European Journal of Protistology 96 (2024) 126119

Contents lists available at ScienceDirect

European Journal of Protistology

journal homepage: www.elsevier.com/locate/ejop

Research Paper

O short-branch Microsporidia, where art thou? Identifying diversity hotspots for future sampling

Megan Gross [a,b,*], Ľubomír Rajter [c], Frédéric Mahé [d,e], David Bass [f,g,h], Cédric Berney [i,j], Nicolas Henry [i,k], Colomban de Vargas [i,j], Micah Dunthorn [a]

[a] Natural History Museum, University of Oslo, 0562 Oslo, Norway
[b] Department of Ecology, University of Kaiserslautern-Landau RPTU, 67663 Kaiserslautern, Germany
[c] Institute for Zoology, University of Cologne, 50923 Cologne, Germany
[d] CIRAD, UMR PHIM, 34398 Montpellier, France
[e] PHIM, Univ Montpellier, CIRAD, INRAE, Institut Agro, IRD, 34398 Montpellier, France
[f] Cefas, International Centre for Aquatic Animal Health, Weymouth, Dorset DT4 8UB, United Kingdom
[g] Sustainable Aquaculture Futures, Biosciences, College of Life and Environmental Sciences, University of Exeter, Stocker Road, Exeter EX4 4QD, United Kingdom
[h] Department of Life Sciences, The Natural History Museum, London SW7 5BD, United Kingdom
[i] CNRS, Sorbonne Université, FR2424, ABiMS, Station Biologique de Roscoff, 29680 Roscoff, France
[j] Sorbonne Université, CNRS, Station Biologique de Roscoff, UMR7144, ECOMAP, 29680 Roscoff, France
[k] Research Federation for the Study of Global Ocean Systems Ecology and Evolution, FR2022/Tara GOSEE, 75016 Paris, France

ARTICLE INFO

Keywords:
Metabarcoding
Parasites
Protists
SSU rRNA
V4 region

ABSTRACT

Short-branch Microsporidia were previously shown to form a basal grade within the expanded Microsporidia clade and to branch near the classical, long-branch Microsporidia. Although they share simpler versions of some morphological characteristics, they do not show accelerated evolutionary rates, making them ideal candidates to study the evolutionary trajectories that have led to long-branch microsporidian unique characteristics. However, most sequences assigned to the short-branch Microsporidia are undescribed, novel environmental lineages for which the identification requires knowledge of where they can be found. To direct future isolation, we used the EukBank database of the global UniEuk initiative that contains the majority of the publicly available environmental V4 SSU rRNA gene sequences of protists. The curated OTU table and corresponding metadata were used to evaluate the occurrence of short-branch Microsporidia across freshwater, hypersaline, marine benthic, marine pelagic, and terrestrial environments. Presence–absence analyses infer that short-branch Microsporidia are most abundant in freshwater and terrestrial environments, and alpha- and beta-diversity measures indicate that focusing our sampling effort on these two environments would cover a large part of their overall diversity. These results can be used to coordinate future isolation and sampling campaigns to better understand the enigmatic evolution of microsporidians' unique characteristics.

1. Introduction

Parasites comprise a large fraction of Earth's biodiversity (Dobson et al., 2008; Loker and Hofkin, 2022; Mahé et al., 2017) and their transition from free-living forms to a parasitic lifestyle occurred in numerous independent events. This convergent evolution among phylogenetically unrelated taxa resulted in similar strategies of host-invasion, transmission between, and survival within their hosts (Poulin, 2011; Poulin and Randhawa, 2015). One clade of obligate intracellular parasites is the Microsporidia (Keeling and Fast, 2002; Vávra and Lukeš, 2013; Weiss and Becnel, 2014). While the diversity of Microsporidia and their infection mechanisms are somewhat understood, it is still not clear how their unique characteristics, and thus how this clade itself, evolved (Keeling and Fast, 2002).

Microsporidians infect a wide range of animals and some protists (Adl et al., 2019; Becnel and Andreadis, 1999; Foissner and Foissner, 1995; Vávra and Lukeš, 2013; Vossbrinck and Debrunner-Vossbrinck, 2005). Thus far, more than 1300 species are described (Franzen, 2008; Weiss and Becnel, 2014), many of them being harmful and emergent pathogens of socio-economic importance (Fries, 1993; Keeling

* Corresponding author at: Department of Ecology, University of Kaiserslautern-Landau RPTU, 67663 Kaiserslautern, Germany.
E-mail address: gross.megan@rptu.de (M. Gross).

https://doi.org/10.1016/j.ejop.2024.126119
Received 26 June 2024; Received in revised form 23 September 2024; Accepted 24 September 2024
Available online 25 September 2024

and Fast, 2002; Kent et al., 1989; Stentiford et al., 2016; Weber et al., 1994). While microsporidians display a wide complexity in their infection pathways using a unique polar filament for host invasion, they otherwise have had a rather reductive evolution in their cellular organization (Dean et al., 2018; Katinka et al., 2001; Keeling and Corradi, 2011). The loss of many DNA repair enzymes (Gill and Fast, 2007) and the presence of many fast-evolving genes (Thomarat et al., 2004), may have enhanced the process of genome reduction, thereby playing a major role in the evolution of Microsporidia (Cuomo et al., 2012).

Accelerated evolutionary rates caused long-branch attraction artefacts and initial phylogenetic inferences placed the microsporidians as an early branching eukaryote clade (Keeling and McFadden, 1998; Vossbrinck et al., 1987). Subsequent phylogenetic analyses that took into account variable base substitutions and included numerous loci changed our understanding of their placement within the eukaryotic tree of life (Corradi and Keeling, 2009; Park and Poulin, 2021). It is now widely accepted that microsporidians are closely related to the Fungi (Brown and Doolittle, 1999; Corradi and Keeling, 2009; Gill and Fast, 2006; Hirt et al., 1999; Keeling, 2014, 2003; Strassert and Monaghan, 2022; Voigt et al., 2021). However, although many studies made important contributions to unravel their diversity and phylogeny, as well as shedding light on the unique characteristics that make the microsporidians such pivotal parasites, it still remains unclear how these characteristics evolved.

