INTRODUCTION

High-throughput sequencing of rRNA gene amplicons is increasingly applied to analyse the microbial diversity in terrestrial (e.g., Mahé et al., 2017), freshwater (e.g., Yi et al., 2017) and marine (e.g., de Vargas et al., 2015) ecosystems. Linking microbial diversity to the functioning and dynamics of microbial communities and ultimately to environmental responses and ecosystem processes requires correct interpretation of ecological traits of the identified taxa (Bouskill, Eveillard, Chien, Jayakumar, & Ward, 2012; Krause et al., 2014). However, information about the functional traits of microbes is limited and difficult to access, being broadly distributed across the literature from many decades. To meet the analytical demands of large-scale sequencing projects, trait-based data must be collated and made more easily accessible. Completing such databases is a priority for microbial ecology, as it has been done for bacteria (FRAPROTAX) (Louca, Parfrey, & Doebeli, 2016), fungi (FUNGuild) (Nguyen et al., 2016) and nematodes (NEMAGuild_db at https://github.com/brendanf/FUNGuildR, last accessed October 2019).

Accordingly, so far most trait-based studies have been conducted on plants, animals (Kunstler et al., 2013; Stuart-Smith et al., 2013) and on a few groups of microorganisms, mostly prokaryotes (Bouskill et al., 2012). As stated by Krause et al. (2014) the major requirements for trait-based studies in microbial ecology are “(a) defining which microbial traits are important with respect..."
to ecosystem functioning or particular ecosystem functions, and (b) measuring these relevant traits.” The last 10 years have seen a welcome increase of focus on the less studied part of the microbial biome, the protists (Bonkowski, Dumack, & Fiore-Donno, 2019). Protists are an immensely heterogeneous, polyphyletic assemblage of unicellular organisms featuring a vast array of functional roles (Pawlowski et al., 2012). Depending on their associated functional traits, different protistan groups will thus play contrasting or complementary roles in essential ecosystem functions and services, including primary production and nutrient cycling. Understanding traits and their redundancy in microbial taxa may thus facilitate the interpretation of biotic, abiotic and global change drivers on the protist community. For example, protistan functional traits have been used to identify the main actors driving the shift between net heterotrophy and autotrophy in the oceans and to establish models predicting phytoplanktonic blooms (Alexander et al., 2015). Protists are an immensely heterogeneous, polyphyletic assemblage of unicellular organisms featuring a vast array of functional traits; nevertheless, we mainly chose traits with known or and thoroughly assessed with an established protocol (Fiore-Donno et al., 2018). Because amplicon-based studies comprise the majority of the environmental surveys conducted so far, and the markers most frequently used for protists are sequences of the rRNA genes (mostly 18S, variable regions V1, V4 and V9) (Pawlowski et al., 2012), we focused on taxa for which 18S sequences are publicly available.

2 | MATERIALS AND METHODS

Our work is based on an extensive literature search. As a starting point, we used the classifications of Adl et al. (2019) and of Cavalier-Smith et al. (2018), and screened for functional traits only for taxa for which an 18S sequence was available.

2.1 | Taxonomy

The main taxonomic subdivisions are: (a) Endomyxa including Vampyrellida, Phytomyxea, Filoreta, Gromia and Ascetospora (Bass et al., 2019; Sierra et al., 2016); (b) Cercozoa including the classes Sarcomonadea, Thecofilosea, Granofilosea, Imbricatea, Chlorarachnea and Metromonadea (Cavalier-Smith et al., 2018); and (c) the incertae sedis rhizarian clades Aquavelonida, Tremulida and Novel Clade 12 (Bass et al., 2018).

Our table is hierarchically structured, with descending taxonomic ranks from left to right, down to the genus. We consider the genus level to be the most pragmatic for the purpose of assigning functional traits, but at the same time recognize that traits may differ significantly between species within each genus (e.g., food and parasite host preferences, morphology and behaviour). However, sequences in reference databases are often not assigned to species. In addition, the resolution of the environmental sequencing data lacks precision at the species level, because similar sequences are usually clustered together in operational taxonomic units (OTUs), to compensate for amplification and sequencing errors. Consequently, closely related species will not (or rarely) be recognized in environmental sequencing data, making the genus level more adequate for our purpose. Each taxonomic rank has attributes to it, and therefore the information is compatible with results from environmental sampling. For example, in the Silva database, sequences can be identified to any taxonomic rank (order, family or genus). A simple command or script can be used to join the user’s taxonomic table with our functional table, based on the taxonomic entries (Table S1).

