for comparison, we built a presence/absence matrix of 695 KEGG orthologs (OGs) from 11 functional categories (5 KEGG maps and 6 BRITE categories) that aimed at including all necessary proteins to perform phagotrophy; i.e. phagolysosome biogenesis, membrane trafficking and the actin cytoskeleton [36] (see Supplementary Methods and Table S3A). A PCoA showed that *Paraphelidium* and *Rozella* are much closer to one another in terms of phagotrophy- than metabolism-related genes, grouping with fungi and far from Microsporidia (Figure 3C and S4A). This pattern was also evident in the presence/absence matrix (Figure 3D and S4B). In both, the two opisthosporidians clustered with early-branching fungi: chytridiomycetes (*Spizellomyces*, *Batrachochytrium*, *Gonapodya*) and blastocladiomycetes (*Allomyces*, *Catenaria*, *Blastocladiella*). A Venn diagram including these 4 lineages (*Rozella*, *Paraphelidium*, chytridiomycetes, blastocladiomycetes) showed that only 494 of the 695

phagotrophy-related OGs were present in the last common ancestor of opisthosporidians and early-branching fungi and only 282 OGs of those were shared by the 4 taxa (Figure S4C). No particular KEGG category varied notably between the 4 lineages, but each lineage had particular differences mainly in membrane-trafficking proteins (Table S3B; Figures S4D-F). Early-branching fungi shared more cytoskeleton, membrane-trafficking and phagotrophy-related OGs with *Paraphelidium* (57) than with *Rozella* (26), in agreement with a more reduced *Rozella* genome. At the same time, the *Paraphelidium* protein set seems neither to overlap with *Rozella* nor to retain more OGs than early-branching fungi, thus arguing for a differential recruitment of ancestral membrane-remodeling proteins for diverse cellular tasks.

In order to gain more insights into the diversification of the actin cytoskeleton toolkit of fungi and opisthosporidians, we analyzed the evolution of myosin motor proteins. The myosin toolkit of *Paraphelidium* contains a mixture of classical fungal families and others previously identified in holozoans (animals and their unicellular relatives; Table S3C). We recovered diversified class I myosins in *Paraphelidium*, *Spizellomyces* and nucleariids (Figure S4G), with paralogs of Ic/h and Ik families, previously described only in holozoans [37]. *Paraphelidium,* nucleariids and *Gonapodya* also possess homologs of class XV myosins, formerly thought holozoan-specific [37]. In addition, the aphelid not only possesses homologs of the V/VII myosin-motor family associated to chitin-synthase (see above; Figure S2A), but also myosins of If (pan-eukaryotic), II (amorphean), and XXII (opisthokont) classes, which clustered with previously described fungal homologs (Table S3C; Figure S4H). Thus, compared with the ancestral opisthokont myosin complement [37], *Paraphelidium* retains all but one myosin class (VI), with homologs of all myosin families present in fungi (If, II, V, XVII - chitin synthase) plus four additional families (Ic/h, Ik, XV and XXII) that were lost in 'canonical' fungi. This suggests an independent step-wise simplification of the myosin complement in fungi, *Rozella* and Microsporidia, with *Paraphelidium,* nucleariids and chytrids retaining ancestral classes.

Further supporting the view of an elaborate actin-cytoskeletal machinery compared to *Rozella*, we identified genes involved in pseudopodia formation in *Paraphelidium*, such as WASP and SCAR/WAVE activators of branched actin assembly proteins [38], which are essential to build actin-based pseudopodia (filopodia) and are fundamental for alpha-motility [39] (Table S1). The aphelid zoospore motility (Video 1) resembles that of chytrid fungi [39]. Interestingly, although *Rozella* also contains a WASP homolog, it lacks the WH1 domain (PF00568). This might explain why filopodia have not been described in rozellid zoospores [40-42].

