ORIGINAL ARTICLE

Clarifying the Relationships between Microsporidia and Cryptomycota

David Bass[a,b], Lucas Czech[c], Bryony A. P. Williams[d], Cédric Berney[e], Micah Dunthorn[f], Frederic Mahé[g], Guifré Torruella[h], Grant D. Stentiford[a] & Tom A. Williams[i]

a Pathology and Microbial Systematics Theme, Centre for Environment, Fisheries and Aquaculture Science (Cefas), Barrack Road, The Nothe, Weymouth DT4 8UB, UK
b Department of Life Sciences, The Natural History Museum, London SW7 5BD, UK
c Heidelberg Institute for Theoretical Studies, Schloß-Wolfsbrunnenweg, Heidelberg 69118, Germany
d Biosciences, College of Life and Environmental Sciences, University of Exeter, Geoffrey Pope Building, Stocker Road, Exeter EX4 4QD, UK
e Sorbonne Université & CNRS, UMR 7144 (AD2M), Station Biologique de Roscoff, Place Georges Teissier, Roscoff 29680, France
f Department of Ecology, University of Kaiserslautern, Kaiserslautern, Germany
g CIRAD, UMR LSTM, Montpellier, France
h Ecologie Systématique Evolution, CNRS, Université Paris-Sud, AgroParisTech, Université Paris-Saclay, Orsay, France
i School of Biological Sciences, University of Bristol, 24 Tyndall Avenue, Bristol BS8 1TQ, UK

ABSTRACT

Some protists with microsporidian-like cell biological characters, including *Mitosporidium*, *Paramicrosporidium*, and *Nucleophaga*, have SSU rRNA gene sequences that are much less divergent than canonical Microsporidia. We analysed the phylogenetic placement and environmental diversity of microsporidian-like lineages that group near the base of the fungal radiation and show that they group in a clade with metchnikovellids and canonical microsporidians, to the exclusion of the clade including *Rozella*, in line with what is currently known of their morphology and cell biology. These results show that the phylogenetic scope of Microsporidia has been greatly underestimated. We propose that much of the lineage diversity previously thought to be cryptomycotan/rozellid is actually microsporidian, offering new insights into the evolution of the highly specialized parasitism of canonical Microsporidia. This insight has important implications for our understanding of opisthokont evolution and ecology, and is important for accurate interpretation of environmental diversity. Our analyses also demonstrate that many opisthosporidian (aphelid+rozellid+microsporidian) SSU V4 OTUs from Neotropical forest soils group with the short-branching Microsporidia, consistent with the abundance of their protist and arthropod hosts in soils. This novel diversity of Microsporidia provides a unique opportunity to investigate the evolutionary origins of a highly specialized clade of major animal parasites.

MICROSPORIDIA are conventionally considered as highly derived parasitic protists sister to *Rozella* or diverging as the next branch below the fungi (James et al. 2013). Microsporidia, Balbiani 1882 display a suite of distinctive cell biological characters related to their obligate parasitic lifestyle, including a characteristic spore-extrusion apparatus (represented most conspicuously by the polar filament and its terminal anchoring disc) (Franzen 2004; Vávra and Lukeš 2013), unwalled intracellular trophic (meront) stages, and multiwalled spores produced by merogony or other forms of proliferation (Vávra and Larsson 1999). Microsporidia lack canonical Golgi apparatus (Beznoussenko et al. 2007; Vávra and Larsson 1999; Vávra and Lukeš 2013) and their mitochondria have been highly reduced to mitosomes (reviewed in Dean et al. 2016). These mitosomes are unable to generate their own ATP through oxidative phosphorylation, requiring energy to be imported from the host via nucleotide transporters.

Microsporidia also lack flagella and an apparent capacity for phagocytosis. The known diversity of Microsporidia is large, comprising approximately 1,300 described species (Vávra and Lukeš 2013), forming a long-branched clade (hereafter referred to as LB-Microsporidia) in SSU rRNA and multigene phylogenies (James et al. 2013). Metchnikovellids (e.g. *Amphiamblys*, *Amphiacantha*) have traditionally been referred to as atypical, "primitive" microsporidians, but share many characters with LB-Microsporidia, and were recently shown to branch as sisters to them (Mikhailov et al. 2017).

Microsporidia are known primarily as parasites of invertebrates and vertebrates (including humans), but are also known as endosymbionts of ciliates (Fokin 2012; Fokin et al. 2008), and hyperparasites in protists: metchnikovellids are parasites of gregarines, protistan gut symbionts of many invertebrates, and *Hyperspora aquatica* is a hyperparasite of the paramyxid *Marteilia cochillia*, a serious pathogen of European cockles (Stentiford et al. 2017). *Rozella* species are zoosporic biotrophic parasites of oomycetes, chytrids, and Blastocladiomycota (Spatafora et al. 2017). Increased attention has recently been given to a large diversity of lineages shown by phylogenetic analyses including environmental sequences to be related to Microsporidia, rozellids, and aphelids. The first to highlight this diversity were Lara et al. (2010) and Jones et al. (2011), who showed a large diversity of environmental sequences which, in the absence of microsporidian sequences, group with *Rozella* in phylogenetic trees. These have been referred to as Rozellida (Fig. 1 in Lara et al. 2010), Rozellomycota (Corsaro et al. 2014a,b), and Cryptomycota (Fig. 1 in Jones et al. 2011).

A few other sequences branching between rozellids and LB-microsporidia represent microsporidia-like protists which have been morphologically characterized: *Nucleophaga* (Corsaro et al. 2014a, 2016), *Paramicrosporidium* (originally described as a microsporidian; Michel et al. 2000, 2009), and *Mitosporidium* (Haag et al. 2014).

These three genera clearly share some features with classical LB-Microsporidia, including forms of polar filaments (not necessarily functional as extrusion apparatus), unwalled intracellular meront stages, and nonflagellated spores; but in other respects, they are dissimilar, *Nucleophaga* and *Paramicrosporidium* being the least structurally similar to LB-Microsporidia. *Mitosporidium*, sometimes referred to as the earliest branching microsporidian (Mikhailov et al. 2017; Quandt et al. 2017), has a mitochondrion, albeit lacking Complex I of the oxidative phosphorylation pathway. *Paramicrosporidium* has a canonical fungal mitochondrial genome, and shares more gene content with distantly related fungi than with its closest relatives (Quandt et al. 2017). *Mitosporidium* and *Nucleophaga* have possibly nonhomologous finger-like extensions in naked intranuclear trophic stages similar to those of *Rozella* (Corsaro et al. 2014b; Haag et al. 2014).