2.2 | Functional traits

We attempted to be as exhaustive as possible in the selection of functional traits; nevertheless, we mainly chose traits with known or potential ecological importance. We considered that the organisms’
morphology, mode of movement, habitat and nutritional preferences were essential features. Thus, we classified the traits of interest in four categories: morphology, locomotion or nutrition modes, and habitats (Table S1).

Each category was divided into discrete units to be compatible with statistical analyses (i.e., each taxon can only be assigned once in each category). We attributed traits to taxa by searching in the relevant literature, starting from the original description, until we could find all attributes for each taxon. All consulted references are provided (Data S1; S2; Table S1 can also be accessed via GitHub [https://github.com/Kenneth-Dumack/Functional-Traits-Cercozoa-Endomyxa]). Suggestions for updates can be addressed to the corresponding author.

The main distinction in all categories was between free-living organisms and parasites. In the morphology of the free-living cells, we separated cells protected by a shell from those that are naked. Two main categories of shells were recognized: organic/agglutinated or siliceous. We made a subcategory for the silica-shell-forming taxa because of their importance in the silica cycle in soil and water (Blard et al., 2016; Wilkinson & Mitchell, 2010). The other subcategory comprised organic and agglutinated shells, because only cells bearing an organic shell may facultatively develop an agglutinated shell around the organic layer (Dumack, Bonkowski, Clauš, & Völker, 2018). We assumed that amoebae, amoeboidflagellates and flagellates would differ in their locomotion and feeding behaviour and thus distinguished between them.

Three main locomotion modes were recognized: organisms bound to the substrate (i.e., gliding) or freely swimming. In nutrition, we first separated autotrophs (most of the marine Chlorarachnea) from the remainder. We further distinguished bacteriovores, eukaryovores (feeding on fungi, microfauna, algae or other protists) and omnivores (feeding on bacteria and eukaryotes). We could not create more precise categories (e.g., fungivores) because of a lack of information (or contradictory reports/probable multiple trophic modes) for most taxa. We considered most Cercomonadidae as omnivores, based on recent observations (Flues, Bass, & Bonkowski, 2017; Geisen et al., 2016), although most characterized taxa were isolated with bacteria as the sole food source (Bass et al., 2009; Howe, Bass, Chao, & Cavalier-Smith, 2011; Howe, Bass, Vickerman, Chao, & Cavalier-Smith, 2009). However, many taxa that have been originally considered as bacteriovorous may actually be omnivorous, feeding on small flagellates, algae or yeasts, as ongoing observations and studies suggest. Further evidence from feeding experiments is required in this respect for many taxa.

We estimated that parasites of plants or stramenopiles (Phytomyxea) and animals (Ascetospora) spent most of their life cycle as nonmotile endoparasites (within hosts), although they may be liberated as spores or ephemeral motile stages (Neuhäuser et al., 2014).

Regarding ecology, we set the main distinction between soil and freshwater and marine taxa, because freshwater and soil-inhabiting taxa may easily switch habitat, while marine taxa are rarely found in terrestrial or limnic habitats (Heger et al., 2010; Shalchian-Tabrizi et al., 2008). Genera known to accommodate species from marine and soil and freshwater habitats were considered ecologically ubiquitous.

3 | RESULTS AND DISCUSSION

To improve our knowledge of ecosystem functioning, we need an enhanced understanding of the functional relationships between components of the microbial communities—bacteria, fungi and protists—and their environment transcending the traditional taxonomic framework. Understanding ecological processes can be facilitated through trait-based approaches that identify linkages between phenotypic and ecological traits of organisms. However, knowledge in this area for protists currently lags behind that for bacteria and fungi. Transcriptomic-based functional assignment is routinely done for bacteria and fungi (Hesse et al., 2015; Lladó, Žifčáková, Větrovský, Eichlerová, & Baldrian, 2016; Tveit, Urih, Frenzel, & Svenning, 2015; Žifčáková, Větrovský, Howe, & Baldrian, 2016), but such studies using protists are scarcer (Alexander et al., 2015; Faure et al., 2019). This is due largely to still thinly and patchily populated reference databases with respect to protists, which are in any case oriented towards parasitic protists of human concern (del Campo et al., 2014). Additionally, culture-based experiments and behavioural/ecological observations of protists are now less widely undertaken than in the past, leading to a gap in such knowledge, which is regrettable as we are increasingly aware of the very high protistan diversity.