Aphelids and the Free-Living Ancestor of Fungi and Opisthosporidia

From an evolutionary perspective, the current situation in the holomycotan branch (including Fungi) of the eukaryotic super-group Opisthokonta mirrors that of the Holozoa (including Metazoa), where the discovery of deeply-branching unicellular protists that possess genes thought unique to animals continues to challenge previous evolutionary schemes about the emergence of Metazoa [15, 43]. Thus, the discovery that widely diverse environmental sequences formed a deeply-branching clade with respect to fungi including the parasitic *Rozella*, rozellids [3], subsequently called Cryptomycota [4], Rozellomycota [44] or Rozellosporidia [45], triggered the discussion of what Fungi actually are [18, 46]. A debate further nourished by the discovery that aphelids, another widely diverse group of parasites of algae [47], formed a large clade with rozellids and Microsporidia based on rRNA and RNA-polymerase genes [1, 5, 45, 48]. This seemingly monophyletic clade was named Opisthosporidia and branched as sister to classical Fungi (the monophyletic clade including from chytrids and their relatives to the Dikarya [46]) in phylogenetic trees [1]. Lately, many mycologists include the three opisthosporidian lineages, Aphelida, Rozellosporidia and Microsporidia, within Fungi [9, 18, 49, 50]. Some authors even incorporate as fungi the free-living phagotrophic chitin-lacking nucleariids, thus pushing the limits of Fungi to incorporate all holomycotan (possibly a confusing name) lineages, despite asserting that "...the kingdom Fungi is characterized by osmotrophic nutrition across a chitinous cell wall..." [51]. However, unlike fungi, the most basal opisthosporidian branches (aphelids and rozellids) are phagotrophs, lack a chitin cell wall at the vegetative stage and are endobiotic, having a unique mode of penetration into the host [1, 2, 46, 48]. Also, because *Rozella* has a very reduced genome, a typical trait of streamlined parasites, some authors inferred a parasitic nature for the ancestor of Fungi [9].

The study of the *Paraphelidium tribonemae* transcriptome helps to clarify this controversy. First, our multi-gene phylogenetic analysis confirms the Opisthosporidia monophyly and its sisterhood to Fungi, placing aphelids as the earliest-branching opisthosporidian lineage (Figure 1). The establishment of a solid phylogenomic framework for Opisthosporidia and the comparative study of the *Paraphelidium* transcriptome allow better inferring ancestral states for Opisthosporidia, Fungi and their common ancestor (Figure 4). *Paraphelidium* has a complex metabolism resembling that of free-living chytrids but feeds by phagotrophy as free-living nucleariids and holozoan protists (Figure 3). This suggests that aphelids are 'borderline' parasites that have not undergone the reduction process that characterizes *Rozella* and all members along the microsporidian branch [6, 8, 9]. This also advocates for a free-living opisthosporidian ancestor that specialized in endobiotic predation and had a complex life cycle that included amoeboflagellate zoospores and chitin-containing infective cysts (Figure 4). By contrast, the fungal ancestor was a free-living osmotroph that had amoeboflagellated zoospores and

chitin in the vegetative stage. Consequently, the common ancestor of both Opisthosporidia and Fungi likely was a free-living phagotroph with differentiated cell stages having at least a free-living phagotrophic amoeboflagellate state and a chitin-containing cyst (Figure 4). Because chitin can no longer be considered a fungal synapomorphy and the two groups share the ancestral occurrence of amoeboflagellate zoospores in a complex life cycle to the exclusion of known nucleariids (which lack zoospores), we propose to call the Fungi + Opisthosporidia clade Zoosporia. From the zoosporian free-living ancestor, the fungal lineage lost phagotrophy, acquiring its ecologically successful osmotrophy, whereas opisthosporidians became endobiotic phagotrophic predators that readily evolved into obligate parasites with more complex life cycles. The acquisition of additional genomes/transcriptomes for deep-branching species at the base of both Zoosporia lineages should help further validate and refine this evolutionary scenario.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures S1-S4, three tables S1 to S3, one video and one supplementary file (all_trees.txt).

AUTHOR CONTRIBUTIONS

G.T., D.M. and P.L-G. conceived and coordinated the study, and wrote the manuscript. G.T. performed culture cleaning, RNA extraction, *de novo* transcriptome assembly, phylogenomics and comparative genomic analyses. G.T. and X. G-B. cleaned the protein set from contamination. X. G-B. and A. S-B. performed myosin comparative genomic analyses. X. G-B. carried out multivariate statistical analyses of metabolic genes. G.T. and J.A.B. carried out phagolysosome protein analyses. G.T. and S.K. maintained cultures and performed WGA staining and imaging. E.V. performed SEM fixation and imaging. All authors commented on the manuscript.

ACKNOWLEDGEMENTS

Research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Program (ERC Grant Agreement 322669 ProtistWorld). GT was also financed by the European Marie Sklodowska-Curie Action (704566 AlgDates). S.K. was supported by RSF grant No 16-04-10302 and "Jean d'Alembert" program (Université Paris-Saclay). We thank Victoria S. Tcvetkova for micromanipulating infected *Tribonema* sp. filaments in view of establishing clean aphelid cultures, Alex de Mendoza for providing CHS alignments, Philippe Deschamps, Rafael Ponce, Guillaume Reboul, Ana Gutiérrez-Preciado, Ares Rocañín-Arjó and Ricardo Rodríguez de la Vega for discussion on bioinformatics and statistics aspects, and Peter Kubicek and Bernard Hernissat for helpful discussion about cellulases.

