https://doi.org/10.35662/unine-thesis-2873

University of Neuchâtel
Faculty of Science
Laboratory of Soil Biodiversity

Exploration and characterization of Amoebozoa diversity and investigation of their diversity patterns at regional and global scales

---

PhD THESIS

to obtain the title of
Doctor of Philosophy in Biological Science
of the University of Neuchâtel

defended by
Quentin Blandenier
on March 12, 2020

Thesis committee:
Dr. Enrique Lara (Supervisor)
Prof. Edward A. D. Mitchell (Co-supervisor)
Dr. Carlos Lado
Dr. Diego Fontaneto
Dr. Guillaume Lentendu

UNIVERSITÉ DE
NEUCHÂTEL

Faculté des Sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel – Suisse
Tél : + 41 (0)32 718 21 00
E-mail : secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

**La Faculté des sciences de l'Université de Neuchâtel
autorise l'impression de la présente thèse soutenue par**

Monsieur Quentin BLANDENIER

Titre:

“Exploration and characterization of *Amoebozoa* diversity and investigation of their diversity patterns at regional and global scales”

sur le rapport des membres du jury composé comme suit:

- Dr Enrique Lara, directeur de thèse, Jardin Botanique Royal, Madrid, Espagne
- Prof. Edward Mitchell, co-directeur de thèse, Université de Neuchâtel, Suisse
- Dr Guillaume Lentendu, Université de Neuchâtel, Suisse
- Dr Diego Fontaneto, Centre National de Recherche, Italie
- Dr Carlos Lado, Jardin Botanique Royal, Madrid Espagne

Neuchâtel, le 27 avril 2020

Le Doyen, Prof. P. Felber

P. Fel

"How could life be boring when there are still rhizopods to describe?"

Leidy

"My work, which I’ve done for a long time, was not pursued in order to gain the praise I now enjoy, but chiefly from a craving after knoweledge, which I notice resides in me more than most other men."

Leeuwenhoek, Letter of 12 June 1716

"Nobody exists on purpose. Nobody belongs anywhere.
We’re all going to die. Come watch TV."

Rick and Morty

Table of contents

Abstract

The world eukaryotic diversity is dominated by (mostly) single-celled organisms referred to as protists. Among them, the Amoebozoa are one of the most numerous, diverse and characteristic groups in soil, thus playing important roles in ecosystems functioning. However, their study has been impeded by the difficulty in detecting them and the lack of stable morphological traits in most groups. Nevertheless, some amoebozoans such as the Hyalospheniformes (Arcellinida) are characterized by a self-constructed *test* (i.e. shell) which facilitates their identification, and are then considered as a suitable model group for investigating diversity patterns of repartition. The recent development of DNA barcoding has helped considerably taxonomic identification, whereas metabarcoding has allowed revealing microbial community composition without observational and cultivation biases. These methods have proved efficient for several microbial groups, but only few studies have been designed for Amoebozoa and available protocols are still rather scarce. The aims of my thesis were then to 1) improve and develop molecular methods to study the amoebozoan diversity and ecology, 2) estimate the taxonomic and functional diversity in the soil, 3) improve the taxonomic and phylogenetic frame for this diversity in order to build a sound basis for further research and 4) characterize the ecological drivers which are likely to influence microbial diversity at local, continental and global scales.

We first identified a new molecular marker to survey arcellinids taxa, which proved to be efficient for discriminating closely-related taxa and simultaneously investigating deep relationships among distant taxa (Chapter 2). In addition, we also adapted a metabarcoding protocol with specific COI primers to survey the diversity within the genus *Nebela* at a fine taxonomical resolution (Chapter 6).

Then, we isolated, cultivated and described the first member of a deep-branching environmental clade of Amoebozoa (Chapter 3). This amoeba, one of the smallest amoeboid species described, presents a unique life cycle with an alternation of phagotrophic active trophozoites and osmotrophic fungi-like ramifications. Its presence has been pervasively reported in many soil metabarcoding studies, but this organism had never been characterized. By contrast, Hyalospheniformes are known since the works of Ehrenberg in the 19$^{\text{th}}$ century. However, their diversity at the species level remains poorly characterized. In chapter 4, we showed that the iconic testate amoeba species *Nebela militaris* did not belong to genus *Nebela* but branched as a separate entity in the Hyalospheniformes tree. Therefore, we erected the new genus *Alabasta* for this species (Chapter 4).

In addition, we demonstrated that Hyalospheniformes diversity had been greatly underestimated. Indeed, our morphological and molecular results have revealed the presence of several species within the genera *Apodera, Alocodera* and *Padaungiella.* This new diversity has implications on microbial biogeography as *Apodera vas* and *Alocodera cockayni* were previously considered as two non-cosmopolite species with very broad geographical ranges and large ecological tolerances. Furthermore, we showed that the situation was far more complex, suggesting the existence of narrow local endemisms and ecological specialists, similarly to genera *Hyalosphenia* and *Nebela* (Chapter 5).