Bass et al. (2018) expanded our understanding of the relationship between microsporidians and many different microsporidian-like protists that were previously assumed to group together with rozellids (parasites of Chytridiomycetes, Blastocladiomycetes, and Oomycetes) within the 'cryptomycotan' clade. They used small subunit rRNA (SSU rRNA) from sequenced isolates and environmental metabarcoding data to infer that these microsporidian-like lineages form a basal grade of parasites that group with metchnikovellids and classical Microsporidia, thereby demonstrating that the phylogenetic scope of the microsporidians is greater than originally assumed. This expanded microsporidian clade includes all classical Microsporidia, Metchnikovellida, and Chytridiopsida, which they collectively named 'long-branch Microsporidia' (referring to the relatively long branches within phylogenetic trees), and the basal grade with less divergent SSU rRNA gene sequences, which they named 'short-branch Microsporidia' (Bass et al., 2018; Corsaro et al., 2019).

The short-branch Microsporidia include the partially-characterized lineages *Mitosporidium*, *Morellospora*, *Nucleophaga*, and *Paramicrosporidium*, which are variously known to be parasites of *Daphnia* (Haag et al., 2014) and different amoebae (Corsaro et al., 2014a, 2014b; Michel et al., 2000, 2009a, 2009b). Some of them share morphological traits with the long-branch Microsporidia, such as simpler versions of the polar filament (with a similar function) but differ greatly in other characteristics (Bass et al., 2018; Corsaro et al., 2014a, 2014b; Haag et al., 2014). For example, they do not show rapid rates of evolution and have less reduced genomes that are more similar to those of *Rozella* and canonical Fungi (Haag et al., 2014; Quandt et al., 2017). The short-branch Microsporidia also include many uncharacterized environmental lineages that may be parasitic as well (Doliwa et al., 2021), such as LMK11. By further investigating the short-branch Microsporidia, we can better understand the intriguing evolution that occurred among the classical long-branch Microsporidia, which ultimately resulted in trait reductions and increased complexity of their extrusion apparatus. However, as the short-branch Microsporidia are undersampled and understudied, we do not yet know where to focus attempts in order to isolate more of them, to further investigate the partially characterized lineages, and to newly characterize the environmental lineages.

In this study, we used data from Berney et al. (2023) that collected and analyzed all available environmental metabarcoding data of the protistan hypervariable V4 region of the nuclear SSU rRNA gene. After extracting short-branch Microsporidia operational taxonomic units (OTUs), we asked in which environments short-branch Microsporidia are present and if there are differences in their abundance and diversity throughout these environments. Our results pave the road for future studies in which we aim to better understand the evolution of the Microsporidia, by directing where we should go to isolate and sequence more short-branch Microsporidia.

2. Material and methods

2.1. Dataset

Microsporidian sequences and their corresponding metadata used in this study came from a beta version of the EukBank database (Berney et al., 2023) as part of the global UniEuk initiative (Berney et al., 2017). Briefly, to develop the EukBank database, metadata from each bioproject was downloaded from public repositories such as the European Nucleotide Archive (EMBL/EBI-ENA) and the Sequence Read Archive (NCBI). Environmental V4 SSU rRNA raw sequences were downloaded from the EMBL/EBI-ENA EukBank umbrella project, clustered with Swarm ver. 3 using default parameters with the fastidious option on (Mahé et al., 2022), checked for chimeras using VSEARCH ver. 2.21.0 (Rognes et al., 2016), merged with MUMU ver. 1.0.1 (Mahé, 2021), and taxonomically assigned using the stampa pipeline (https://github.com/frederic-mahe/stampa/) and EukRibo ver. 1 reference database (Berney et al., 2023). From this, an OTU (operational taxonomic unit) occurrence table was generated. OTU, taxonomy, and metadata tables can be found in the supplements (Suppl. Files 1–3). All further steps were conducted using R ver. 4.3.1 (R Development Core Team, 2012). The codes are available in HTML format (Suppl. File 4).

2.2. Modification of OTU table

Modifications of the dataset were performed using the R package tidyverse ver. 1.3.1 (Wickham et al., 2019). The OTU, taxonomy, and metadata tables were matched against each other to make sure that OTUs and samples were identical. To verify that the dataset only contained OTUs annotated to the short-branch Microsporidia, the taxonomy table was searched for the keywords 'canonical' and 'classical' Microsporidia. All OTUs to which this applied were removed. To correctly infer the environment in which short-branch Microsporidia can be found, every sample was assigned to either of the five different environments, i.e., 'freshwater', 'hypersaline', 'marine benthic', 'marine pelagic' or 'terrestrial' using the information provided in the metadata table. However, for some samples, there was not enough information provided in the metadata table itself to correctly assign it to one of the environments. To address this knowledge gap, we searched inside the EMBL-EBI Ontology Code for equivalent information and added it to the metadata table.

2.3. Statistical analysis

Analyses and visualizations were performed with the R packages tidyverse ver. 2.0.0 (Wickham et al., 2019), ggpubr ver. 0.6.0 (Kassambara, 2020), rstatix ver. 0.7.2 (Kassambara, 2021), maps ver. 3.4.2 (Brownrigg, 2021), phyloseq ver. 1.46.0 (McMurdie and Holmes, 2013), metagMisc ver. 0.5.0 (Mikryukov, 2021), and venn ver. 1.12 (Dusa, 2021). All analyses were performed on the unrarefied dataset to ensure that samples were not lost due to poor sample coverage. However, as the sequencing depth varies between different studies due to different sampling efforts and procedures, all analyses were tested on the rarefied dataset as well, to make sure that this did not influence the outcome. The quality of the taxonomic annotation was examined based on the number of OTUs per similarity value to the reference database and was visualized as bar charts. Sampling maps, based on latitude/longitude information, were created to compare the total amount of samples against samples for which short-branch microsporidian OTUs were found. Samples were color-coded based on the five different

environments. To compare the number of unique and shared OTUs present in the different environments, Venn diagrams were created. Additionally, OTUs occurring in all investigated environments were aligned with MUSCLE ver. 3.8.425 (Edgar, 2004) to check for sequence similarity. For community analysis, OTU, metadata, and taxonomy tables were transformed and then merged into a phyloseq object. Observed and estimated richness (Chao1) was compared between the different environments and results were visualized as boxplots. Wilcoxon rank sum tests were used to assess the pairwise differences in richness among the environments (we recognized the dataset was not normally distributed because many samples were low abundant, but we retained all samples to get a complete view of the distributions across the environments). To investigate similarity patterns between the five different environments, the OTU table was transformed into presence/