REFERENCES

1. Karpov, S.A., Mamkaeva, M.A., Aleoshin, V.V., Nassonova, E., Lilje, O., and Gleason, F.H. (2014). Morphology, phylogeny, and ecology of the aphelids (Aphelidea, Opisthokonta) and proposal for the new superphylum Opisthosporidia. Front Microbiol *5*, 112.
2. Karpov, S.A., Tcvetkova, V.S., Mamkaeva, M.A., Torruella, G., Timpano, H., Moreira, D., Mamanazarova, K.S., and Lopez-Garcia, P. (2017). Morphological and genetic diversity of Opisthosporidia: new aphelid *Paraphelidium tribonemae* gen. et sp. nov. J Eukaryot Microbiol *64*, 204-212.
3. Lara, E., Moreira, D., and Lopez-Garcia, P. (2010). The environmental clade LKM11 and Rozella form the deepest branching clade of fungi. Protist *161*, 116-121.
4. 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.
5. Karpov, S.A., Mikhailov, K.V., Mirzaeva, G.S., Mirabdullaev, I.M., Mamkaeva, K.A., Titova, N.N., and Aleoshin, V.V. (2013). Obligately phagotrophic aphelids turned out to branch with the earliest-diverging fungi. Protist *164*, 195-205.
6. Corradi, N. (2015). Microsporidia: Eukaryotic intracellular parasites shaped by gene loss and horizontal gene transfers. In Annual Review of Microbiology, Vol 69, Volume 69. (Palo Alto: Annual Reviews), pp. 167-183.
7. Letcher, P.M., Lopez, S., Schmieder, R., Lee, P.A., Behnke, C., Powell, M.J., and McBride, R.C. (2013). Characterization of *Amoeboaphelidium protococcarum*, an algal parasite new to the Cryptomycota isolated from an outdoor algal pond used for the production of biofuel. Plos One *8*, 15.
8. Mikhailov, K.V., Simdyanov, T.G., and Aleoshin, V.V. (2017). Genomic survey of a hyperparasitic microsporidian *Amphiamblys* sp. (Metchnikovellidae). Genome Biol Evol *9*, 454-467.
9. 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. Current Biology *23*, 1548-1553.
10. Simao, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics *31*, 3210-3212.
11. Torruella, G., de Mendoza, A., Grau-Bove, X., Anto, M., Chaplin, M.A., del Campo, J., Eme, L., Perez-Cordon, G., Whipps, C.M., Nichols, K.M., et al. (2015). Phylogenomics reveals convergent evolution of lifestyles in close relatives of animals and fungi. Curr Biol *25*, 2404-2410.
12. Haag, K.L., James, T.Y., Pombert, J.F., Larsson, R., Schaer, T.M., Refardt, D., and Ebert, D. (2014). Evolution of a morphological novelty occurred before genome compaction in a lineage of extreme parasites. Proc Natl Acad Sci U S A *111*, 15480-15485.
13. Lartillot, N., and Philippe, H. (2004). A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol Biol Evol *21*, 1095-1109.
14. Le, S.Q., Lartillot, N., and Gascuel, O. (2008). Phylogenetic mixture models for proteins. Philosophical Transactions of the Royal Society B: Biological Sciences *363*, 3965-3976.
15. Grau-Bove, X., Torruella, G., Donachie, S., Suga, H., Leonard, G., Richards, T.A., and Ruiz-Trillo, I. (2017). Dynamics of genomic innovation in the unicellular ancestry of animals. Elife *6*, pii: e26036.
16. Minh, B.Q., Nguyen, M.A., and von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol *30*, 1188-1195.
17. Wang, H.C., Minh, B.Q., Susko, E., and Roger, A.J. (2017). Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. Syst Biol.
18. Richards, T.A., Leonard, G., and Wideman, J.G. (2017). What Defines the "Kingdom" Fungi? Microbiology spectrum *5*.
19. Ruiz-Herrera, J., and Ortiz-Castellanos, L. (2010). Analysis of the phylogenetic relationships and evolution of the cell walls from yeasts and fungi. FEMS yeast research *10*, 225-243.