Finally, we explored the diversity patterns of the genus *Nebela* along an elevation gradient (Chapter 6). We observed a decrease of abundance and diversity in high elevation corresponding to a typical mid-domain effect. Our study also revealed several unknown phylotypes restricted to the higher elevation that seemed to present competitive exclusion with the generalist taxa from lower elevation.

In conclusion, this thesis highlights that molecular methods associated to robust morphological observations are efficient to reveal and describe the diversity of Amoebozoa. Furthermore, these micro-

bial organisms display biogeographical and macroecological patterns similarly to animals, plants and fungi, when all groups are studied at the same taxonomical rank, i.e. species level.

Résumé

La diversité mondiale des eucaryotes est dominée par des organismes (principalement) unicellulaires appelés protistes. Parmi eux, les Amoebozoa sont l'un des groupes les plus abondants, diversifiés et caractéristiques du sol, jouant ainsi des rôles importants dans le fonctionnement des écosystèmes. Cependant, leur étude a été entravée par la difficulté de les détecter et le manque de traits morphologiques stables dans la plupart des groupes. Toutefois, certains amibozoaires comme les Hyalospheniformes (Arcellinida) produisent une *thèque* (c.-à-d. une coquille) caractéristique qui facilite leur identification, et sont donc considérées comme un groupe modèle approprié pour étudier les schémas de répartition de la diversité. Le développement récent du barcoding moléculaire a considérablement aidé pour l'identification taxonomique, tandis que le métabarcoding a permis de révéler la composition des communautés microbiennes sans biais d'observation et de culture. Ces méthodes se sont révélées efficaces pour plusieurs groupes microbiens, mais seulement quelques études ont été conçues pour les Amoebozoa et les protocoles disponibles sont encore assez rares. Les objectifs de ma thèse étaient alors 1) améliorer et développer des méthodes moléculaires pour étudier la diversité et l'écologie des amibozoaires, 2) estimer la diversité taxonomique et fonctionnelle présente dans le sol, 3) améliorer la taxonomie et phylogénie de cette diversité afin d'établir une base solide pour de futures recherches et 4) caractériser les facteurs écologiques susceptibles d'influencer la diversité microbienne à l'échelle locale, continentale et mondiale.

Nous avons d'abord conçu un nouveau marqueur moléculaire pour étudier plusieurs groupes d'arcellinides, qui s'est révélé efficace pour discriminer des taxons proches et étudier simultanément les relations phylogénétiques profondes entre des taxons éloignés (chapitre 2). De plus, nous avons également adapté un protocole de métabarcoding pour étudier le genre *Nebela* avec des amorces COI spécifiques et une résolution taxonomique fine (chapitre 6).

Ensuite, nous avons isolé, cultivé et décrit le premier membre d'un clade environnemental d'amibozoaires évolutivement très divergent (chapitre 3). Cette amibe, l'une des plus petites espèces d'amibes décrites, présente un cycle de vie unique avec une alternance de trophozoïtes actifs phagotrophes et de ramifications osmotrophes ressemblant aux champignons. Sa présence a été fréquemment reportée dans de nombreuses études de métabarcoding du sol, mais cet organisme n'avait jamais été caractérisé auparavant. En revanche, les Hyalospheniformes sont connus depuis les travaux d'Ehrenberg au XIXe siècle. Cependant, leur diversité au niveau de l'espèce reste mal caractérisée. Dans le chapitre 4, nous avons montré que l'espèce emblématique d'amibe à thèque, *Nebela militaris*, n'appartenait pas au genre *Nebela*, mais constituait une entité distincte dans l'arbre des Hyalospheniformes. Par conséquent, nous avons érigé le nouveau genre *Alabasta* pour cette espèce (chapitre 4).

De plus, nous avons montré que la diversité des Hyalospheniformes avait été largement sous-estimée. En effet, nos résultats morphologiques et moléculaires ont révélé la présence de plusieurs espèces au sein des genres *Apodera, Alocodera* et *Padaungiella.* Cette nouvelle diversité a des impacts sur la biogéographie microbienne, car *Apodera vas* et *Alocodera cockayni* étaient auparavant considérées comme deux espèces non-cosmopolites avec des aires de répartition géographique très étendues et de grandes tolérances écologiques. Par conséquent, nous avons montré que la situation était beaucoup plus complexe, suggérant l'existence d'endémismes locaux étroits et de spécialistes écologiques, à l'instar des genres *Hyalosphenia* et *Nebela* (chapitre 5).

Finalement, nous avons exploré la diversité du genre *Nebela* le long d'un gradient d'élévation (chapitre 6). Nous avons observé une diminution de l'abondance et de la diversité en haute altitude ce

qui correspond à un effet typique de « milieu de domaine ». Notre étude a également révélé plusieurs phylotypes inconnus limités à de hautes altitudes qui semblent présenter une exclusion réciproque avec des taxons généralistes présents à des altitudes inférieures.