Most published phylogenetic analyses that include Crypto/Rozellomycota/rozellid environmental sequences do not also include LB-Microsporidia. Those that do (Corsaro et al. 2016 (Fig. 3); Tedersoo et al. 2017) suggest that

Rozella is sister to a highly diverse clade comprising LB-Microsporidia, the microsporidian-like protists described above, and a large diversity of uncharacterized environmental sequences. In this study we investigate these relationships further, integrating morphological, phylogenetic, and sequence diversity data, to determine the phylogenetic and taxonomic boundaries of microsporidia and their immediate relatives.

MATERIALS AND METHODS

The nr nucleotide GenBank database was blastn searched using seed SSU rRNA gene sequences of characterized microsporidia-like protists, metchnikovellids, deeply branching LB-Microsporidia (clades 2 and 3 Stentiford et al. 2017), and the phylogenetic diversity of "Cryptomycota" and aphelids in Karpov et al. (2014). The top 50 matches for each sequence were downloaded, aligned with mafft e-ins-i (Katoh and Standley 2013), deduplicated, and a preliminary tree constructed on the basis of which the number of closely related sequences was reduced, retaining the longest possible sequences. The shorted branched LB-Microsporidia *Janacekia*, *Trichonosema*, and *Bacillidium* were selected to represent LB-Microsporidia in order to reduce the possibility of phylogenetic artefacts caused by LBA. Published phylogenies (e.g. Stentiford et al. 2017) show that these form a very robust monophyletic group with all other LB-Microsporidia, so it is reasonable to use them as a proxy for the whole group in this study. A Bayesian phylogeny was inferred under the CAT+GTR+Gamma (4) model in PhyloBayes-MPI 1.7 (Lartillot et al. 2013). Convergence among four MCMC chains was assessed by comparing the discrepancies in bipartition frequencies and in a range of continuous model parameters, along with the effective sample sizes of the continuous parameters. A consensus tree was built once all discrepancies were < 0.1, with sample sizes > 100. A maximum likelihood phylogeny was estimated under the GTR+Gamma(4)+F model in IQ-Tree (Nguyen et al. 2015), with 200 traditional nonparametric bootstraps. The distribution of microsporidian characters (Table 1) was constructed from the literature. OTUs assigned to any of "Opisthosporidia" (Karpov et al. 2014); "Cryptomycota" (Jones et al. 2011), "Holomycota" (Liu et al. 2009), and "Microsporidia" and otherwise unassigned "fungi" by the taxonomic assignment algorithms of each study were extracted from environmental amplicon sequencing data of tropical forest soils clustered into OTUs by Swarm v2.1.5 (Mahé et al. 2015) and European coastal water and sediment samples (Logares et al. 2014). We inferred a maximum likelihood tree from our reference database using RaxML v8.2.8 (Stamatakis 2014). The OTUs were then aligned to the reference database using PaPaRa (Berger and Stamatakis 2011) and placed on the tree by RaxML-EPA (Berger et al. 2011). The distribution of placements (Fig. 2, 3) was created with Genesis (http://genesis-lib.org/) and visualized with FigTree (http://tree.bio.ed.ac.uk/software/figtree/). Based on the 49 flagellar toolkit proteins assembled for a previous study (Torruella et al. 2015), we searched by

BLAST in available early-branching microsporidian pre-dicted protein sets from genomic data, in *Mitosporidium daphniae* (Haag et al. 2014), *Paramicrosporidium* (Quandt et al. 2017), and *Amphiamblys* sp. (Mikhailov et al. 2017), using an e-value threshold of 1e-10 and manual scrutiny.

RESULTS AND DISCUSSION

An expanded Microsporidia

We constructed an SSU rRNA gene sequence alignment including a comprehensive selection of the microsporidian-like protists and their related environmental sequences, metchnikovellids, and representatives of short-branch LB-Microsporidia (see Methods), aphelids, rozellids, and related unclassified groups. Maximum Likelihood and Bayesian phylogenetic analyses (Fig. 1) showed that the expected metchnikovellid+LB-Microsporidia clade emerges from within a diverse and well-supported (Bayesian poste-rior probability (BPP) = 0.98; Maximum Likelihood boot-strap 96%) clade including *Paramicrosporidium*, *Nucleophaga*, *Mitosporidium*, LKM-11, and many other environmental sequences referred to as rozello-/cryptomy-cotans. *Rozella* formed a clade with high support with a relatively small number of environmental sequences.

It is well recognized that SSU rRNA gene data alone are unable to resolve deep phylogenetic relationships. How-ever, although not a substitute for multigene data, the high taxon sampling afforded by SSU provides phyloge-netic information currently inaccessible for phylogenomic analyses (Berney et al. 2004; Cavalier-Smith 2004). We calculated a range of trees of varying taxon samplings (e.g. Fig. 2; other data not shown), and although the branching order within the clade was not consistent between them (although many subclades were consis-tently recovered), the microsporidian clade as marked on Fig. 1 was invariably and strongly recovered.

These analyses provide additional evidence for an expanded and strongly supported microsporidian clade, including all of the LB-Microsporidia, (metchnikovellids, the "microsporidian-like" protists discussed above, and almost all of the environmental "crypto/rozellomycotan" diversity indicated in Lara et al. 2010; Jones et al. 2011; Corsaro et al. 2014a,b, 2016; Karpov et al. 2014; Lazarus and James 2015; Tedersoo et al. 2017), but excluding rozellids (=*Rozella*), NAMAKO-37, and NCLC1 (Basal Clone Group 1). We therefore propose that this clade including the large diversity of environmental sequences, are all actually microsporidia, and we refer to them (excluding LB-Microsporidia) here as short-branched Microsporidia (SB-Microsporidia). For the purposes of this study we also exclude metchnikovellids from the definition of SB-Micro-sporidia as sequenced metchnikovellids have significantly longer branches than the SB-Microsporidia shown on Fig. 2. However, it is likely that as more related lineages are discovered a more gradual continuum of branch lengths between SB-Microsporidia such as *Nucleophaga* (Fig. 2), metchnikovellids, and LB-Microsporidia will be revealed.

This "expanded" Microsporidia concept is consistent with the original descriptions of *Paramicrosporidium* (Michel et al. 2000, 2009) as "microsporidian" or "microsporidian-like", and *Mitosporidium*, which exhibits merogony and a coiled polar filament, the latter exclusive to Microsporidia, being "profoundly morphologically similar to Microsporidia" (Haag et al. 2014). This phylogenetically broader circum-scription of Microsporidia is morphologically distinct because all characterized lineages in the microsporidian clade possess the key morphological features of Microspor-idia: spores with multilayered cell walls containing polar fila-ment apparatus (Richards et al. 2017; Vávra and Lukeš 2013), and merogony, whereas *Rozella*, aphelids, and fungi do not possess these characters.