20. James, T.Y., and Berbee, M.L. (2012). No jacket required--new fungal lineage defies dress code: recently described zoosporic fungi lack a cell wall during trophic phase. Bioessays *34*, 94-102.
21. Merzendorfer, H. (2011). The cellular basis of chitin synthesis in fungi and insects: common principles and differences. European journal of cell biology *90*, 759-769.
22. Goncalves, I.R., Brouillet, S., Soulie, M.C., Gribaldo, S., Sirven, C., Charron, N., Boccara, M., and Choquer, M. (2016). Genome-wide analyses of chitin synthases identify horizontal gene transfers towards bacteria and allow a robust and unifying classification into fungi. BMC Evol Biol *16*, 252.
23. Vavra, J., and Lukes, J. (2013). Microsporidia and 'the art of living together'. Advances in parasitology *82*, 253-319.
24. Zhao, Y., Park, R.D., and Muzzarelli, R.A. (2010). Chitin deacetylases: properties and applications. Mar Drugs *8*, 24-46.
25. Viens, P., Lacombe-Harvey, M.E., and Brzezinski, R. (2015). Chitosanases from Family 46 of Glycoside Hydrolases: From Proteins to Phenotypes. Mar Drugs *13*, 6566-6587.
26. Udaya Prakash, N.A., Jayanthi, M., Sabarinathan, R., Kangueane, P., Mathew, L., and Sekar, K. (2010). Evolution, homology conservation, and identification of unique sequence signatures in GH19 family chitinases. J Mol Evol *70*, 466-478.
27. Cleare, M., and Percival, E. (1972). Carbohydrates of the fresh water alga *Tribonema aequale.* I. Low molecular weight and polysaccharides. British Phycological Journal *7*, 185-193.
28. Kubicek, C.P., Starr, T.L., and Glass, N.L. (2014). Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. Annu Rev Phytopathol *52*, 427-451.
29. Murphy, C., Powlowski, J., Wu, M., Butler, G., and Tsang, A. (2011). Curation of characterized glycoside hydrolases of fungal origin. In Database (Oxford). p. bar020.
30. McAndrew, R.P., Park, J.I., Heins, R.A., Reindl, W., Friedland, G.D., D'Haeseleer, P., Northen, T., Sale, K.L., Simmons, B.A., and Adams, P.D. (2013). From soil to structure, a novel dimeric beta-glucosidase belonging to glycoside hydrolase family 3 isolated from compost using metagenomic analysis. J Biol Chem *288*, 14985-14992.
31. Aspeborg, H., Coutinho, P.M., Wang, Y., Brumer, H., 3rd, and Henrissat, B. (2012). Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol *12*, 186.
32. Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., Rattei, T., Mende, D.R., Sunagawa, S., Kuhn, M., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res *44*, D286-293.
33. Okuda, S., Yamada, T., Hamajima, M., Itoh, M., Katayama, T., Bork, P., Goto, S., and Kanehisa, M. (2008). KEGG Atlas mapping for global analysis of metabolic pathways. Nucleic Acids Res *36*, W423-426.
34. Roshchina, V.V. (2016). New Trends and Perspectives in the Evolution of Neurotransmitters in Microbial, Plant, and Animal Cells. Adv Exp Med Biol *874*, 25-77.
35. Yutin, N., Wolf, M.Y., Wolf, Y.I., and Koonin, E.V. (2009). The origins of phagocytosis and eukaryogenesis. Biol Direct *4*, 9.
36. Boulais, J., Trost, M., Landry, C.R., Dieckmann, R., Levy, E.D., Soldati, T., Michnick, S.W., Thibault, P., and Desjardins, M. (2010). Molecular characterization of the evolution of phagosomes. Mol Syst Biol *6*, 423.
37. Sebe-Pedros, A., Grau-Bove, X., Richards, T.A., and Ruiz-Trillo, I. (2014). Evolution and classification of myosins, a paneukaryotic whole-genome approach. Elife *6*, 290-305.
38. Sebe-Pedros, A., Burkhardt, P., Sanchez-Pons, N., Fairclough, S.R., Lang, B.F., King, N., and Ruiz-Trillo, I. (2013). Insights into the origin of metazoan filopodia and microvilli. Mol Biol Evol *30*, 2013-2023.
39. Fritz-Laylin, L.K., Lord, S.J., and Mullins, R.D. (2017). WASP and SCAR are evolutionarily conserved in actin-filled pseudopod-based motility. J Cell Biol *216*, 1673-1688.
40. Held, A.A. (1975). Zoospore of *Rozella allomycis* - ultrastructure. Can J Bot *53*, 2212-2223.