En conclusion, cette thèse met en évidence que des méthodes moléculaires associées à des observations morphologiques robustes sont efficaces pour révéler et décrire la diversité des Amoebozoa. De plus, ces organismes microbiens possèdent des schémas biogéographiques et macro-écologiques similaires aux animaux, plantes et champignons, dès lors que ces groupes sont étudiés au même rang taxonomique, c'est-à-dire au niveau de l'espèce.

Acknowledgements/Remerciements

La thèse de doctorat représente une quantité de travail gigantesque, incluant naturellement la rédaction, mais également l'acquisition de nombreuses données ainsi que leurs analyses et interprétations. Heureusement, j'ai eu la chance de pouvoir compter sur le soutien et l'aide de nombreuses personnes compétentes et disponibles que je tiens à remercier du fond du cœur, cette thèse est également la vôtre!

J'aimerais commencer ces remerciements par mon superviseur de thèse, Enrique Lara, qui m'a partagé sa passion pour les protistes dès mon bachelor et qui m'a formé depuis en tant que chercheur. J'ai eu énormément de plaisir à apprendre et travailler à tes côtés, mais également durant nos activités en dehors du monde scientifique. J'espère avoir encore d'autres occasions de collaborer avec toi dans le futur et je suis sûr qu'on pourra garder contact ! Je remercie également mon co-superviseur, Edward Mitchell, pour ses précieuses qualités scientifiques et humaines ainsi que sa grande disponibilité durant toutes ces années. Tu as su faire de ton laboratoire un cadre de travail épanouissant qui va beaucoup me manquer à l'avenir !

I would like to express my gratitude to the members of my jury, Carlos Lado, Diego Fontaneto and Guillaume Lentendu, who accepted to read and evaluate my thesis.

Durant ma thèse, j'ai eu la chance d'être entouré d'une « Dream Team » incroyable. J'aimerais tout d'abord remercier David qui a joué son rôle de grand frère à la perfection, toujours présent pour m'aider dans tous les domaines de la science avec souvent de très bons conseils. Ensuite Clément (Von Duky), mon éternel binôme avec qui j'ai pu partager tous les bons moments ainsi que toutes les galères de ma thèse. Nos supers Schtroumpfettes Amandine, Molly, Coralie ainsi que toutes les « petites mains » qui m'ont largement aidées à l'obtention de mes résultats, malheureusement généralement très récalcitrants. Mes remerciements vont ensuite à ma geek team Chris, notre super Mulot, David et Duke, qui m'ont énormément aidé lors de mon apprentissage des analyses bioinformatiques et statistiques, j'aurais sûrement balancé mon PC plusieurs fois par la fenêtre sans votre soutien. Isa pour ses précieuses discussions et son aide en particulier pour LaTeX. I am also thankful to Luciana my wonderful argantinian lab mate for all the great moment spent together, I wish you could have come earlier ! Je remercie également Estelle, Olivia, Julie, Maj, Guillaume, Anush, Joris, Fanny, Raph ainsi que toutes les personnes qui ont fait partie du laboratoire de biodiversité du sol pour leur soutien et leur extrême gentillesse, cette bonne ambiance et franche camaraderie, au labo comme en dehors, a certainement été la clé du succès de cette thèse !

Je tiens également à remercier le Fond National Suisse de la Recherche Scientifique pour leur soutien financier (SNF: 31003A_163254) ainsi que l'université de Neuchâtel pour le soutien technique et financier. Un grand merci en particulier à Delphine Lab qui a toujours pu trouver les références dont j'avais besoin, Christiane et Agnès pour leur bienveillance à la cafétéria, les secrétaires de l'institut de biologie et de la faculté des sciences pour leur aide précieuse concernant les questions administratives, et finalement Mostafa et Werner pour leur sympathie et leur disponibilité.

I am also grateful to all my colleagues from other laboratories with whom I collaborated or that I met in congresses. In particular, Daniel Lahr and his students Alfredo Porfirio de Sousa, Giulia Magri Ribeiro and Paulo Hofstatter for they friendly welcome during my stay in the Laboratory of Evolutionary Protistology, and also Stefan Geisen, Andrew Macumber, Antoine Guisan and his team, Mariusz Lamentowicz and his students Monika Reczuga, Anna Basinska, Katarzyna Marcisz, Katarzyna Kajukalo and Dominika Łuców, Kenneth Dumack, Maximilian Ganzer, Michael Gruber, Alexandra Rybarski, Alexandra Schoenle, Valentyna Krashevska, Milcho

Todorov and all my colleagues from DGP and ISTA. You are all amazing scientist and people and I really appreciate to meet and interact with you! Also a special thanks to Ferry Siemensma and his wonderful updated website about amoeboid organisms (https://www.arcella.nl/), which helped me to grasp the protists diversity. I am very grateful to all the people who sent me samples from various localities, which were included, or not in this thesis!