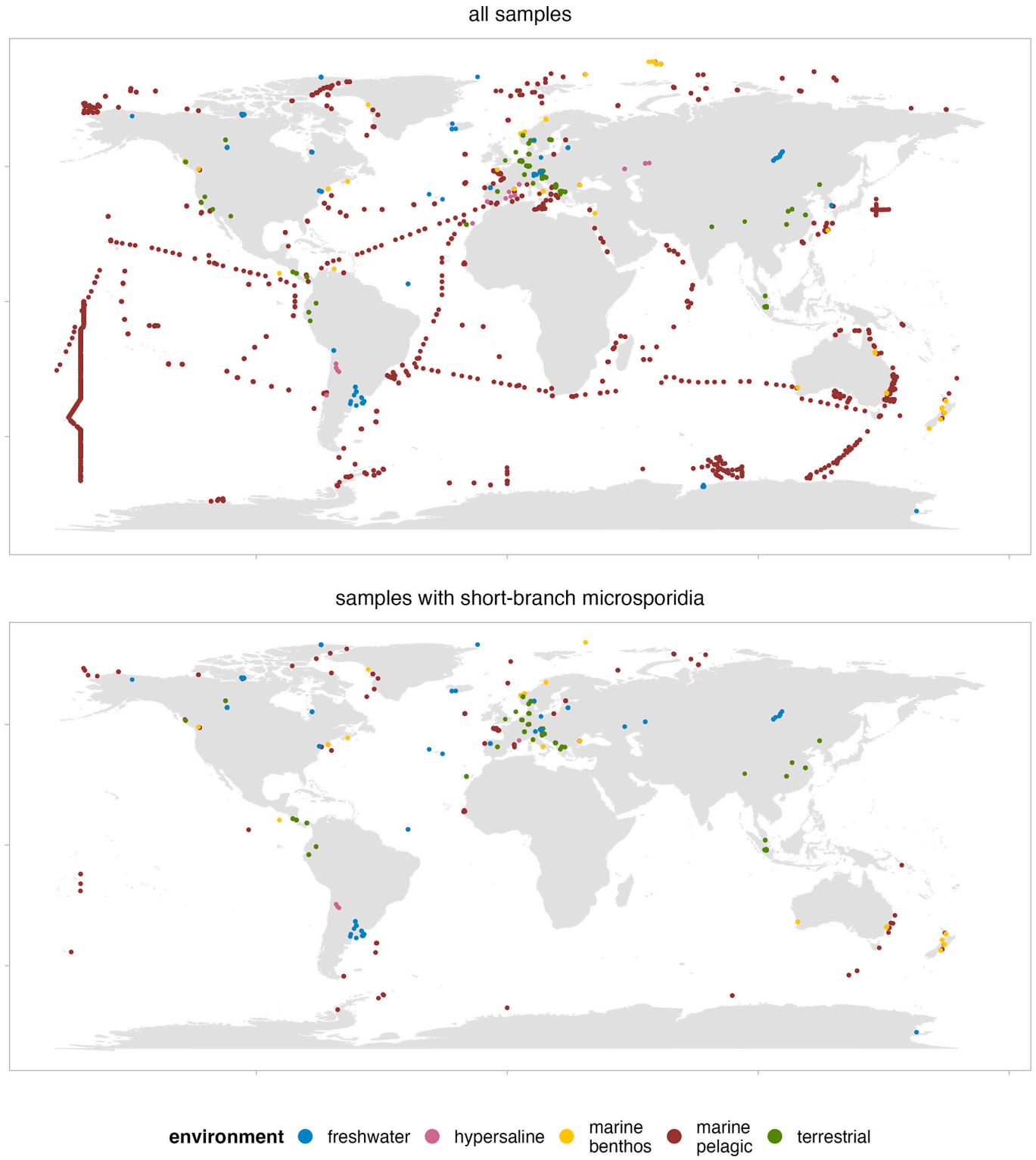

Fig. 1. Map of all sample locations (upper panel) and map of samples that include short-branch Microsporidia (lower panel). Colors represent the environment in which the samples were collected.

absence data, and a non-metric multidimensional scaling (NMDS), using the Jaccard metric, was performed.

3. Results

3.1. Sampling and sequencing

Of the 13,045 samples taken from across the globe, the majority came from marine pelagic environments (62.89 %), followed by terrestrial (19.36 %), marine benthos (13.58 %), freshwater (3.92 %), and hypersaline (0.25 %) environments (Fig. 1). The majority of the terrestrial samples were taken in areas in the northern hemisphere, especially within Europe. There were also many countries for which no samples targeting the V4 region were available including the whole African continent, Australia, and also parts of South America such as Brazil. From the initial dataset, 1,403,019,176 environmental V4 sequencing reads clustered into 460,147 OTUs that were taxonomically assigned to the protists in general. After filtering out all non-short-branch microsporidian sequences, 6,796,304 (0.48 %) reads remained that were clustered into 1,741 short-branch microsporidian OTUs. From these short-branch Microsporidia data, 94.9 % of the reads and 69.8 % of the OTUs were more than 90 % similar to references in the EukRibo (Berney et al., 2022) reference database (Fig. 2).

3.2. Distributions across environments

Of the 13,045 samples, we found short-branch Microsporidia OTUs within 3,063 (23.27 %) of them (Fig. 1). A majority of these samples (1,794) and OTUs (1,279) were found in terrestrial environments (Fig. 3). Although there were more samples from marine environments, freshwater samples accounted for 879 OTUs compared to 578 for marine pelagic and 563 OTUs for marine benthos. A similar picture was found when searching for unique and shared OTUs across the different environments. Terrestrial environments had the highest number of unique OTUs (604) followed by freshwater (192). There were 155 OTUs shared between freshwater and terrestrial environments, 216 that occurred within every environment with the exception of hypersaline, and two OTUs that occurred in all five environments. These two OTUs found in all five environments had a high similarity to accessions in the taxonomic reference database (>98 %). Both had *Paramicrosporidium* clone LKM-46 (van Hannen et al., 1999) as their closest reference and they were 98.26 % similar to each other.

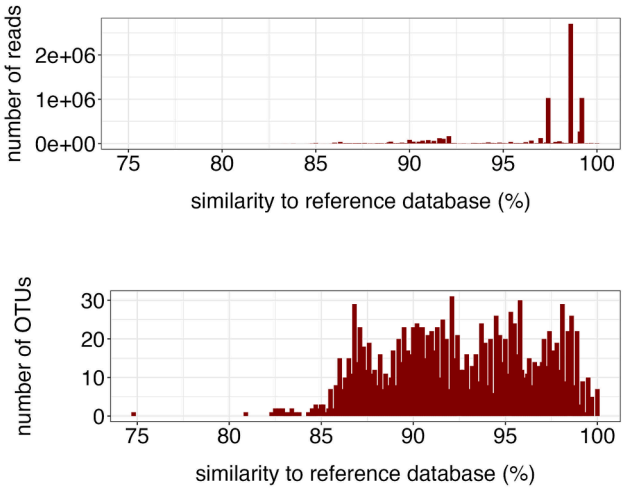

Fig. 2. Similarity of short-branch Microsporidia to the taxonomic reference database for number of reads (**A**) and number of OTUs (**B**).