Morphological vs. genomic evolution in Microsporidia

Corresponding morphological and genomic datasets are available for only a very small proportion of lineages repre-senting SB-Microsporidia and metchnikovellids. Until very recently this applied to only *Mitosporidium* (Haag et al. 2014) and *Amphiamblys* (Mikhailov et al. 2017), but the addition of nuclear and mitochondrial genomes of *Parami-crosporidium saccamoebae* (Quandt et al. 2017) have pro-vided several intriguing new perspectives on microsporidian evolution. It is increasingly apparent that, although all SB-microsporidia exhibit microsporidian-defining morphological characters, their genomic evolution appears far more mosaic (Quandt et al. 2017); Table 1. *Rozella* and LB-Micro-sporidia both have horizontally acquired *Rickettsia*-like NTT ATP/ADP transporters, but metchnikovellids, *Mitosporidium* and *Paramicrosporidium* do not (Table 1). LB-Microsporidia lack mitochondrial genomes, which are present in *Mitosporidium* and *Paramicrosporidium*. The mitochondrial genomes of *Mitosporidium* and *Rozella* lack Complex I of the oxidative phosphorylation pathway, are degenerate, and AT-rich, but that of *Paramicrosporidium* includes all genes of that pathway typically found in fungi; in fact, the total gene complement of both *Paramicrosporidium* and *Mitosporidium* have more in common with fungi than with its closest relatives (Haag et al. 2014; Quandt et al. 2017). The very fast rates of sequence evolution (and so long branches in the phylogeny) observed for LB-Microsporidia correlate with extensive cellular and genomic reduction, including strong mutational bias to AT and the loss of some DNA replication and repair genes that, in other eukaryotes, help to promote genome stability (Williams et al. 2016). While genome data are currently very sparse, the slower evolutionary rates observed for the 18S genes of SB-Micro-sporidia suggest that, while these organisms are also para-sites, reductive evolution has not proceeded to the same extreme degree as in the LB clade. Another instance of potentially horizontally acquired genes (the distribution of which may understandably not correlate with phylogeny) are thymidine kinases found in *Rozella* and LB-Microspori-dia, but not *Paramicrosporidium* (Alexander et al. 2016; Quandt et al. 2017). *Paramicrosporidium* possesses a full set of meiosis genes and significant evidence for diploidy, more similar to LB-Microsporidia than is known to be the

Table 1. Molecular and genomic characteristics of Microsporidia and their relatives. Two long-branch microsporidia infecting nematodes (*Nematocida parisii*) and vertebrates (*Encephalitozoon cuniculi*) and *Rozella* (Rozellida) are shown in comparison. Dashes represent unavailable data. ✓/x indicate presence/absence of a character; where uncertain these are shown in parentheses. Note that, following ancestral reduction, some LB-Microsporidia have experienced independent secondary expansions in genome size (e.g. up to 51.3 Mb in *Edhazardia aedis* Desjardins et al. 2015), but the additional material largely comprises noncoding and repetitive sequences, without a concomitant increase in cellular or metabolic complexity. The polar filament was reported as absent in *Nucleophaga amoebae*, although a form of it is present in the very closely related *N. terricolae*. It is possible that it has been lost in *N. amoebae* or was simply not seen in the cells investigated by Corsaro et al. (2014b)

	Flagellate stage	Number of Genome size (Mb)	Gene density genes/kbp	protein coding genes	ATP/DP translocases (HGT)	Mitochondrial genome	Electron transport chain	Polar filament Regular	Atypical	Absent	Spore Exo+endospore	Posterior vacuole
Encephalitozoon cuniculi (LB-Microsporidia)[1,2]	×	2.9	0.81	1,997	✓	×	×	✓			✓	✓
Nematocida parisii ERTm1 (LB-Microsporidia)[3,4]	×	4.1	0.65	2,661	✓	×	×	✓			✓	✓
Amphiamblys sp. WSBS 2006 (metchnikovellid)[5]	×	5–7	0.66	2,529 (nr)	x*	×	×	-	-	-	-	-
Amphiamblys capitellides (metchnikovellid)[6]	×	-	-	-	-	-	-		✓		✓	×
Mitosporidium daphniae (SB-Microsporidia)[7]	×	5.64	0.585	3,300	×	✓	Partial	✓			✓	×
Paramicrosporidium saccamoebae (SB-Microsporidia)[8,9]	×	7.3	-	3,750	×	✓	✓		✓		✓	×
LKM11-01; CM1.1 (SB-Microsporidia)[10]	✓	-	-	-	-	-	-				-	-
Nucleophaga terricolae (SB-Microsporidia)[11,12]	×	-	-	-	-	-	-		✓		✓	×
Nucleophaga amoebae (SB-Microsporidia)[13]	×	-	-	-	-	-	-			✓?†	‡	×
Rozella allomycis (rozellid)[14]	✓	11.9	0.535	6,350	✓	✓	Partial			✓	×	×

nr, nonredundant. 1. Katinka et al. (2001), 2. Karpov et al. (2014), 3. Cuomo et al. (2012), 4. Troemel et al. (2008), 5. Mikhailov et al. (2017), 6. Larsson and Koie (2006), 7. Haag et al. (2014), 8. Corsaro et al. (2014a), 9. Quandt et al. (2017), 10. Jones et al. (2011), 11. Corsaro et al. (2016), 12. Michel et al. (2012), 13. Corsaro et al. (2014b), 14. James et al. (2013), 15. Letcher and Powell (2005), and 16. Ajello (1948).