En plus de mon labo, j'ai également eu la chance d'avoir beaucoup d'amis au sein de l'université que je remercie pour tous les bons moments d'échange et les nombreux apéros ! Notamment l'ensemble du LEF et en particulier Claire, Adrienne, Pili, Sarah, Manu et Alan, Anaële, Wafa, Coco, Aislinn, mes deux bouchemousses préférées Mélissa et Alice, Vinciane, Stella, Emilie, Jonas ainsi que tous les autres! Vous avez tous été super !

J'en arrive à tous mes amis de longue date en dehors de l'université qui m'ont permis de prendre de la distance avec la recherche quand j'en avais besoin. Je remercie donc ma team des bioleux Damien, David, Zazou, Maellechou, La Cata, Gianni et Chichi pour toutes les soirées passées en votre super compagnie. Mes Sabines préférées (Valou, Loutre et Scuty) pour tous les paléos épiques, les escapes room d'anthologies et autres chouettes moments passés avec vous. Benj pour ces plus de 20 ans d'incroyable amitié ainsi que Noémie, Corinne et Jan pour les soirées d'organisation de mariage et autres superbes balades. Tous mes colocs pour les bons moments de détentes après le travail. Sarah et Oriane pour les voyages et concerts de folie. Muzy, Marta et Ahmed pour nos sympathiques discussions autour d'une bière ou d'un repas indien. Merci à tous mes autres amis pour les bons moments et fêtes passées avec vous. Toutes les personnes du ColombierVolley et en particulier mes nombreux coéquipiers pour tout le temps passé en salle avec vous à taper dans des ballons et à me défouler, que de beaux moments vécus! Un merci spécial à Alexandre Astier, Yann Marguet et toutes les autres personnes qui m'ont fait rire au quotidien lors de nos traditionnelles pauses Kaamelott et Cie.

Finalement, j'aimerais exprimer ma plus profonde gratitude à ma famille qui m'a encouragé et soutenu tout au long de mes études. Ma famille d'Auvernier : Patricia, Lucie, Elie et Nicolas. Ma famille de Savagnier : Vincent, Paty, Boris, Océane et Léa. Et ma famille de Fenin avec mes beaux-parents Patrick et Françoise ainsi que Laeticia. Je remercie également mes grands-parents, tante, oncles, cousines et cousin d'avoir toujours été si présent et fiers de moi sans vraiment comprendre ce que j'étudiais. Mon dernier remerciement est pour Aurélie et son soutien inconditionnel au quotidien en particulier dans les moments durs de nos fins de thèse. Je me réjouis de partir à l'aventure découvrir un nouveau pays à tes côtés durant mon futur postdoc ! Je t'aime <3

Chapter 1

Introduction

1.1 Diversity

1.1.1 What are micro-eukaryotes and protists?

The eukaryotes are all organisms that possess a nucleus compartmented by endomembrane systems, in contrast to prokaryotes (i.e. bacteria and archaea) which have no organelles. Most known eukaryotic groups are macroscopic, i.e. plants, animals and many fungi as we can easily observe them and interact with them. Nevertheless, it is now well accepted that these three groups represent only a minor fraction of the total eukaryotic diversity. Indeed, since the discovery of the first micro-eukaryote in the 17th century by Van Leewenhoeck (1676) the interest for microbial diversity has gradually increased in the scientific community, especially around the end of the 20th century following the development of molecular methods. The retrieval of molecular data from a wide panel of microbial-sized eukaryotes has evidenced that macroscopical groups were interspersed in the tree of eukaryotes among unicellular forms, which form the bulk of eukaryotic diversity (Fig 1.1) (Simpson and Roger, 2004). This microbial diversity is referred to as protists, which include all eukaryotic organisms that do not belong to plants, animals or fungi.

Protists display diverse and important functions in all ecosystems. For example, heterotrophic forms control bacterial (Azam et al., 1983; Clarholm, 2005), but also eukaryotic microbial populations (Geisen et al., 2016, 2015c). In addition, marine phototrophic protists are responsible of half of the global carbon intake (Field et al., 1998), whereas soils protists probably also play a non-negligible role (Seppey et al., 2017; Zancan et al., 2006). Parasitic forms are very diversified and abundant in all systems, and most likely play fundamental roles in regulating marine plankton communities (de Vargas et al., 2015) and soil invertebrates (Mahé et al., 2017). This high diversity of functions and environments is associated with an important morphological variability. Indeed, protists from all clades of the eukaryotic tree display a wide variety of flagella, cilia, and may have more or less amoeboid shapes with pseudopodia from all sizes and shapes. Some protists even produce complex mineral skeletons of varying compositions (e.g. Acantharia, Polycystinea and Foraminifera).