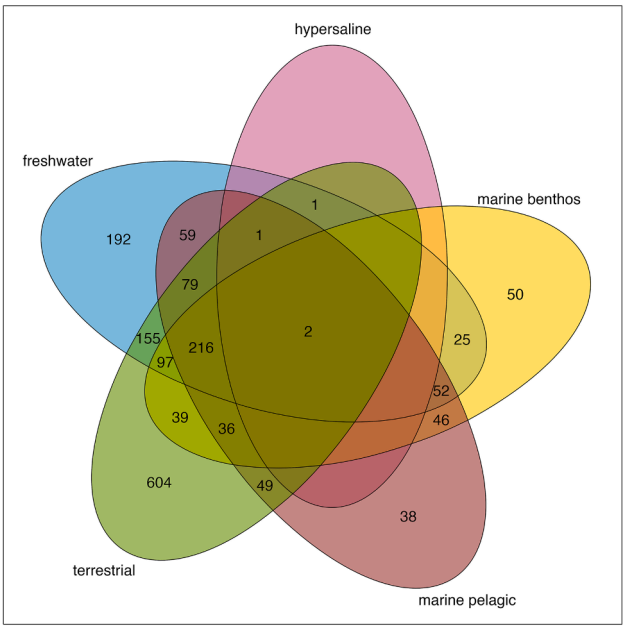

Fig. 3. Venn-diagram showing the number of unique and shared OTUs across the five investigated environments. Sections in which no number appears indicate zero OTUs.

3.3. Community richness and heterogeneity

The observed richness of short-branch Microsporidia OTUs showed that the average number of OTUs per sample was highest for terrestrial, followed by freshwater environments (Fig. 4). Estimated richness using Chao1 predicted slightly higher diversity values for the different environments. Terrestrial environments had the highest diversity (Chao1: 15 ± 18.82 SE) followed by freshwater (Chao1: 11 ± 44.13 SE). Marine benthos, marine pelagic, and hypersaline environments showed very low average diversity estimates (respectively 3 ± 30.39 SE, 2 ± 15.54 SE, and 1 ± 1.50 SE) and differed significantly from freshwater and terrestrial environments (Suppl. Table 1). Although most samples showed an observed richness of <50 OTUs per sample, 11 freshwater and six marine benthos samples were found with >100 OTUs per sample. Among these, three out of 11 freshwater sample sites were from lake Pollevann in Norway (unpublished study), seven came from lake Sanabria in Spain (unpublished study), and one sample originated from lake Augstsee in Austria (unpublished study). All of the six marine benthos samples were from Norwegian fjords (unpublished study). In contrast to the differences in observed and estimated OTU richnesses, there were no clear differences in the communities of short-branch Microsporidia across the environments (Fig. 5). Non-metric multidimensional scaling showed that the majority of samples clustered together with the exception of some marine benthic, marine pelagic, and terrestrial samples.

4. Discussion

Our data show that short-branch Microsporidia are ubiquitous protists that occur in all of the investigated environments with many of them not closely related to already sequenced isolates (Figs. 1 and 2). Although widespread, we nevertheless found differences in their abundance across environments, with terrestrial accounting for the highest number of samples with short-branch Microsporidia, followed by freshwater (Figs. 1 and 4). However, since there were no clear community differences between environments (Fig. 5), we can infer that sampling in terrestrial and freshwater environments would cover a large fraction of the overall diversity of the short-branch Microsporidia. These

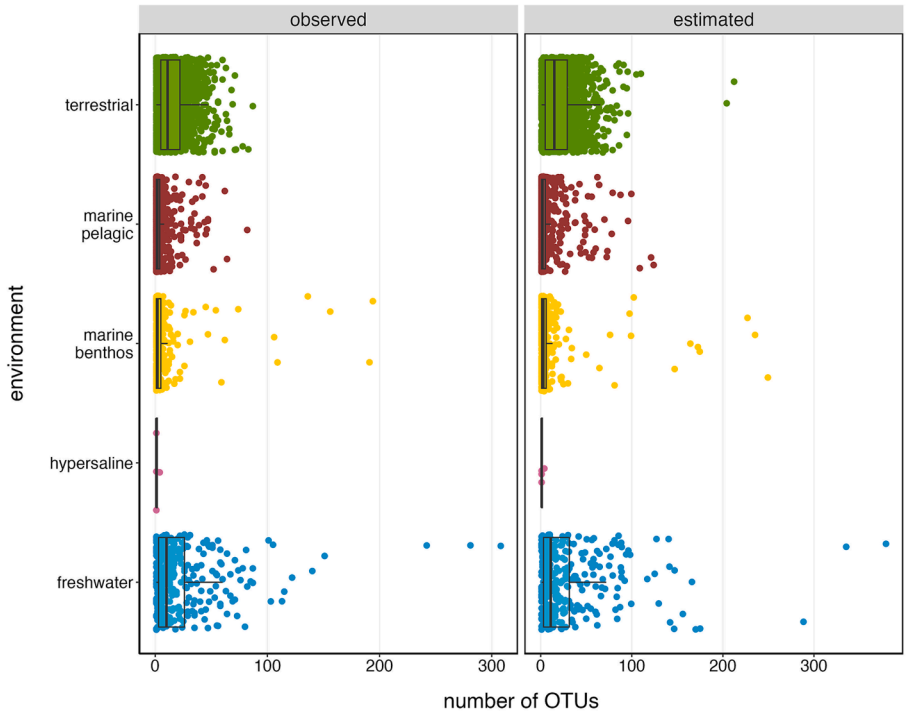

Fig. 4. Alpha-diversity per environment for the observed number of OTUs and estimated number of OTUs using Chao1.

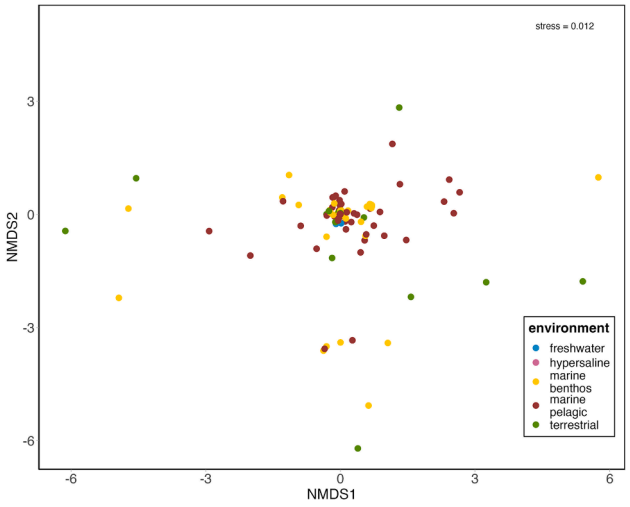

Fig. 5. Non-metric multidimensional scaling plot of all samples across the five different environments using the Jaccard index as the dissimilarity measure.

findings can be used to direct future isolation of short-branch Microsporidia.