*Incomplete genome (estimated 90% complete).
†Undetected or potential loss.
‡Two-layered cell wall.
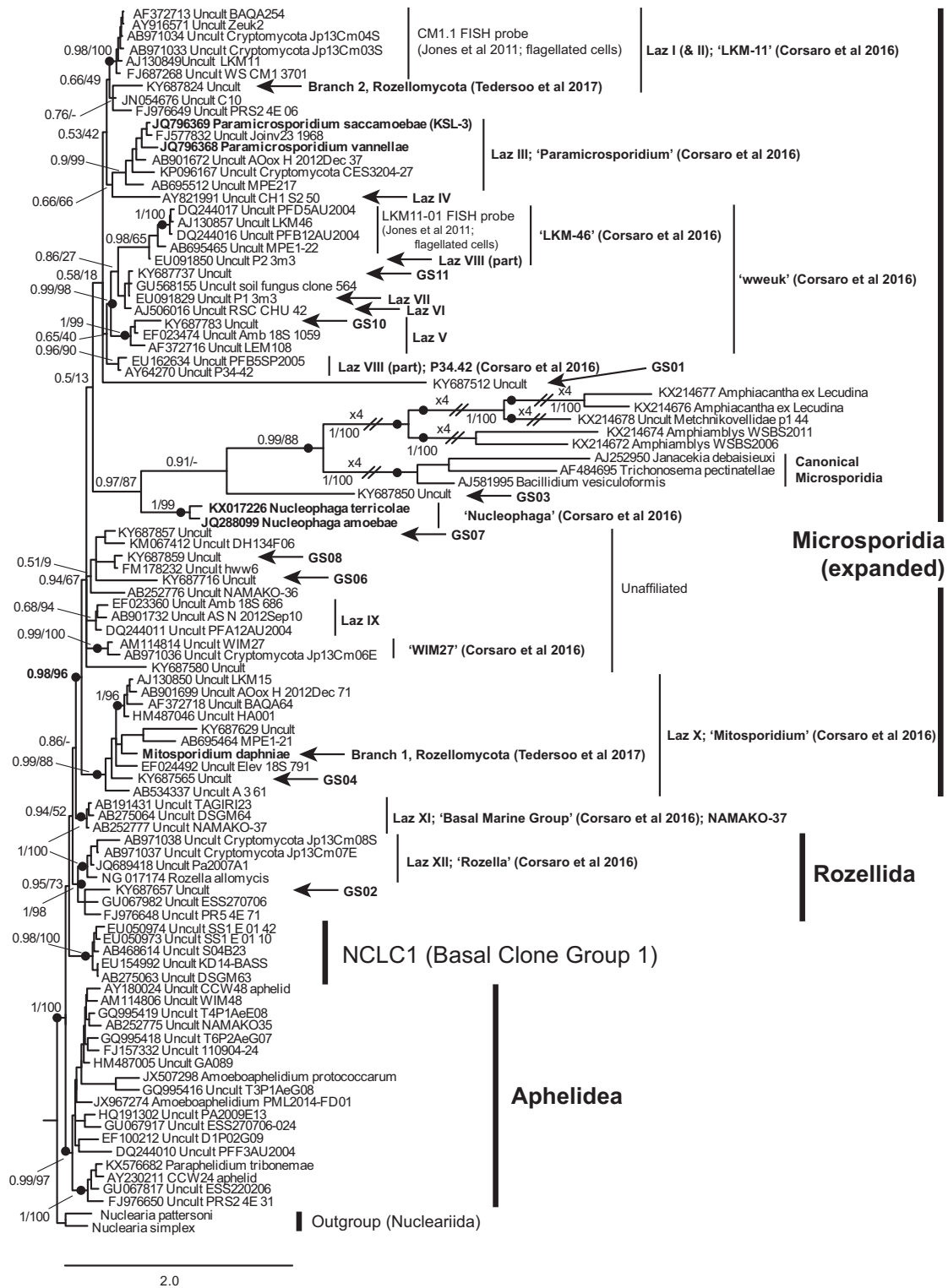
Figure 1 Phylogenetic relationships among canonical, long-branching (LB) Microsporidia, metchnikovellids, Rozellida, and a diversity of related short-branching lineages. LB-Microsporidia form a clade with a diversity of short-branching lineages (SB-Microsporidia) that share key cell biological characters defining the microsporidian clade. Lineages that have been labelled in other studies are labelled: GSxx from Tedersoo et al. (2017), Laz x from Lazarus and James (2015); others as marked. The extent of the expanded Microsporidia is shown by the bracket on the far right. The tree was inferred under the CAT+GTR model in PhyloBayes-MPI, on a final alignment of 1729 sites from the SSU rRNA gene. Black blobs indicate support values of > 0.96 BPP and > 95% ML bootstrap (actual values also shown). Branch lengths are proportional to the expected number of substitutions per site, as denoted by the scale bar.
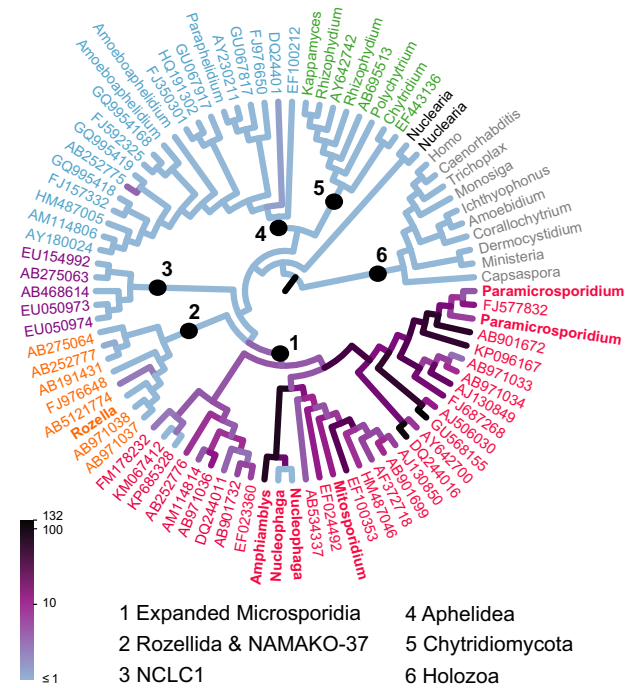
Figure 2 Evolutionary placement of tropical soil microsporidian OTUs on the opisthosporidian reference tree. The branch colours correspond to the distribution of OTU placement: The darker a branch, the more OTUs are placed on it. The OTU frequency scaling is logarithmic due to the large range of placement density. The branch labels for Microsporidia are red (characterized taxa in bold), rozellids orange, NCLC1 purple, aphelids blue, chytrids green, and the holozoan outgroup grey. Nucleariids are the sister clade to Fungi+Opisthosporidia.

earlier-diverging *Mitosporidium* (or the metchnikovellid *Amphiamblys* sp.), but were unable to find any significant hits (E < 1e-10). We were also unable to find significant hits to most of the proteins associated with flagellar structure and function detected in *Paramicrosporidium* by Quandt et al. (2017). These analyses suggest that none of these lineages has a cryptic flagellum that might have been missed by microscopy. However, the branching position of the lineages targeted by FISH in Jones et al. (2011) is unresolved within the microsporidian clade. If these lineages are actually more deeply branching than other characterized SB-Microsporidia then it is possible that the flagellar apparatus in microsporidia was lost early in the diversification of the clade. If this is the case, it would be very interesting to know whether this pre-dated or overlapped with the evolution of the polar filament apparatus, and whether the latter arose at the origin of the microsporidian clade. Alternatively, it is possible that only some SB-Microsporidia lost their flagella apparatus (and associated genes) and that those lineages represented in Jones et al. (2011) are exceptions to the generality suggested by the other characterized lineages in this part of the tree. In that case the use of FISH enabled detection of life-stages and lineages that have so far eluded cell isolation-based methods of investigation. Screening genomic data alone may not provide all such information: even though *Paramicrosporidium* has an obvious polar filament, polar filament proteins (PFPs) known from LB-Microsporidia were not found in the *Paramicrosporidium* or *Mitosporidium* genomes (Haag et al. 2014; Quandt et al. 2017), suggesting rapid evolution of PFPs in the microsporidian clade, earlier forms of these proteins being too dissimilar to their highly derived homologs to be detectable by gene similarity searches. This observation is supported by the low number (589–664; 24–27%) of orthologous genes shared between *Paramicrosporidium* and all sequenced LB-Microsporidia. Haag et al. (2014) found only four orthologs shared between *Mitosporidium* and LB-Microsporidia, but not with other fungi.