For environmental microbial diversity and ecology studies, the gold standard is currently amplicon-based surveys using high-throughput sequencing (de Vargas et al., 2015; Mahé et al., 2017) allowing the necessary sequencing depth for large sampling. Our database has been designed to add ecological meaning to environmental studies reporting rhizarian sequences. For this, we compiled a reference database for cercozoan and endomyxan genera categorized by morphological and ecological features here termed functional traits. For full transparency, we provide the sources consulted for each functional trait.

With respect to soil Cercozoa, we revealed several interesting and new interactions between taxonomic or functional groups, and the ecological processes shaping them (Fiore-Donno et al., 2019). For example, we showed a correlation between abundances of OTUs of (a) testate cells and soil moisture, (b) bacteriovores and bacteria, (c) freely swimming protists and soil bulk density, and (d) plant parasites and C microbial biomass (Figure 1). We compiled a table comprising each so far sequenced cercozoan and endomyxan genus and its associated traits in soil, marine and limnic environments (Table S1). For the table to be used in statistical analyses, categories represented by only one taxon had to be avoided. For example, there are only a few taxa that spend the majority of their life cycles as colonies (some Spongomonadidae). We did not consider this trait, although it could confer significant behavioural and/or ecological differences from noncolonial lineages, as indicated by Fiore-Donno et al. (2019). Another example
is Thaumatomonadida, which we classified as shelled, although they have interlocking silica scales, not forming an entirely rigid shell. It has been shown that they do not display the negative correlation with soil moisture as do other shelled taxa in this data set (Fiore-Donno et al., 2019). For such particular cases, it is advisable to cross-check the taxonomic and functional assignments and also to refer to the comments of the table.

Further research is likely to reveal a broader repertoire of traits for many of the taxa listed in Table S1 (updated versions will be accessible under https://github.com/Kenneth-Dumack/Functional-Traits-Cercozoa-Endomyxa). It is also desirable that trait-based data sets are eventually informative at the species level. We hope that this work, based on relevant and significant functional information at a detailed taxonomic scale, will facilitate the establishment of new models of microbial interactions in microbial ecology and more generally in ecology.

ACKNOWLEDGEMENTS
This work was funded by the DFG Priority Program 1374 “Infrastructure-Biodiversity-Exploratories” by grants BO 1907/18-1 PATHOGENS and BO 1907/13-2 MicroSystem to A.M.F.D. and M.B. This work was also supported by NERC grants NE/H009426/1 and NE/H000887/1 to D.B.

AUTHOR CONTRIBUTIONS
A.M.F.D. conceived the project. K.D. and A.M.F.D. compiled the table and A.M.F.D. and K.D. wrote the manuscript. D.B. and M.B. revised the table and the manuscript. All co-authors approved the final version of the manuscript.

ORCID
Kenneth Dumack https://orcid.org/0000-0001-8798-0483
Anna Maria Fiore-Donno https://orcid.org/0000-0001-6265-1907

David Bass https://orcid.org/0000-0002-9883-7823
Michael Bonkowski https://orcid.org/0000-0003-2656-1183

DATA AVAILABILITY STATEMENT
The functional table will be published with the manuscript as Supporting information data and on GitHub with ‘The functional table has been published as Supporting information data and on GitHub at https://github.com/Kenneth-Dumack/Functional-Traits-Cercozoa-Endomyxa’

REFERENCES
Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., … Zhang, Q. (2019). Revisions to the classification, nomenclature, and diversity of eukaryotes. *Journal of Eukaryotic Microbiology*, 66, 4–119. https://doi.org/10.1111/jeu.12691
Alexander, H., Rouco, M., Haley, S. T., Wilson, S. T., Karl, D. M., & Dyhrman, S. T. (2015). Functional group-specific traits drive phytoplankton dynamics in the oligotrophic ocean. *Proceedings of the National Academy of Sciences of the United States of America*, 112, E5972–E5979. https://doi.org/10.1073/pnas.1518165112
Bass, D., Howe, A. T., Mylnikov, A. P., Vickerman, K., Chao, E. E., Edwards Smallbone, J., … Cavalier-Smith, T. (2009). Phylogeny and classification of Cercomonadida (Protozoa, Cercozoa): Cercomonas, Eocercomonas, Paracercomonas, and Cavernomonas gen. nov. *Protist*, 160, 483–521. https://doi.org/10.1016/j.protis.2009.01.004
Bass, D., Tikhonenkov, D. V., Foster, R., Dyal, P., Janouškovec, J., Keeling, P. J., … Berney, C. (2018). Rhizarian ‘Novel Clade 10’ revealed as abundant and diverse planktonic and terrestrial flagellates, including Aquavolona n. gen. *Journal of Eukaryotic Microbiology*, 65, 828–842. https://doi.org/10.1111/jeu.12524
Bass, D., Ward, G., & Burki, F. (2019). *Ascetospora*. *Current Biology*, 29, R1–R15. https://doi.org/10.1016/j.cub.2018.11.025
Biard, T., Stemmann, L., Picheral, M., Mayot, N., Vandromme, P., Hauss, H., … Not, F. (2016). In situ imaging reveals the biomass of giant protists in the global ocean. *Nature*, 532, 504–517. https://doi.org/10.1038/nature17652
Bjorbækmo, M. F. M., Evenstad, A., Røsæg, L. L., Krabberød, A. K., & Logares, R. (2019). The planktonic protist interactome: Where do
peatlands. *Scientific Reports*, 6, 33907. https://doi.org/10.1038/srep33907