Using environmental sequencing data targeting the SSU rRNA gene has revealed an incredible diversity of unicellular eukaryotes, including microsporidians (Logares et al., 2020; Massana et al., 2015; Santoferrara et al., 2020; de Vargas et al., 2015). Many of these studies led to novel findings and highlighted that Microsporidia are widespread parasites with a large part of their diversity that remains undescribed (Ardila-Garcia et al., 2013; Dubuffet et al., 2021; Murareanu et al., 2021; Williams et al., 2018). However, since short-branch Microsporidia are generally referred to as 'microsporidian-like', and because they are frequently considered to group indistinguishably from the rozellids, published studies often did not include them when investigating the diversity and distribution of long-branch Microsporidia (Bass et al., 2018). Additionally, it was not until Bass et al. (2018) that the relationship between long-branch and short-branch Microsporidia was solidly known. Here, with our dataset of 3,063 samples and only selecting short-branch Microsporidia sequences, our results show that this high diversity not only applies to the long-branch Microsporidia, but also applies to the short-branch Microsporidia (Fig. 1). We also found that around 30 % of the OTUs analyzed here had no close reference (<90 % similarity) within the taxonomic reference database (Fig. 2), and may therefore represent further undescribed lineages, underpinning the need to find and sequence more short-branch Microsporidia.

Beyond the standard problems of using an environmental metabarcoding approach to evaluate protistan environmental diversity (Santoferrara et al., 2020), the uneven sampling in our study could have led to an underestimation of short-branch Microsporidia in some parts of Earth. In particular, large parts of Africa, Australia as well as of North and South America are underrepresented or not represented at all in this study, and there was uneven sampling in the different layers of the oceans. Additionally, this study analyzed just V4 data derived from general eukaryotic primers. General eukaryotic primers are known to amplify many of the short-branch Microsporidia (Bass et al., 2018), while Microsporidia-specific primers tend just to amplify the long-branch Microsporidia (Doliwa et al., 2023). Even with these limitations in mind, our results still show that terrestrial and freshwater environments harbor a large diversity of short-branch Microsporidia that warrant further exploration.

We found that 216 of the short-branch Microsporidia OTUs were shared between freshwater, marine benthos, marine pelagic, and terrestrial environments, and 155 OTUs were shared between freshwater and terrestrial environments, thus increasing our level of confidence in the reality of this diversity (Fig. 3). Although early classifications divided Microsporidia into the 'Aquasporidia', 'Marinosporidia' and 'Terresporidia' (Vossbrinck and Debrunner-Vossbrinck, 2005), more recent studies have shown that many species are associated with more than one environment (Murareanu et al., 2021). This heterogeneity may be the result of their hosts being widespread and occurring in different environments themselves (Park et al., 2020). For example, Murareanu et al. (2021) found that 27.4 % of microsporidian hosts were classified to occur in more than one environment.

Furthermore, some microsporidians have been classified as generalists and can infect more than one host species in different environments (Stentiford et al., 2016). Both of these scenarios could apply to many lineages of short-branch Microsporidia, especially for those two OTUs that were associated with all environments. Despite the many shared OTUs, we also found a considerable number of unique OTUs within freshwater and terrestrial environments (Fig. 3; 192 and 604 OTUs, respectively) that may allow to resolve some of the uncharacterized diversity. In addition, although the large diversity of environmental lineages of short-branch Microsporidia was first shown in neotropical rainforest soils by Bass et al. (2018), we found more diversity within 17 samples taken from freshwater and marine benthos environments of the northern hemisphere. These samples came from the freshwater lake Pollevann and fjords close to Oslo, Norway as well as from lake Sanabria, Spain.

The greater diversity and abundance of short-branch Microsporidia in freshwater lakes is also in agreement with the assumption that evolutionary transitions to long-branch Microsporidia have occurred in freshwater environments due to the accessibility to free-living amoebae as transfer hosts (Corsaro et al., 2019). However, our results also showed that many short-branch Microsporidia OTUs were found within samples from Norwegian fjords. One possible explanation may be that many of the free-living amoebae that have been identified as hosts have also been described from marine environments (Page, 1977), or related amoebae in marine and non-marine environments are similarly infected by short-branch Microsporidia. This could also apply to the two OTUs that were shared across all five investigated environments and which had *Paramicrosporidium* as the closest genus. These references were originally found in freshwater environments. The original sequence assigned to *Paramicrosporidium* clone LKM-46 was isolated from cultures containing water from lake Ketelmeer, Netherlands, and various sources of detritus (van Hannen et al., 1999). However, it was not clear in the original study that the sequences were derived from parasitic species and what their potential hosts were.

The findings of this study also further our understanding of what is known about the partially characterized lineages so far. For example, *Mitosporidium daphniae* infects the water flea *Daphnia*, which occurs in freshwater habitats, and *Nucleophaga terricolae* infects free-living amoebae that were isolated from the bark of trees (Corsaro et al., 2016; Michel et al., 2012). The same applies to other partially-characterized lineages that were also isolated from hosts derived from freshwater or terrestrial samples (Corsaro et al., 2016, 2020; Michel et al., 2000, 2009a). A recent study investigating host-parasite interactions in a freshwater lake found that phytoplankton and microzooplankton (e.g., ciliates) are potential hosts of Microsporidia and suggested that they could be of greater importance for the functioning of lake ecosystems than previously known (Chauvet et al., 2022). The same could apply to the short-branch Microsporidia since many of their identified hosts are protists and microscopic animals.

5. Conclusion

Using environmental metabarcoding data from samples taken across the world, our study shows that there is a tremendous diversity of short-branch Microsporidia, especially in freshwater and terrestrial environments. In specific, some samples originating from Norway and Spain were particularly rich in short-branch Microsporidia. These data can be used to direct future sampling campaigns, with the goal of isolating and further characterizing partially described as well as novel environmental lineages. Furthering our knowledge about these lineages of short-branch Microsporidia may allow us to better understand the evolution that occurred among the long-branch Microsporidia.