The small number of genomic comparisons currently possible between members of Opisthosporidia shows a mosaic evolution, at least partly mediated by horizontal gene transfer, independent gene losses, and perhaps multiple transitions to parasitism. This is a fascinating situation, worthy of intense study, but does not detract from a simple and robust classification as proposed here. As more lineages are detected and characterized, this heterogeneity is set to increase. By adopting a phylogenetically driven, character-based classification structure based on the monophyly of an expanded microsporidian clade, our rationalization provides a clearer set of hypotheses on which to base future studies by pinpointing the origin of the microsporidian radiation, the relation to which genomic and cellular characters can be ascertained.

case for other SB-Microsporidia. In summary, as Quandt et al. (2017) note, "shared gene content is clearly not correlated with evolutionary relationships"; instead, gene content evolution within the microsporidian clade depicted in Fig. 1 appears to be characterized by repeated, lineage-specific gene losses rather than a stepwise trend towards genome reduction in LB-Microsporidia.

Flagella and polar filament evolution in Microsporidia

While rozellids have a lifecycle of alternating zoosporic and nonflagellate trophic stages, flagellate (zoosporic) stages are unknown for any lineages in the expanded microsporidian clade, with two intriguing exceptions. Jones et al. (2011) showed that members of two SB-microsporidia lineages (Laz I and LKM-46, indicated on Fig. 1) have zoosporic stages without chitin in their cell walls. Assuming the cells reconstructed from those FISH experiments represent the branches indicated on Fig. 1 and were not false-positive FISH signals, this shows that some SB-Microsporidia do possess flagella at some stage of their life cycle. In this study, we used BLASTP to search for homologs of the specialized epsilon and delta tubulins, intraflagellar transport system, or flagellar-specific motor molecules in the genomic datasets for the putatively

Classification of Microsporidia and their relatives

Taxonomic circumscriptions of Rozellomycota and, particularly, Cryptomycota, vary significantly (see Berbee et al.

2017; Spatafora et al. 2017; Richards et al. 2017; and examples cited throughout this study). Some authors include aphelids within Cryptomycota (e.g. Letcher et al. 2013), but more normally Crypto/Rozellomycota are used to encompass *Paramicrosporidium*, *Mitosporidium*, *Nucleophaga*, the lineages detected by FISH in Jones et al. (2011), the strongly supported clade containing *Rozella*, and the large diversity of environmental sequences branching around and among these lineages. This is the most frequently used classification, informally described by Letcher et al. (2018) and shown in Jones et al. (2011), Lazarus and James (2015), and (excluding *Mitosporidium*, which is classified as microsporidian) Quandt et al. (2017).

We suggest that the frequently referred to "paraphyly of Rozello/cryptomycota" is both misleading and avoidable. The defining morphological characters of microsporidia are spores with multilayered cell walls containing polar filament apparatus homologs (not necessarily functioning in extrusion). The phylogenetic distribution of these characters is coincident with the clade containing only SB- and LB-Microsporidia, which is recovered by both SSU and multigene phylogenetic analyses. Therefore, all members of that clade are in fact Microsporidia, and there is no need for them to be regarded as belonging to a paraphyletic group. The closest known relatives of microsporidia, again according to SSU (Fig. 1) and multigene phylogenies, are rozellids, which are restricted to a robustly supported clade including the genus *Rozella*. *Rozella* shares some microsporidian-like features with the more distantly related LB-Microsporidia (e.g. horizontally acquired *Rickettsia*-like NTT ATP/ADP transporters, degenerate, and AT-rich mitochondrial genomes lacking Complex I of the oxidative phosphorylation pathway, thymidine kinases), but not with all of their shorter-branch relatives, and are highly morphologically distinct.

What then are Cryptomycota (Jones et al. 2011) and Rozellomycota (Corsaro et al. 2014a)? On the basis that they refer to the same assemblage of lineages (which is usually the case) two names are unnecessary and confusing. The purpose of this study is neither to decide between them nor suggest an alternative, nor even to invalidate them. However, to be consistent with recent usage this label could be applied to an uncharacterized monophyletic group branching somewhere in the opisthosporidian clade, excluding aphelids, the expanded microsporidia, or the clade including *Rozella*. The latter already has order or class status, according to different authorities (Lara et al. 2010; Ruggiero et al. 2015). Figure 1 indicates the diversity revealed by general eukaryote-wide environmental sequencing studies, including those clades coded and labelled by Corsaro et al. (2016), Tedersoo et al. (2017), and the diversity detected by the targeted PCR approach of Lazarus and James (2015). Of this very substantial environmental diversity, only two groups within Opisthosporidia do not branch within the expanded microsporidian, rozellid, or aphelid clades: (1) the NAMAKO-37 clade (which in Corsaro et al. (2016) branches outside of the microsporidian clade before the divergence of *Rozella*, in Jones et al. (2011) is sister to

Rozella, and in Fig. 1 of this study [under the best-fitting CAT+GTR phylogenetic model], and in Lazarus and James (2015) is sister to microsporidia), and (2) NCLC1 (Basal Clone Group I), which in Fig. 1 branches between NAMAKO-37 and aphelids, but whose actual branching position is unresolved (Richards et al. 2015, 2017). On the basis of their phylogenetic position and existing provenance data we suggest that the last two groups are zoosporic parasites of marine microbial eukaryotes.

We infer that the "unaffiliated" group (Fig. 1) is microsporidian (and therefore not cryptomycotan) based on tree topology and bipartition support. These cells could provide key insights into early microsporidian evolution. We hypothesize that they possess mitochondria, and possibly flagellar structures or/and simpler cell extrusion apparatus than in other SB-Microsporidia. The NAMAKO-37 clade cannot be classified until more data are available. There is negligible support for their being microsporidian or belonging to any other recognized group, so the label "Cryptomycota" could be used for this monophyletic lineage, at least for the time being. However, the message of this article remains unchanged whether the whole Rozellida + Microsporidia clade is referred to as Rozello- or Cryptomycota or whether these names fall out of use, what the relationships between aphelids, rozellids, and microsporidia actually are, and where the boundary between fungi and other protists is. Crucially, acknowledging an expanded Microsporidia as an evolutionarily and morphologically coherent unit both unambiguously clarifies the taxonomy/classification of this very interesting clade in the eukaryote tree, and provides a clear framework for future research.