41. Held, A.A. (1980). Development of Rozella in Allomyces - A single zoospore produces numerous zoosporangia and resistant sporangia. Can J Bot *58*, 959-979.
42. Letcher, P.M., Longcore, J.E., Quandt, C.A., Leite, D.D., James, T.Y., and Powell, M.J. (2017). Morphological, molecular, and ultrastructural characterization of *Rozella rhizoclosmatii*, a new species in Cryptomycota. Fungal biology *121*, 1-10.
43. Hehenberger, E., Tikhonenkov, D.V., Kolisko, M., Del Campo, J., Esaulov, A.S., Mylnikov, A.P., and Keeling, P.J. (2017). Novel predators reshape Holozoan phylogeny and reveal the presence of a two-component signaling system in the ancestor of animals. Curr Biol *27*, 2043-2050.e2046.
44. Corsaro, D., Walochnik, J., Venditti, D., Steinmann, J., Muller, K.D., and Michel, R. (2014). Microsporidia-like parasites of amoebae belong to the early fungal lineage Rozellomycota. Parasitology research *113*, 1909-1918.
45. Karpov, S.A., Torruella, G., Moreira, D., Mamkaeva, M.A., and Lopez-Garcia, P. (2017). Molecular phylogeny of *Paraphelidium letcheri* sp. nov. (Aphelida, Opisthosporidia). J Eukaryot Microbiol *64*, 573-578.
46. 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.
47. Gromov, B.V. (2000). Algal parasites of the genera *Aphelidium, Amoeboaphelidium,* and *Pseudaphelidium* from the Cienkovski's "monadinea" group as representatives of a new class. Zoologichesky Zhurnal *79*, 517-525.
48. Karpov, S.A., Mamkaeva, M.A., Benzerara, K., Moreira, D., and Lopez-Garcia, P. (2014). Molecular phylogeny and ultrastructure of *Aphelidium* aff. *melosirae* (Aphelida, Opisthosporidia). Protist *165*, 512-526.
49. Tedersoo, L., and Lindahl, B. (2016). Fungal identification biases in microbiome projects. Environ Microbiol Rep.
50. Spatafora, J.W., Aime, M.C., Grigoriev, I.V., Martin, F., Stajich, J.E., and Blackwell, M. (2017). The fungal tree of life: from molecular systematics to genome-scale phylogenies. Microbiology spectrum *5*.
51. Berbee, M.L., James, T.Y., and Strullu-Derrien, C. (2017). Early diverging Fungi: Diversity and impact at the dawn of Terrestrial life. Annu Rev Microbiol *71*, 41-60.

Figure legends

Figure 1. Phylogenomic Tree Showing the Position of the Aphelid *Paraphelidium tribonemae* Within Opisthosporidia

Bayesian phylogenetic tree based on single-copy protein domains inferred using a CAT-Poisson model. Statistical supports indicated at each node correspond, from left to right, to PhyloBayes Bayesian posterior probabilities; IQ-TREE ML ultrafast-bootstrap support using the C60 model; and IQ-TREE ML non-parametric bootstrap (BS) using the PMSF model. Branches with maximum support values (BPP = 1 and UFBS/BS = 100%) are indicated by black circles. The inlet shows a schematic cell cycle of *P. tribonemae*. Briefly, infecting cysts (red wall), deliver an amoeboid trophont to an algal filament cell via an infection tube; the trophont engulfs the algal cytoplasm by phagocytosis, leaving a growing residual body (dark red particle); after nuclear and cell division, a mature sporangium releases amoeboflagellated zoospores (occasionally seen as amoeboid only) that get out the algal cell wall and close the life cycle [2]. To see the uncondensed tree and additional phylogenomic trees, see Figure S1.

Figure 2. Zoospores and Chitin-bearing Infective Cysts of *Paraphelidium tribonemae*

(A) Filament of *Tribonema gayanum* infected by *P. tribonemae* seen under optical microscopy and (B) the same filament stained with fluorescent wheat germ agglutinin (WGA) showing the presence of chitin in infective cysts under epifluorescence microscopy. (C) False-colored scanning-electron microscopy image of a filament infected by several *P. tribonemae* cysts (pedunculated rounded structures). The algal host filament is colored in green and parasite cysts in pink. Note that one cyst germ tube is penetrating the host cell by a cell-wall juncture (white arrowhead) and two cysts are broken (black arrowheads), showing the penetration channel. (D) Amoeboid zoospore (infrequent). (E) Amoeboflagellated zoospore. Scale bar: A = 5 µm, B-F = 1 µm. Phylogenetic trees related to chitin and other cell-wall synthesis and degradation-related enzymes are shown in Figure S2. Zoospore motility can be seen in Video 1.

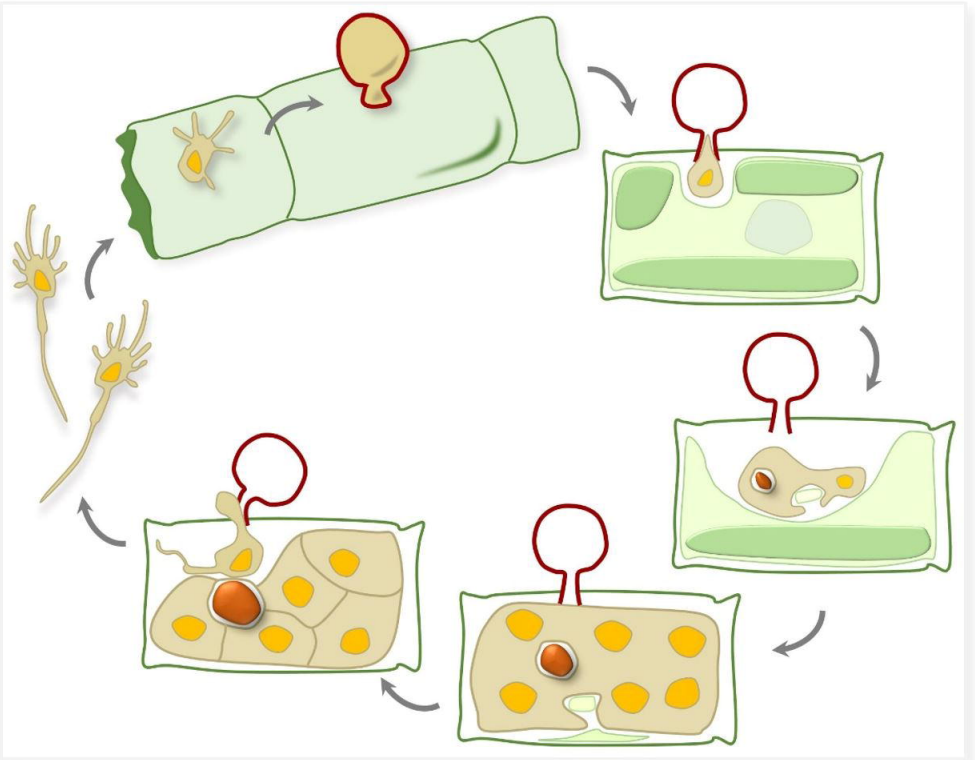
Figure 3. Complexity of *Paraphelidium tribonemae* Metabolism and Cytoskeleton-Trafficking-Phagotrophy-related Proteome