CRediT authorship contribution statement

Megan Gross: Writing – original draft, Visualization, Investigation, Funding acquisition, Formal analysis. **Ľubomír Rajter:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis. **Frédéric Mahé:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis, Data curation, Conceptualization. **David Bass:** Writing – review & editing, Writing – original draft, Conceptualization. **Cédric Berney:** Writing – review & editing, Data curation. **Nicolas Henry:** Writing – review & editing, Data curation. **Colomban de Vargas:** Writing – review & editing. **Micah Dunthorn:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Conceptualization.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

This work was funded by an Erasmus + grant to M.G. EukBank was made possible by funds from the Gordon and Betty Moore Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejop.2024.126119.

References

Adl, S.M., Bass, D., Lane, C.E., Lukeš, J., Schoch, C.L., Smirnov, A., Agatha, S., Berney, C., Brown, M.W., Burki, F., Cárdenas, P., Čepička, I., Chistyakova, L., Del Campo, J., Dunthorn, M., Edvardsen, B., Eglit, Y., Guillou, L., Hampl, V., Heiss, A.A., Hoppenrath, M., James, T.Y., Karnkowska, A., Karpov, S., Kim, E., Kolisko, M., Kudryavtsev, A., Lahr, D.J.G., Lara, E., Le Gall, L., Lynn, D.H., Mann, D.G., Massana, R., Mitchell, E.A.D., Morrow, C., Park, J.S., Pawlowski, J.W., Powell, M.J., Richter, D.J., Rueckert, S., Shadwick, L., Shimano, S., Spiegel, F.W., Torruella, G., Youssef, N., Zlatogursky, V., Zhang, Q., 2019. Revisions to the classification, nomenclature, and diversity of eukaryotes. J. Eukaryot. Microbiol. 66, 4–119. https://doi.org/10.1111/jeu.12691.

Ardila-Garcia, A.M., Raghuram, N., Sihota, P., Fast, N.M., 2013. Microsporidian diversity in soil, sand, and compost of the Pacific Northwest. J. Eukaryot. Microbiol. 60, 601–608. https://doi.org/10.1111/jeu.12066.

Bass, D., Czech, L., Williams, B.A.P., Berney, C., Dunthorn, M., Mahé, F., Torruella, G., Stentiford, G.D., Williams, T.A., 2018. Clarifying the relationships between Microsporidia and Cryptomycota. J. Eukaryot. Microbiol. 65, 773–782. https://doi.org/10.1111/jeu.12519.

Becnel, J.J., Andreadis, T.G., 1999. Microsporidia in insects. In: Wittner, M., Weiss, L.M. (Eds.), The Microsporidia and Microsporidiosis. ASM Press, Washington DC, pp. 447–501.

Berney, C., Ciuprina, A., Bender, S., Brodie, J., Edgcomb, V., Kim, E., Rajan, J., Parfrey, L. W., Adl, S., Audic, S., Bass, D., Caron, D.A., Cochrane, G., Czech, L., Dunthorn, M., Geisen, S., Glöckner, F.O., Mahé, F., Quast, C., Kaye, J.Z., Simpson, A.G.B., Stamatakis, A., del Campo, J., de Yilmaz, P., Vargas, C., 2017. UniEuk: time to speak a common language in protistology! J. Eukaryot. Microbiol. 64, 407–411. https://doi.org/10.1111/jeu.12414.

Berney, C., Henry, N., Mahé, F., Richter, D.J., de Vargas, C., 2022. EukRibo: a manually curated eukaryotic 18S rDNA reference database to facilitate identification of new diversity. bioRxiv 11. doi: 10.1101/2022.11.03.515105.

Berney, C., Mahé, F., Henry, N., Lara, E., de Vargas, C., EukBank consortium, 2023. EukBank 18S V4 dataset [Data set]. Zenodo. Doi: 10.5281/zenodo.7804946.

Brown, J.R., Doolittle, W.F., 1999. Gene descent, duplication, and horizontal transfer in the evolution of glutamyl- and glutaminyl-tRNA synthetases. J. Mol. Evol. 149, 485–495. https://doi.org/10.1007/PL00006571.

Brownrigg, R., 2021. maps: Draw Geographical Maps. R Package Version 3 (4). https://CRAN.R-project.org/package=maps.

Chauvet, M., Debroas, D., Moné, A., Dubuffet, A., Lepère, C., 2022. Temporal variations of Microsporidia diversity and discovery of new host–parasite interactions in a lake ecosystem. Environ. Microbiol. 24, 1672–1686. https://doi.org/10.1111/1462-2920.15950.

Corradi, N., Keeling, P.J., 2009. Microsporidia: a journey through radical taxonomical revisions. Fungal. Biol. Rev. 23, 1–8. https://doi.org/10.1016/j.fbr.2009.05.001.

Corsaro, D., Walochnik, J., Venditti, D., Müller, K.D., Hauröder, B., Michel, R., 2014a. Rediscovery of *Nucleophaga* amoebae, a novel member of the Rozellomycota. Parasitol. Res. 113, 4491–4498. https://doi.org/10.1007/s00436-014-4138-8.

Corsaro, D., Walochnik, J., Venditti, D., Steinmann, J., Müller, K-D., Michel, R., 2014b. Microsporidia-like parasites of amoebae belong to the early fungal lineage Rozellomycota. Parasitol. Res. 113, 1909–1918. <https:/doi: 10.1007/s00436-014-3838-4>.

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. https://doi.org/10.1007/s00436-016-5055-9.

Corsaro, D., Wylezich, C., Venditti, D., Michel, R., Walochnik, J., Wegensteiner, R., 2019. Filling gaps in the microsporidian tree: rDNA phylogeny of *Chytridiopsis typographi* (Microsporidia: Chytridiopsida). Parasitol. Res. 118, 169–180. https://doi.org/10.1007/s00436-018-6130-1.

Corsaro, D., Walochnik, J., Venditti, D., Hauröder, B., Michel, R., 2020. Solving an old enigma: *Morellospora saccamoebae* gen. nov., sp. nov. (Rozellomycota), a *Sphaerita*-like parasite of free-living amoebae. Parasitol. Res. 119, 925–934. https://doi.org/10.1007/s00436-020-06623-5.

Cuomo, C.A., Desjardins, C.A., Bakowski, M.A., Goldberg, J., Ma, A.T., Becnel, J.J., Didier, E.S., Fan, L., Heiman, D.I., Levin, J.Z., Young, S., Zeng, Q., Troemel, E.R., 2012. Microsporidian genome analysis reveals evolutionary strategies for obligate intracellular growth. Genome. Res. 22, 2478–2488. https://doi.org/10.1101/gr.142802.112.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J. M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Coordinators, T.O., Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P., Karsenti, E., 2015. Eukaryotic plankton diversity in the sunlit ocean. Science 348, e1261605. https://doi.org/10.1126/science.1261605.