High diversity of SB-Microsporidia in neotropical soils

An important consequence of our analyses is that the taxonomic affiliations of large-scale SSU rRNA amplicon sequencing studies must now be revisited. Most of the sequences annotated as "crypto/rozellomycota" in taxonomically curated databases such as SILVA and PR[2] (Guillou et al. 2013) are likely to be SB-Microsporidia. The rozellid annotation should be restricted to the clade indicated on Fig. 1. This is not purely a matter of classification: Microsporidia and rozellids are each monophyletic and distinct in terms of biology and ecology, but are currently conflated in the major sequence databases. Clarifying the composition of the two clades will enable much more accurate and high-resolution analyses, and interpretation of microsporidian diversity and function.

To demonstrate the particular relevance of this more inclusive definition of microsporidia to the annotation and interpretation of environmental sequencing studies, we reevaluated a recent SSU rRNA high-throughput sequencing study of three Neotropical rainforest soils (Mahé et al. 2017), using the Evolutionary Placement Algorithm as implemented in RAxML (Stamatakis 2014) to place operational taxonomic units (OTUs) previously annotated as Opisthosporidia, crypto/rozellomycota, or unassigned fungi. We similarly analysed an OTU dataset from

European coastal water and sediment samples (BioMarKs: Logares et al. 2014; Massana et al. 2015).

Of the 1,279 candidate tropical soils OTUs, 94% clustered within the microsporidian clade (Fig. 2, 3). Figure 3 shows the branching positions of clades of OTUs relative to the reference sequences. OTUs clustering with the metchnikovellid branch are putatively parasites of gregarines. Others grouping nearer *Nucleophaga*, at the base of the branch leading to metchnikovellids and LB-Microsporidia, may be parasites of amoebae or/and other protists (perhaps including hyperparasites). Mahé et al. (2017) show that both gregarines and Amoebozoa are highly diverse in their Neotropical soil samples. However, until more lineages are found in nature we will be unable to conclusively determine whether microsporidian diversity branching nearer *Paramicrosporidium* or *Mitosporidium* are parasites of protists and arthropods respectively. Although there were far fewer opisthosporidian OTUs in the BioMarKs data, the majority of these also branched within the microsporidian clade (Fig. 3). In both datasets a small proportion of OTUs was placed in the rozellid clade (two from soil), and the NAMAKO-37 clade (four from BioMarKs), consistent with the apparently limited diversity of these clades compared to Microsporidia.

LB-Microsporidia are predominantly parasites of animals, but *Nucleophaga* and *Paramicrosporidium* are parasites of protists (amoebozoan hosts are only known so far, which are relatively scarce in marine habitats). Metchnikovellids are parasites of gregarines (protistan gut parasites of a wide range of invertebrates; Desportes and Schrével 2013). *Mitosporidium* is a parasite of *Daphnia* (Haag et al. 2014), and perhaps other lineages in the *Mitosporidium* clade are also parasites of arthropods. SB-Microsporidia therefore appear to occupy a broad and little understood set of niches. The phylogenetic distribution of characters typically associated with LB-Microsporidia may be determined at least as much by host-specific adaptation as phylogenetic relatedness. For instance, shorter/reduced polar filaments may be more characteristic of Microsporidia infecting protists than invertebrates. On the other hand, if at least in some cases cell host invasion is mediated by host phagocytosis, the length/complexity of the polar filament may not always be directly related to the physical barriers it must cross to invade the host (Franzen 2004).

CONCLUSIONS

We analysed the phylogenetic placement and environmental diversity of microsporidian-like lineages that group near the base of the fungal radiation. These lineages form a monophyletic group including canonical Microsporidia and metchnikovellids, but excluding a strongly supported rozellid clade. This topology is concordant with the phylogenetic distribution of defining microsporidian cell characters, but not shared gene content across Opisthosporidia and fungi. The genetic diversity of Microsporidia is far higher than previously realized, and includes the SB-microsporidian taxa *Mitosporidium*, *Paramicrosporidium*, and *Nucelophaga*, and
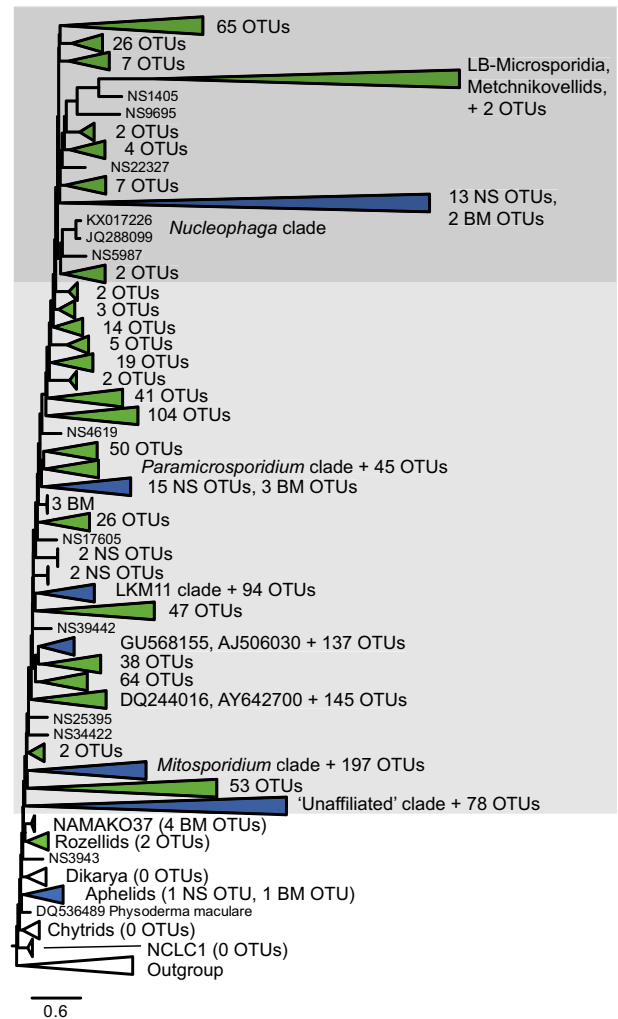
Figure 3 Maximum Likelihood phylogeny of SSU rRNA gene V4 region OTUs from tropical soil and European marine (Biomarks) samples. The OTUs are aligned to reference sequence dataset of long- and short-branch Microsporidia, metchnikovellids, rozellids, other clades of environmental sequences, and representative aphelids and fungi with a holozoan outgroup. Light and dark grey shading indicates extent of expanded Microsporidia; dark grey box encloses all members of the subclade including long-branch Microsporidia. Green triangles (collapsed clades) represent OTUs from Neotropical soil samples only, 'BM' from Biomarks only, and blue from both.

many uncharacterized environmental sequence types. The concept of Rozellomycota/Cryptomycota requires revision to avoid encompassing lineages that are actually microsporidian. Our analyses suggest hypotheses for investigations into the relative timings of acquisition of the polar filament apparatus and loss of flagella, key microsporidian characteristics. We show that this revised classification has major implications for our understanding of microsporidian diversity as inferred from environmental sequencing surveys. The large diversity and abundance of SB-Microsporidia offer unique opportunities to study the evolution of the highly specialized cells and genomes of canonical

Microsporidia, and in particular their propensity for direct parasitism of animals.