(A) Principal Coordinate Analysis (PCoA) and (B) binary heat-map and species clustering based on the presence/absence of 933 orthologous genes (OGs) belonging to 8 primary metabolism eggNOG categories across 40 eukaryotic genomes/transcriptomes (see Supplementary Methods). (C) PCoA and (D) binary heat-map and species clustering based on the presence/absence of 695 KEGG orthologs (OGs) related to cytoskeleton, membrane-trafficking and phagotrophy, which were selected from 11 KEGG categories (see Supplementary Methods). Species are color-coded according to their taxonomic

assignment; particular members within Opisthosporidia are highlighted. See also Figures S3-S4 and Tables S2-S3.

Figure 4. Early evolution of Fungi and Opisthosporidia

Schematic representation of evolutionary relationships between Fungi and Opisthosporidia, forming the Zoosporia, within the holomycotan branch of Opisthokonts. Inferred key ancestral features and life cycle stages are depicted at the corresponding ancestral nodes.



Outgroup

Holozoa

Nucleariids

Parvularia atlantis

Fonticula alba

Fungi

Allomyces macrogynus

Catenaria anguillulae

Blastocladiella emersonii

1/90/46

Gonapodya prolifera

Spizellomyces punctatus

Batrachochytrium dendrobatidis

1/96/94

Conidiobolus coronatus

Coemansia reversa

Schizosaccharomyces pombe

Saitoella complicata

Ustilago maydis

Phanerochaete chrysosporium

Rhizophagus irregularis

Mortierella verticillata

1/94/85

Mucor circineloides

Lichtheimia hyalospora

Paraphelidium tribonemae **(Aphelida)**

Rozella allomycis

1/79/80

Mitosporidium daphniae

Amphiamblys sp.