Dean, P., Sendra, K.M., Williams, T.A., Watson, A.K., Major, P., Nakjang, S., Kozhevnikova, E., Goldberg, A.V., Kunji, E.R.S., Hirt, R.P., Embley, T.M., 2018. Transporter gene acquisition and innovation in the evolution of Microsporidia intracellular parasites. Nat. Commun. 9, e1709.

Dobson, A., Lafferty, K.D., Kuris, A.M., Hechinger, R.F., Jetz, W., 2008. Homage to Linnaeus: How many parasites? How many hosts? PNAS 105 (Suppl. 1), 11482–11489. https://doi.org/10.1073/pnas.0803232105.

Doliwa, A., Dunthorn, M., Rassoshanska, E., Mahé, F., Bass, D., Duarte Ritter, C., 2021. Identifying potential hosts of short-branch Microsporidia. Microb. Ecol. 82, 549–553. https://doi.org/10.1007/s00248-020-01657-9.

Doliwa, A., Grabner, D., Sures, B., Dunthorn, M., 2023. Comparing Microsporidia-targeting primers for environmental DNA sequencing. Parasite 30, e52.

Dubuffet, A., Chauvet, M., Moné, A., Debroas, D., Lepère, C., 2021. A phylogenetic framework to investigate the microsporidian communities through metabarcoding and its application to lake ecosystems. Environ. Microbiol. 23, 4344–4359. https://doi.org/10.1111/1462-2920.15618.

Dusa, A., 2021. venn: Draw Venn diagrams. R Package Vers. 1, 10. https://CRAN.R-project.org/package=venn.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl. Acids Res. 32, 1792–1797. https://doi.org/10.1093/NAR/GKH340.

Foissner, I., Foissner, W., 1995. *Ciliatosporidium platyophryae* nov. gen., nov. spec. (Microspora incerta sedis), a parasite of *Platyophrya terricola* (Ciliophora, Colpodea). Eur. J. Protistol. 31, 248–259. https://doi.org/10.1016/S0932-4739(11)80088-X.

Franzen, C., 2008. Microsporidia: a review of 150 years of research. Open. Parasitol. J. 2, 1–34. https://doi.org/10.2174/1874421400802010001.

Fries, I., 1993. *Nosema apis*—a parasite in the honey bee colony. Bee World 74, 5–19. https://doi.org/10.1080/0005772X.1993.11099149.

Gill, E.E., Fast, N.M., 2006. Assessing the microsporidia-fungi relationship: combined phylogenetic analysis of eight genes. Gene 375, 103–109. https://doi.org/10.1016/j.gene.2006.02.023.

Gill, E.E., Fast, N.M., 2007. Stripped-down DNA repair in a highly reduced parasite. BMC Mol. Biol. 8, e24.

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. PNAS 111, 15480–15485. https://doi.org/10.1073/pnas.1410442111.

Hirt, R.P., Logsdon, J.M., Healy, B., Dorey, M.W., Doolittle, W.F., Embley, T.M., 1999. Microsporidia are related to Fungi: evidence from the largest subunit of RNA polymerase II and other proteins. PNAS 96, 580–585. https://doi.org/10.1073/pnas.96.2.580.

Kassambara, A., 2020. ggpubr: ggplot2 Based Publication Ready Plots. R Package Version (4). https://CRAN.R-project.org/package=ggpubr.

Kassambara, A., 2021. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R Package Version (7). https://CRAN.R-project.org/package=rstatix.

Katinka, M.D., Duprat, S., Cornillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., Vivarès, C.P., 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414, 450–453. https://doi.org/10.1038/35106579.

Keeling, P.J., 2003. Congruent evidence from α-tubulin and β-tubulin gene phylogenies for a zygomycete origin of microsporidia. Fungal Genet. Biol. 38, 298–309. https://doi.org/10.1016/S1087-1845(02)00537-6.

Keeling, P.J., 2014. Phylogenetic place of Microsporidia in the tree of eukaryotes. In: Weiss, L.M., Becnel, J.J. (Eds.), Microsporidia: Pathogens of Opportunity, first ed. John Wiley & Sons Ltd., pp. 195–202

Keeling, P.J., Corradi, N., 2011. Shrink it or lose it: balancing loss of function with shrinking genomes in the microsporidia. Virulence 2, 67–70. https://doi.org/10.4161/viru.2.1.14606.

Keeling, P.J., Fast, N.M., 2002. Microsporidia: biology and evolution of highly reduced intracellular parasites. Annu. Rev. Microbiol. 56, 93–116. https://doi.org/10.1146/annurev.micro.56.012302.160854.

Keeling, P.J., McFadden, G.I., 1998. Origins of microsporidia. Trends. Microbiol. 6, 19–23. https://doi.org/10.1016/S0966-842X(97)01185-2.

Kent, M.L., Elliott, D.G., Groff, J.M., Hedrick, R.P., 1989. *Loma salmonae* (Protozoa: Microspora) infections in seawater reared coho salmon *Oncorhynchus kisutch*. Aquaculture 80, 211–222. https://doi.org/10.1016/0044-8486(89)90169-5.

Logares, R., Deutschmann, I.M., Junger, P.C., Giner, C.R., Krabberød, A.K., Schmidt, T.S. B., Rubinat-Ripoll, L., Mestre, M., Salazar, G., Ruiz-González, C., Sebastián, M., de Vargas, C., Acinas, S.G., Duarte, C.M., JGasol, J.M., Massana, R., 2020. Disentangling the mechanisms shaping the surface ocean microbiota. Microbiome 8, 1–17. https://doi.org/10.1186/s40168-020-00827-8.

Loker, E.S., Hofkin, B.V., 2022. Parasitology: A Conceptual Approach, second ed. CRC Press. doi:10.1201/9780429277405.

Mahé, F., 2021. mumu: post-clustering curation tool for metabarcoding data. Version 1.0.2. https://github.com/frederic-mahe/mumu.

Mahé, F., de Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., Singer, D., Mayor, J., Bunge, J., Sernaker, S., Siemensmeyer, T., Trautmann, I., Romac, S., Berney, C., Kozlov, A., Mitchell, E.A.D., Seppey, C.V.W., Egge, E., Lentendu, G., Wirth, R., Trueba, G., Dunthorn, M., 2017. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. Nat. Ecol. Evol. 1, e0091.