Eukaryotes are currently divided into several monophyletic supergroups with common characteristics, such as Rhizaria and Opisthokonta. These supergroups are then organized into an eukaryotic tree of life (Fig 1.1) (e.g. (Keeling et al., 2005; Pawlowski et al., 2012; Simpson and Roger, 2004)) which is still in progress (Burki et al., 2019). Despite being very diverse and playing essential functions in all ecosystems, protists are comparatively less studied and understood than macroscopical eukaryotes or other microorganisms such as fungi and bacteria. Therefore, several important fundamental and applied questions still await to be answered (see Appendix E (Geisen et al., 2017)). In particular, terrestrial free-living organisms of non-economical values are particularly understudied groups (del Campo et al., 2014).

1.1.2 What are Amoebozoa?

The Amoebozoa (Lühe 1913), represent one of the oldest branching supergroup (Simpson and Roger, 2004) and are considered as the most diverse amoeboid supergroup together with the Rhizaria (Pawlowski, 2008; Pawlowski and Burki, 2009). This supergroup is considered as one of the most dominant protist lineages in several ecosystems as recently reported from a beech leaf litter and agricultural soil surveys (Santos et al., 2020; Voss et al., 2019). In the Amoebozoa, about 5000 morphospecies were described (Lahr et al., 2011c), which is most likely an underestimation of the real diversity as these organisms usually display few morphological characters for idendification (Bovee, 1953; Smirnov, 2008; Smirnov and Goodkov, 1999). All amoebozoans are characterized by a flexible shape associated with cytoplasmic projections called pseudopodia (Smirnov, 2008) and with a phagocytosis feeding strategy (Avery et al., 1995). Some taxa possess flagella, like for instance the Myxomycetes and the genera *Mastigamoeba*, *Phalansterium* and *Multicilia* all regrouped within the Evosea (Kang et al., 2017; Lauterborn, 1895). Most amoebozoans are active heterotrophs in soils, e.g. *Acanthamoeba* and *Vermamoeba* (Clarholm, 1981; Ekelund and Rønn, 1994), or in freshwater, e.g. Dermamoebida (Smirnov et al., 2011a), but also in marine environments e.g. Dactylopodida (Volkova and Kudryavtsev, 2017). Furthermore, some amoebozoans are also osmotrophs, at least during a certain period of their life cycle, such as *Acanthamoeba* and the Myxomycetes (Adl and Gupta, 2006; Neff, 1957). Under specific conditions, some groups can become human parasites, e.g. *Acanthamoeba* and *Entamoeba* (Qvarnstrom et al., 2009; Visvesvara et al., 2007), and are then considered as opportunistic parasites. Some organisms associated with photosym-