ACKNOWLEDGMENTS

LITERATURE CITED

Ajello, L. 1948. A cytological and nutritional study of polychytrium aggregatum. 1. CytologY. *Am. J. Bot.*, 35:1–12.

Alexander, W. G., Wisecaver, J. H., Rokas, A. & Hittinger, C. T. 2016. Horizontally acquired genes in early-diverging pathogenic fungi enable the use of host nucleosides and nucleotides. *Proc. Natl Acad. Sci. USA*, 113:4116–4121.

Berbee, M. L., James, T. Y. & Strullu-Derrien, C. 2017. Early diverging fungi: diversity and impact at the dawn of terrestrial life. *Annu. Rev. Microbiol.*, 71:41–60.

Berger, S. A., Krompass, D. & Stamatakis, A. 2011. Performance, accuracy, and web server for evolutionary placement of short sequence reads under Maximum Likelihood. *Syst. Biol.*, 60:291–302.

Berger, S. A. & Stamatakis, A. 2011. Aligning short reads to reference alignments and trees. *Bioinformatics*, 27:2068–2075.

Berney, C., Fahrni, J. & Pawlowski, J. 2004. How many novel eukaryotic 'kingdoms'? Pitfalls and limitations of environmental DNA surveys. *BMC Biol.*, 2:13.

Beznoussenko, G. V., Dolgikh, V. V., Seliverstova, E. V., Semenov, P. B., Tokarev, Y. S., Trucco, A., Micaroni, M., Di Giandomenico, D., Auinger, P., Senderskiy, I. V., Skarlato, S. O., Snigirevskaya, E. S., Komissarchik, Y. Y., Pavelka, M., De Matteis, M. A., Luini, A., Sokolova, Y. Y. & Mironov, A. A. 2007. Analogs of the Golgi complex in microsporidia: structure and avesicular mechanisms of function. *J. Cell Sci.*, 120:1288–1298.

Cavalier-Smith, T. 2004. Only six kingdoms of life. *Proc. Biol. Sci.*, 271:1251–1562.

Corsaro, D., Michel, R., Walochnik, J., Venditti, D., Müller, K.-D., Hauröder, B. & Wylezich, C. 2016. Molecular identification of *Nucleophaga terricolae* sp. nov. (Rozellomycota), and new insights on the origin of the Microsporidia. *Parasitol. Res.*, 115:3003–3011.

Corsaro, D., Walochnik, J., Venditti, D., Müller, K.-D., Hauröder, B. & Michel, R. 2014b. Rediscovery of *Nucleophaga amoebae*, a novel member of the Rozellomycota. *Parasitol. Res.*, 113:4491–4498.

Corsaro, D., Walochnik, J., Venditti, D., Steinmann, J., Müller, K.-D. & Michel, R. 2014a. Microsporidia-like parasites of amoebae belong to the early fungal lineage Rozellomycota. *Parasitol. Res.*, 113:1909–1918.

Cuomo, CA, Desjardins, CA, Bakowski, MA, Goldberg, J, Ma, AT, Becnel, JJ, Didier, ES, Fan, L, Heiman, DI, Levin, JZ & Young, S. 2012. Microsporidian genome analysis reveals evolutionary strategies for obligate intracellular growth. *Genome Res.*, 22:2478–2488.

Dean, P., Hirt, R. P. & Embley, T. M. 2016. Microsporidia: why make nucleotides if you can steal them? *PLoS Pathogens*, 12: e1005870.

Desjardins, C. A., Sanscrainte, N. D., Goldberg, J. M., Heiman, D., Young, S., Zeng, Q., Madhani, H. D., Becnel, J. J. & Cuomo, C. A. 2015. Contrasting host–pathogen interactions and genome evolution in two generalist and specialist microsporidian pathogens of mosquitoes. *Nat. Comm.*, 6:7121.

Desportes, I. & Schrével, J. 2013. Treatise on Zoology—Anatomy, Taxonomy, Biology: The Gregarines (2 vols). Brill, The Netherlands.

Fokin, S. I. 2012. Frequency and biodiversity of symbionts in representatives of the main classes of Ciliophora. *Eur. J. Protistol.*, 48:138–148.

Fokin, S. I., Di Giuseppe, G., Erra, F. & Dini, F. 2008. *Euplotespora binucleata* n. gen., n. sp. (Protozoa: Microsporidia), a parasite infecting the hypotrichous ciliate *Euplotes woodruffi*, with observations on microsporidian infections in ciliophora. *J. Euk. Microbiol.*, 55:214–228.

Franzen, C. 2004. Microsporidia: how can they invade other cells? *Trends Parasitol.*, 20:275–279.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J. & Del Campo, J. 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucl. Acids Res.*, 41:D597–D604.

Haag, K. L., James, T. Y., Pombert, J.-F., Larsson, R., Schaer, T. M. M., Refardt, D. & Ebert, D. 2014. Evolution of a morphological novelty occurred before genome compaction in a lineage of extreme parasites. *Proc. Natl Acad. Sci. USA*, 111:15480–15485.

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

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

Karpov, S. A., Mamkaeva, M. A., Aleoshin, V. V., Nassonova, E., Lilje, O. & Gleason, F. H. 2014. Morphology, phylogeny, and ecology of the aphelids (Aphelidea, Opisthokonta) and proposal for the new superphylum Opisthosporidia. *Front. Microbiol.*, 5:112.

Katinka, MD, Duprat, S, Cornillot, E, Méténier, G, Thomarat, F, Prensier, G, Barbe, V, Peyretaillade, E, Brottier, P, Wincker, P & Delbac, F. 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature*, 414:450–453.

Katoh, K. & Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, 30:772–780.

Lara, E., Moreira, D. & Lopez-García, P. 2010. The environmental clade LKM11 and *Rozella* form the deepest branching clade of fungi. *Protist*, 161:116–121.