Mahé, F., Czech, L., Stamatakis, A., Quince, C., de Vargas, C., Dunthorn, M., Rognes, T., 2022. Swarm v3: towards tera-scale amplicon clustering. Bioinformatics 38, 267–269. https://doi.org/10.1093/bioinformatics/btab493.

Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., Chambouvet, A., Christen, R., Claverie, J.-M., Decelle, J., Dolan, J.R., Dunthorn, M., Edvardsen, B., Forn, I., Forster, D., Guillou, L., Jaillon, O., Kooistra, W.H.C.F., Logares, R., Mahé, F., Not, F., Ogata, H., Pawlowski, J., Pernice, M.C., Probert, I., Romac, S., Richards, T., Santini, S., Shalchian-Tabrizi, K., Siano, R., Simon, N., Stoeck, T., Vaulot, D., Zingone, A., de Vargas, C., 2015. Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. Environ. Microbiol. 17, 4035–4049. https://doi.org/10.1111/1462-2920.12955.

McMurdie, P.J., Holmes, S., 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE 8, e61217.

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. https://doi.org/10.1007/s004360050704.

Michel, R., Hauröder, B., Zöller, L., 2009a. Isolation of the amoeba *Thecamoeba quadrilineata* harbouring intranuclear spore forming endoparasites considered as fungus-like organisms. Acta Protozool. 48, 41–49.

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

Michel, R., Müller, K.D., Schmid, E.N., Theegarten, D., 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.

Mikryukov, V., 2021. metagMisc: miscellaneous functions for metagenomic analysis. R Package Vers. 4. https://doi.org/10.5281/zenodo.571403.

Murareanu, B.M., Sukhdeo, R., Qu, R., Jiang, J., Reinke, A.W., 2021. Generation of a microsporidia species attribute database and analysis of the extensive ecological and phenotypic diversity of microsporidia. mBio 12, e01490–e01521. https://doi.org/10.1128/mbio.01490-21.

Page, F.C., 1977. The genus *Thecamoeba* (Protozoa, Gymnamoebia) species distinctions, locomotive morphology, and protozoan prey. J. Nat. Hist. 11, 25–63. https://doi.org/10.1080/00222937700770031.

Park, E., Poulin, R., 2021. Revisiting the phylogeny of microsporidia. Int. J. Parasitol. 51, 855–864. https://doi.org/10.1016/j.ijpara.2021.02.005.

Park, E., Jorge, F., Poulin, R., 2020. Shared geographic histories and dispersal contribute to congruent phylogenies between amphipods and their microsporidian parasites at regional and global scales. Mol. Ecol. 29, 3330–3345. https://doi.org/10.1111/MEC.15562.

Poulin, R., 2011. The many roads to parasitism. A tale of convergence. Adv. Parasitol. 74, 1–40. https://doi.org/10.1016/B978-0-12-385897-9.00001-X.

Poulin, R., Randhawa, H.S., 2015. Evolution of parasitism along convergent lines: from ecology to genomics. Parasitology 142, 6–15. https://doi.org/10.1017/S0031182013001674.

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.

R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org.

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584.

Santoferrara, L., Burki, F., Filker, S., Logares, R., Dunthorn, M., McManus, G.B., 2020. Perspectives from ten years of protist studies by high-throughput metabarcoding. J. Eukaryot. Microbiol. 67, 612–622. https://doi.org/10.1111/jeu.12813.

Stentiford, G.D., Becnel, J.J., Weiss, L.M., Keeling, P.J., Didier, E.S., Bjornson, S., Freeman, M.A., Brown, M.J.F., Roesel, K., Sokolova, Y., 2016. Microsporidia – emergent pathogens in the global food chain. Trends Parasitol. 32, 336–348. https://doi.org/10.1016/j.pt.2015.12.004.

Strassert, J.F.H., Monaghan, M.T., 2022. Phylogenomic insights into the early diversification of fungi. Curr. Biol. 32, 3628–3635. https://doi.org/10.1016/j.cub.2022.06.057.

Thomarat, F., Vivarès, C.P., Gouy, M., 2004. Phylogenetic analysis of the complete genome sequence of *Encephalitozoon cuniculi* supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. J. Mol. Evol. 59, 780–791. https://doi.org/10.1007/s00239-004-2673-0.

van Hannen, E.J., Mooij, W., van Agterveld, M.P., Gons, H.J., Laanbroek, H.J., 1999. Detritus-dependent development of the microbial community in an experimental system: qualitative analysis by denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. 65, 2478–2484. https://doi.org/10.1128/AEM.65.6.2478-2484.1999.

Vávra, J., Lukeš, J., 2013. Microsporidia and 'the art of living together'. Adv. Parasitol. 82, 253–319. https://doi.org/10.1016/B978-0-12-407706-5.00004-6.

Voigt, K., James, T.Y., Kirk, P.M., da Santiago, A.L.C.M., Waldman, B., Griffith, G.W., Fu, M., Radek, R., Strassert, J.F.H., Wurzbacher, C., Jerónimo, G.H., Simmons, D.R., Seto, K., Gentekaki, E., Hurdeal, V.G., Hyde, K.D., Nguyen, T.T.T., Lee, H.B., 2021. Early-diverging fungal phyla: taxonomy, species concept, ecology, distribution, anthropogenic impact, and novel phylogenetic proposals. Fungal Divers. 109, 59–98. https://doi.org/10.1007/s13225-021-00480-y.

Vossbrinck, C.R., Debrunner-Vossbrinck, B.A., 2005. Molecular phylogeny of the Microsporidia: ecological, ultrastructural and taxonomic considerations. Folia Parasitol. 52, 131–142. https://doi.org/10.14411/fp.2005.017.

Vossbrinck, C.R., Maddox, J.V., Friedman, S., Debrunner-Vossbrinck, B.A., Woese, C.R., 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326, 411–414. https://doi.org/10.1038/326411a0.

Weber, R., Bryan, R.T., Schwartz, D.A., Owen, R.L., 1994. Human microsporidial infections. Clin. Microbiol. Rev. 7, 426–461. https://doi.org/10.1128/CMR.7.4.426.

Weiss, L.M., Becnel, J.J., 2014. Microsporidia: Pathogens of Opportunity, first ed. John Wiley & Sons Inc.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. J. Open Source Softw. 4, e1686.

Williams, B.A.P., Hamilton, K.M., Jones, M.D., Bass, D., 2018. Group-specific environmental sequencing reveals high levels of ecological heterogeneity across the microsporidian radiation. Environ. Microbiol. Rep. 10, 328–336. https://doi.org/10.1111/1758-2229.12642.