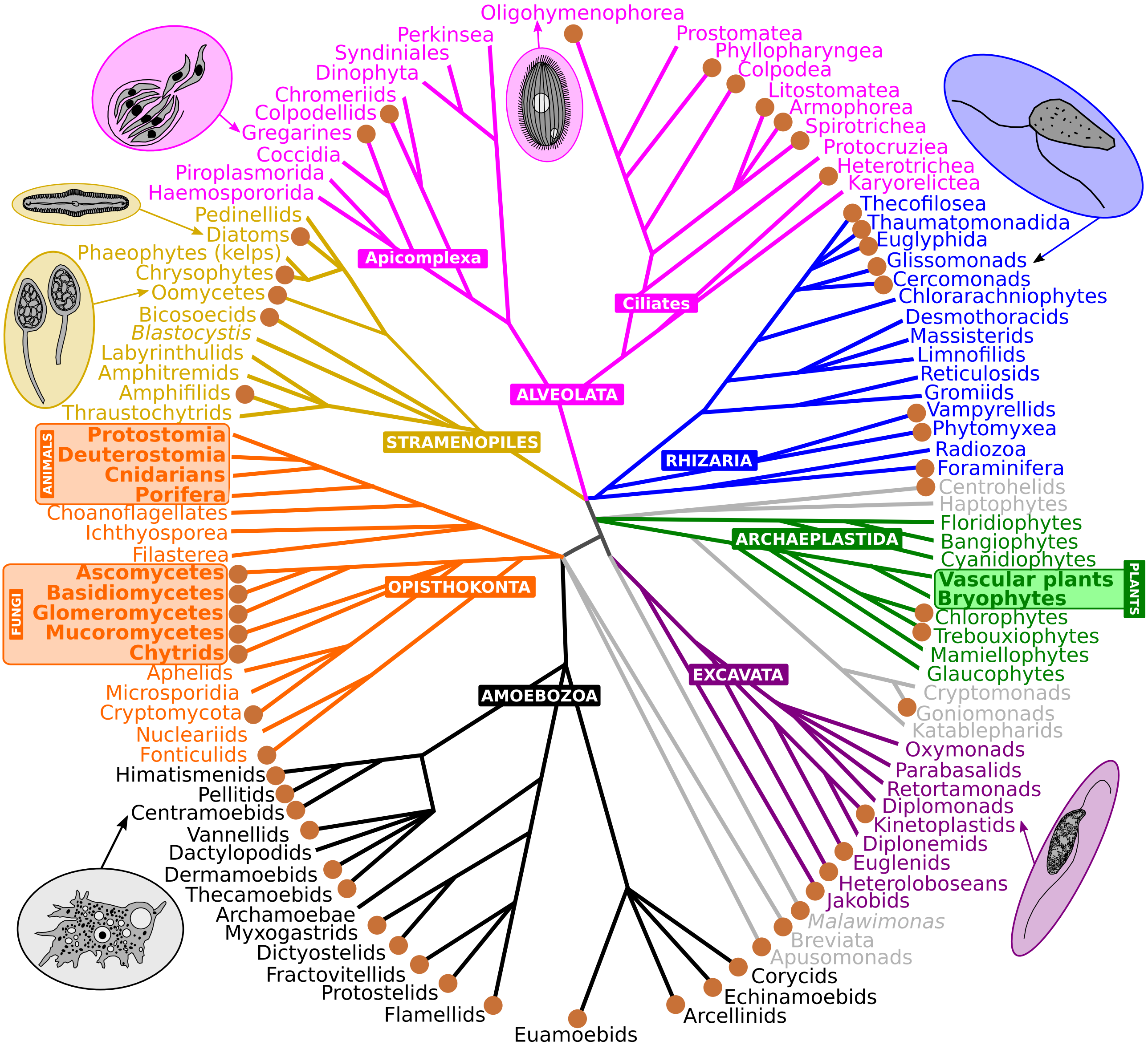

Figure 1.1 – Schematic representation of the major lineages in the eukaryotic tree of life from Geisen et al. (2018). Multicellular non-protist taxa are highlighted in boxes, whereas brown dots represent abundant and diverse terrestrial groups.

bionts can take advantage from the photosynthesis by-products, and behave as mixotrophs e.g. *Hyalosphenia papilio* (Gomaa et al., 2014) (Fig 1.2). Amoebozoa are capable of colonizing any habitat on Earth that can host eukaryotic life and have been found in extreme environments such as anoxic habitats (*Pelomyxa palustris*, *Mastigamoeba*) (Margulis et al., 1998) or hydrothermal springs (*Echinamoeba thermarum*) (Baumgartner et al., 2003).

An increasing number of Amoebozoa is found to display particular and complex life cycles including the formation of fructification structures that allow the dispersion of propagules (Shadwick et al., 2009). The ontogeny of these dispersal-related structures varies from an aggregation of cells, such as in Dictyostelids (Raper, 1984), to structures formed from a single cell ("protosteloids") (Shadwick et al., 2009). Sex, in the sense of genomic recombination, often occurs within these life cycles. Indeed, although amoebozoans such as *Amoeba proteus* were traditionally considered as mainly asexual organisms (Hurst et al., 1992), at least under cultural conditions, several taxa (e.g. Myxomycetes, *Dictyostelium* and *Cochliopodium*) are known to present evidences of sex (i.e. meiosis and karyogamy) (Lahr et al., 2011c; Wood et al., 2017). Hofstatter et al. (2018) reported the whole set of meiosis related genes, using transcriptomics data obtain from a wide sampling of amoebozoan, suggesting the widespread existence of sex-related processes in this group. Hence, these authors proposed that all amoebozoans should be considered as sexual unless proven otherwise. However, other studies argued that protists are rather mainly asexual and that meiosis-specific genes are used in other processes, including polyploidy control and homologous recombination, independent of meiosis (Maciver, 2019). New research based on a combination of cultivation and molecular characterization (see Methods section of this chapter) is constantly revealing new species and clades of Amoebozoa, and in the same time, describing uncharacterized life cycles ((Kudryavtsev et al., 2019; Tice et al., 2016) and Chapter 3 (Blandenier et al., 2017b)).

1.1.3 What are Arcellinida?

The Arcellinida (Kent, 1880), one particular order of Amoebozoa, are characterized by a self-constructed *test* (i.e. a shell) and belong to the polyphyletic group of testate amoebae, together with the Euglyphida and some other Cercozoan clades (Imbricatea) (Cavalier-Smith and Chao, 2003) and the Amphitremida (Poche, 1913) (Stramenopiles) (Adl et al., 2019). The arcellinid test is made of a proteinaceous matrix (i.e. cement) (e.g. *Hyalosphenia* and *Arcella*). In addition, this matrix is generally reinforced either by 1) self-secreted (i.e. idiosomes) siliceous plates (e.g. *Quadrulella*) siliceous rods (e.g. *Lesquereusia*) or calcareous plates (e.g. *Paraquadrula*), 2) foreign agglutinated particles (i.e. xenosomes) (e.g. *Difflugia*), and finally, 3) agglutinated plates taken from preys (e.g. *Nebela*) (Lahr et al., 2015a; Meisterfeld, 2002). The morphology, size and structure of the test are the most common criteria for genus and species determination (Tsyganov et al., 2016) (Fig 1.2), but the number, size and shape of nuclei (Raikov, 1982; Penard, 1902) or pseudopodia (Bonnet, 1961) are also very useful traits, especially for deep phylogeny (Dumack et al., 2019).