0.95/1/1

Opisthosporidia

Nematocida parisii

Edhazardia aedis

Pseudoloma neurophilia

Vavraia culicis

Trachipleistophora hominis

0.85/56/83

Antonospora locustae

Anncaliia algerae

Vittaforma corneae

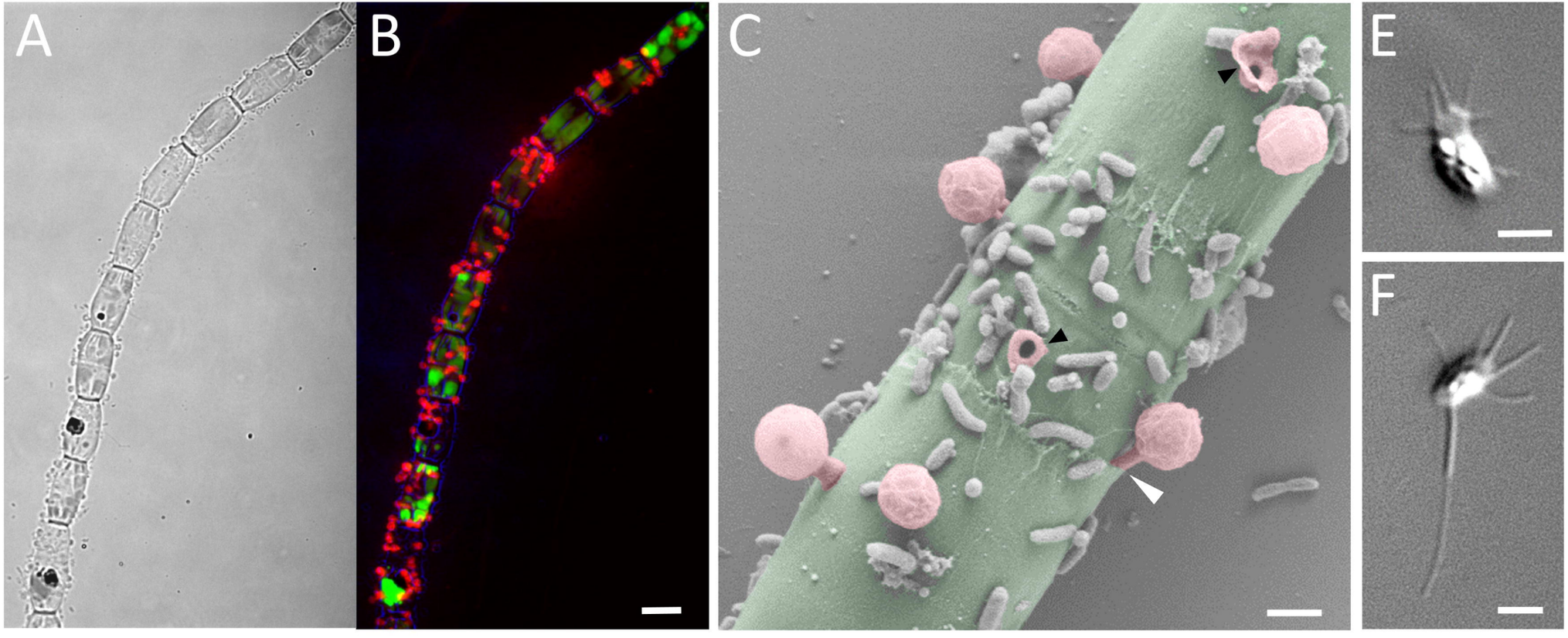
Enterocytozoon bieneusi

Nosema ceranae

Ordospora colligata

Encephalitozoon intestinalis

0.3


A
B
C
E
F

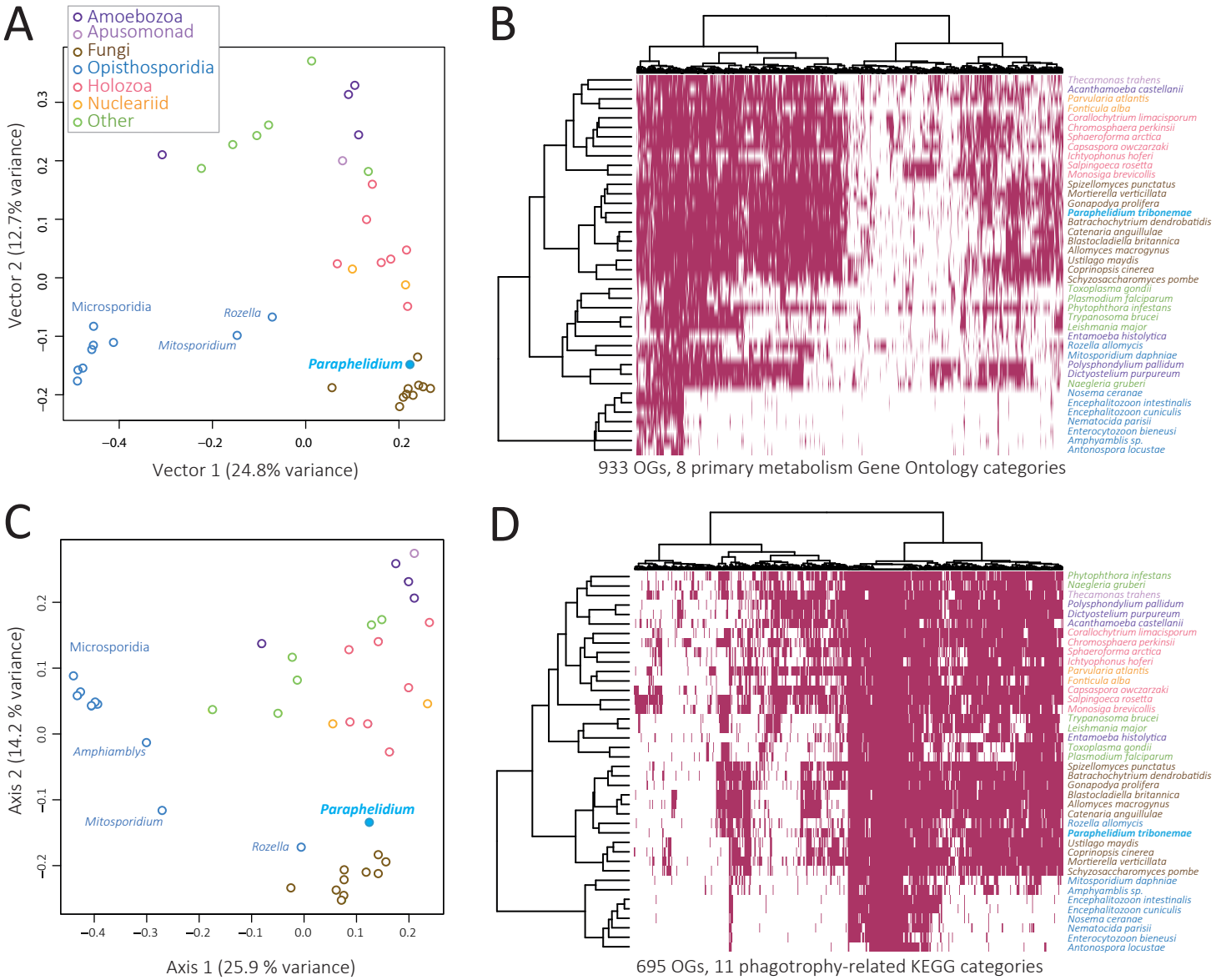
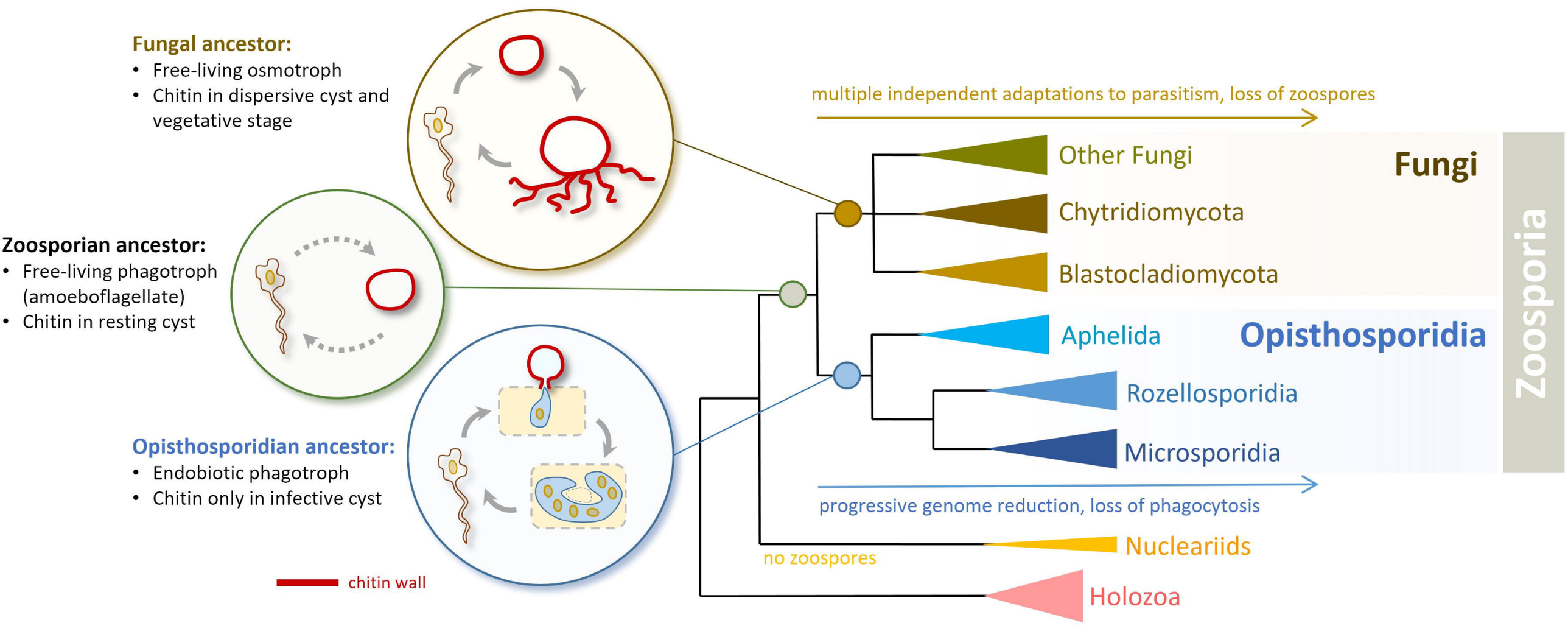


Figure 3. Torruella *et al.*


Fungal ancestor:
• Free-living osmotroph
• Chitin in dispersive cyst and vegetative stage
Zoosporian ancestor:
• Free-living phagotroph (amoeboflagellate)
• Chitin in resting cyst
Opisthosporidian ancestor:
• Endobiotic phagotroph
• Chitin only in infective cyst
chitin wall
multiple independent adaptations to parasitism, loss of zoospores
Other Fungi
Chytridiomycota
Blastocladiomycota
Fungi
Aphelida
Opisthosporidia
Rozellosporidia
Microsporidia
progressive genome reduction, loss of phagocytosis
no zoospores
Nucleariids
Holozoa
Zoosporia