Larsson, J. I. & Koie, M. 2006. The ultrastructure and reproduction of *Amphiamblys capitellides* (Microspora, Metchnikovellidae), a parasite of the gregarine *Ancora sagittata* (Apicomplexa, Lecudinidae), with redescription of the species and comments on the taxonomy. *Eur. J. Protistol.*, 42:233–248.

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.

Lazarus, K. L. & James, T. Y. 2015. Surveying the biodiversity of the Cryptomycota using a targeted PCR approach. *Fung. Ecol.*, 14:62–70.

Letcher, P. M., Lopez, S., Schmieder, R., Lee, P. A., Behnke, C., Powell, M. L. & 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:e56232.

Letcher, P. M., Longcore, J. E., James, T. Y., Leite, D. S., Simmons, D. R. & Powell, M. J. 2018. Morphology, ultrastructure, and molecular phylogeny of *Rozella multimorpha*, a new species in Cryptomycota. *J. Euk. Microbiol.* 65:180–190.

Letcher, P. A. & Powell, M. J. 2005. *Kappamyces*, a new genus in the Chytridiales (Chytridiomycota). *Nova Hedwigia*, 80:115–133.

Liu, Y., Steenkamp, E. T., Brinkmann, H., Forget, L., Philippe, H. & Lang, B. F. 2009. Phylogenomic analyses predict sistergroup relationship of nucleariids and fungi and paraphyly of zygomycetes with significant support. *BMC Evol. Biol.*, 9:272.

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

Mahé, F., de Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., Mayor, J., Bunge, J., Sernaker, S., Siemensmeyer, T., Trautmann, I. 2017. Soil protists in three Neotropical rainforests are hyperdiverse and dominated by parasites. *Nature Ecol. Evol.*, 1:0091.

Mahé, F., Rognes, T., Quince, C., de Vargas, C. & Dunthorn, M. 2015. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ*, 3:e1420.

Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., Chambouvet, A., Christen, R., Claverie, J. M., Decelle, J. & Dolan, J. R.2015. Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ. Microbiol.*, 17:4035–4049.

Michel, R., Müller, K.-D. & Hauröder, B. 2009. A novel microsporidian endoparasite replicating within the nucleus of *Saccamoeba limax* isolated from a pond. *Endocytobiosis Cell Res.*, 19:120–126.

Michel, R., Müller, K-D., Schmid, E.N., Thee-garten, N.N., Hauröder, B. & Corsaro, D 2012. Isolation of *Thecamoeba terricola* from bark of *Platanus occidentalis* harbouring spore-forming eukaryotic endoparasites with intranuclear development. *Endocytobios. Cell Res.*, 22:37–42.

Michel, R., Schmid, E. N., Böker, T., Hager, D. G., Müller, K.-D., Hoffmann, R. & Seitz, H. M. 2000. *Vannella* sp. harboring Microsporidia-like organisms isolated from the contact lens and inflamed eye of a female keratitis patient. *Parasitol. Res.*, 86:514–520.

Mikhailov, K. V., Simdyanov, T. G. & Aleoshin, V. V. 2017. Genomic survey of a hyperparasitic microsporidian *Amphiamblys* sp. (Metchnikovellidae). *Genome Biol. Evol.*, 9:454–467.

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.

Quandt, C. A., Beaudet, D., Corsaro, D., Walochnik, J., Michel, R., Corradi, N. & James, T. Y. 2017. The genome of an intranuclear parasite, Paramicrosporidium saccamoebae, reveals alternative adaptations to obligate intracellular parasitism. *Elife*, 6: e29594.

Richards, T. A., Leonard, G., Mahé, F., del Campo, J., Romac, S., Jones, M. D. M., Maguire, F., Dunthorn, M., De Vargas, C., Massana, R. & Chambouvet, A. 2015. Molecular diversity and distribution of marine fungi across 130 European environmental samples. *Proc. R. Soc. B*, 282:20152243.

Richards, T. A., Leonard, G. & Wideman, J. G. 2017. What defines the "kingdom" fungi?. *Microbiol. Spectr.*, 5. https://doi.org/10.1128/microbiolspec.FUNK-0044-2017

Ruggiero, M. A., Gordon, D. P., Orrell, T. M., Bailly, N., Bourgoin, T., Brisca, R. C., Cavalier-Smith, T., Guiry, M. D. & Kirk, P. M. 2015. A higher level classification of all living organisms. *PLoS ONE*, 10:e0119248.

Spatafora, J. W., Aime, M. C., Grigoriev, I. V., Martin, F., Stajich, J. E. & Blackwell, M. 2017. The fungal tree of life: from molecular systematics to genome-scale phylogenies. *Microbiol. Spectr.*, 5. https://doi.org/10.1128/microbiolspec.FUNK-0053-2016

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30:1312–1313.

Stentiford, G. D., Ramilo, A., Abollo, E., Kerr, R., Bateman, K. S., Feist, S. W., Bass, D. & Villalba, A. 2017. *Hyperspora aquatica* n.gn., n.sp. (Microsporidia), hyperparasitic in *Marteilia cochillia* (Paramyxida), is closely related to crustacean-infecting microspordian taxa. *Parasitology*, 144:186–199.

Tedersoo, L., Bahram, M., Puusepp, R., Nilsson, R. H. & James, T. Y. 2017. Novel soil-inhabiting clades fill gaps in the fungal tree of life. *Microbiome*, 5:42.

Torruella, G., de Mendoza, A., Grau-Bové, X., Antó, M., Chaplin, M. A., del Campo, J., Eme, L., Pérez-Cordón, G., Whipps, C. M., Nichols, K. M., Paley, R., Roger, A. J., Sitjà-Bobadilla, A., Donachie, S. & Ruiz-Trillo, I. 2015. Phylogenomics reveals convergent evolution of lifestyles in close relatives of animals and fungi. *Curr. Biol.*, 25:2404–2410.

Troemel, E. R., Félix, M. A., Whiteman, NK, Barrière, A & Ausubel, FM 2008. Microsporidia are natural intracellular parasites of the nematode *Caenorhabditis elegans*. *PLoS Biol.*, 6:2736–2752.

Vávra, J. & Larsson, J. I. R. 1999. Structure of the microsporidia. *In*: Wittner, M. & Weiss, L. M. (ed.), The Microsporidia and Microsporidiosis. ASM Press, Washington, D.C. p. 7–84.

Vávra, J. & Lukeš, J. 2013. Microsporidia and 'the art of living together'. *Adv. Parasitol.*, 82:253–319.

Williams, T. A., Nakjang, S., Campbell, S. E., Freeman, M. A., Eydal, M., Moore, K., Hirt, R. P., Embley, T. M. & Williams, B. A. 2016. A recent whole-genome duplication divides populations of a globally distributed microsporidian. *Mol. Biol. Evol.*, 33:2002–2015.