Arcellinida are an old group of organisms which originate during the Neoproterozoic (Morais et al., 2017; Porter and Knoll, 2000; Riedman et al., 2018). Indeed, this hypothesis has been supported by phylogenomic data (Lahr et al., 2019). Arcellinida are common in all environments such as lakes (e.g. *Difflugia*), mosses (e.g. *Arcella*), peatlands (e.g. *Hyalosphenia*) and soil (e.g. *Cyclopyxis*). These protists act as top predators in microbial foodwebs (Gilbert et al., 2000; Jassey et al., 2013), preying on a wide array of organisms such as bacteria, algae, fungi (Couteaux, 1985), protists including testate amoebae (Chardez, 1985; Deflandre, 1936) and even small metazoans (Geisen et al., 2015c). Some species are also decomposers of humic detritus (Bonnet, 1964) and obligatory mixotrophs ((Gomaa et al., 2014) see Appendix A (Macumber et al., 2019)). In contrast to other protists, Arcellinida have a long generation times varying from days to weeks (Schönborn, 1992), and the majority of these species displays narrow ecological tolerances. For example, communities react to changes in the water table depth (WTD) in a peatbog within days to weeks (Charman and Warner, 1997; Mitchell et al., 2008).

A key characteristic of testate amoebae is that the test is generally preserved in peat and sediments for many years and even millennia (Warner, 1990). Indeed, Arcellinida subfossils in peat are currently used in palaeoecological studies in order to reconstruct or infer the past environmental conditions based on local current species composition. These studies have allowed to highlight climatic fluctuations until the industrial era (Lamentowicz et al., 2010), or even in the Bronze Age (Amesbury et al., 2008), and also bioindication of past and current pollution events (Patterson et al., 2012; Payne et al., 2012). These researches, however, advocated for an improved taxonomical frame that sets phylogenetic relationships between clades and also divides the observed diversity into biologically-relevant species. Molecular phylogenies have permitted to split two large "wastebin" genera, i.e. *Nebela* (Kosakyan et al., 2012) and *Difflugia* (Gomaa et al., 2012) see Appendix A (Macumber et al., 2019), into relevant monophyletic units. Finally, a recent study based on multigene phylogenies lead to the reconstruction of deep phylogenetic relationships within Arcellinida (Lahr et al., 2019).

1.1.4 What are Hyalospheniformes?

The Hyalospheniidae family (Schultze, 1877) was recently raised to the infraorder rank Hyalospheniformes Lahr et al. (2019). They represent one of the largest clades of Arcellinida in terms of number of described species and their abundance in the environment. Like most other arcellinids, the Hyalospheniformes display a rigid test, largely used for the species identification, which is characterized by a flask-vase shape, oval or piriform, usually dorso-ventrally compressed. In addition, an organic rim surrounding the aperture has been suggested as the synapomorphy of the group (Kosakyan et al., 2012, 2016b). The Hyalospheniformes include some of the first described and most iconic testate amoeba species, e.g. *Nebela collaris* (Ehrenberg, 1848), *Hyalosphenia papilio* (Leidy, 1879) and *Apodera vas* (Certes, 1889). Most species were described between the end of 19th and the first half of