Neuhauser, S., Kirchmair, M., Bulman, S. R., & Bass, D. (2014). Cross-kingdom host shifts of phytomyxid parasites. *BMC Evolutionary Biology*, 14, 33. https://doi.org/10.1186/1471-2148-14-33

Nguyen, N. H., Song, Z., Bates, S. T., Branco, S., Tedersoo, L., Menke, J., ... Kennedy, P. G. (2016). FUNGuild: An open annotation tool for parsing fungal community datasets by ecologial guild. *Fungal Ecology*, 20, 241–248. https://doi.org/10.1016/j.funeco.2015.06.006

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., ... de Vargas, C. (2012). CBOL Protist Working Group: Barcoding eukaryotic richness beyond the animal, plant and fungal kingdoms. *PLoS Biology*, 10, e1001419. https://doi.org/10.1371/journal.pbio.1001419

Shalchian-Tabrizi, K., Bråte, J., Logares, R., Klaveness, D., Berney, C., & Jakobsen, K. S. (2008). Diversification of unicellular eukaryotes: Cryptomonad colonizations of marine and fresh waters inferred from revised 18S rRNA phylogeny. *Environmental Microbiology*, 10, 2635–2644. https://doi.org/10.1111/j.1462-2920.2008.01685.x

Sierra, R., Cañas-Duarte, S. J., Burk, F., Schwelm, A., Fogelqvist, J., Dixielius, C., ... Pawlowski, J. (2016). Evolutionary origins of rhizarian parasites. *Molecular Biology and Evolution*, 33, 980–983. https://doi.org/10.1093/molbev/msv340

Stuart-Smith, R. D., Bates, A. E., Lefcheck, J. S., Duffy, J. E., Baker, S. C., Thomson, R. J., ... Edgar, G. J. (2013). Integrating abundance and functional traits reveals new global hotspots of fish diversity. *Nature*, 501, 539–542. https://doi.org/10.1038/nature12529

Tirok, K., & Gaedke, U. (2010). Internally driven alternation of functional traits in a multispecies predator-prey system. *Ecology*, 91, 1748–1762. https://doi.org/10.1890/09-1052.1

Turner, T. R., Ramakrishnan, K., Walshaw, J., Heavens, D., Alston, M., Swarbreck, D., ... Poole, P. S. (2013). Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *The ISME Journal*, 7, 2248–2258. https://doi.org/10.1038/ismej.2013.119

Tveit, A., Urih, T., Frenzel, P., & Svenning, M. M. (2015). Metabolic and trophic interactions modulate methane production by Arctic peat microbiota in response to warming. *Proceedings of the National Academy of Sciences of the United States of America*, 112, e2507–e2516. https://doi.org/10.1073/pnas.1420797112

Wilkinson, D. M., & Mitchell, E. A. D. (2010). Testate amoebae and nutrient cycling with particular reference to soils. *Geomicrobiology Journal*, 27, 520–533. https://doi.org/10.1080/01490451003702925

Žifčáková, L., Větrovský, T., Howe, A., & Baldrian, P. (2016). Microbial activity in forest soil reflects the changes in ecosystem properties between summer and winter. *Environmental Microbiology*, 18, 288–301. https://doi.org/10.1111/1462-2920.13026

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Dumack K, Fiore-Donno AM, Bass D, Bonkowski M. Making sense of environmental sequencing data: Ecologically important functional traits of the protistan groups Cercozoa and Endomyxa (Rhizaria). *Mol Ecol Resour*. 2020;20:398–403. https://doi.org/10.1111/1755-0998.13112