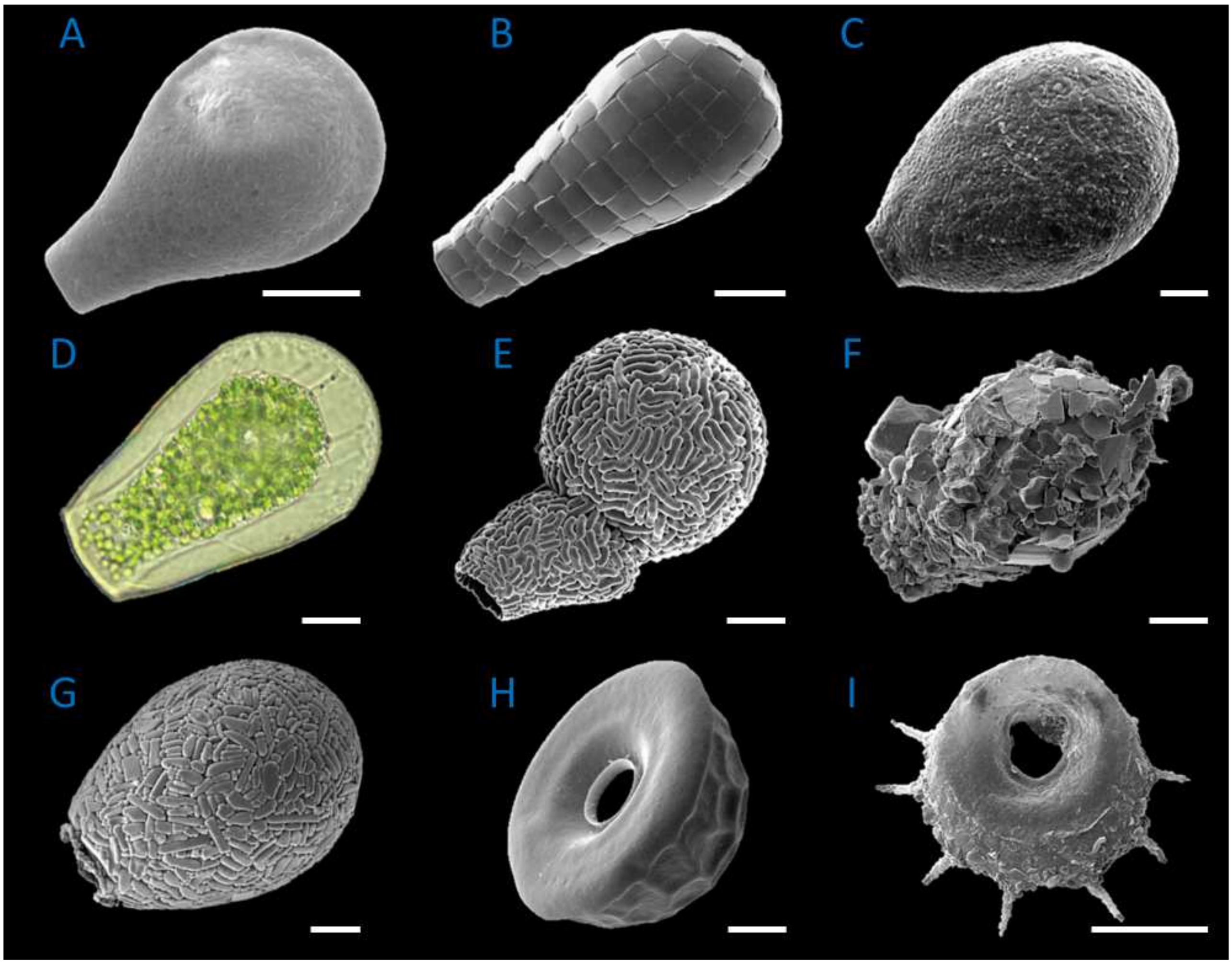

Figure 1.2 – Scanning electron and light micrographs of several genera of Arcellinida illustrating the diversity of shape and the variability of the test composition and organization. **A.** *Gibbocarina* sp.; **B.** *Quadrulella longicollis*; **C.** *Nebela collaris*; **D.** *Hyalosphenia papilio*; **E.** *Lesquereusia epistomium*; **F.** *Difflugia elegans*; **G.** *Netzelia oviformis*; **H.** *Arcella gibbosa*; **I.** *Centropyxis spinosa*. Image **A** by Q. Blandenier and images **B-I** modified from Todorov and Bankov (2019). Scale bars: **A,I** = 50 µm, **B-H** = 20 µm.

20th centuries in important faunistic records from different regions and environments, e.g. North America fresh-water (Leidy, 1879), Lake Geneva in Switzerland (Penard, 1890, 1902), soil and mosses from Congo (Decloître, 1965; Gauthier-Liévre, 1957) and New Zealand (Van Oye, 1956). These description were then reviewed in particular for genera *Nebela* and *Hyalosphenia* by Deflandre (1936) and Grospietsch (1965). In addition, the systematic of Hyalospheniformes has been frequently improved with several important works such as Jung (1942); Schultze (1877); Taranek (1882). The more recent taxonomic reviews, focusing on Arcellinida, that include Hyalospheniformes are from (Meisterfeld, 2002; Tsyganov et al., 2016) but none of them present molecular phylogeny or barcoded taxa.

The Hyalospheniformes are one of the dominant groups in peatlands (Amesbury et al., 2018; Gilbert and Mitchell, 2006; Koenig et al., 2018) but are also abundant in forest litter (Bobrov and Kosakyan, 2015; Todorov, 2002) as well as in other wet soil systems with high amounts of humic acids (Bonnet, 1964). To our knowledge, only one strictly aquatic species have been described, i.e. *Hyalosphenia punctata* (Penard, 1891, 1899). On the contrary, one species (*Quadrulella texcalense*) is apparently restricted to biocrusts in a Mexican tropical desert where the main humidity is assumed to be related to specific dry-resistant mosses belonging to the family Pottiaceae (Pérez-Juárez et al., 2017). The Hyalospheniformes are chiefly predators of other eukaryotes such as fungi, algae, protists and even larger organisms including nematodes (Gilbert et al., 2000; Yeates and Foissner, 1995). Most Hyalospheniformes are known for hunting specifically smaller testate amoebae such as Euglyphida and using their silica scales to build their own tests (Lahr et al., 2015a).

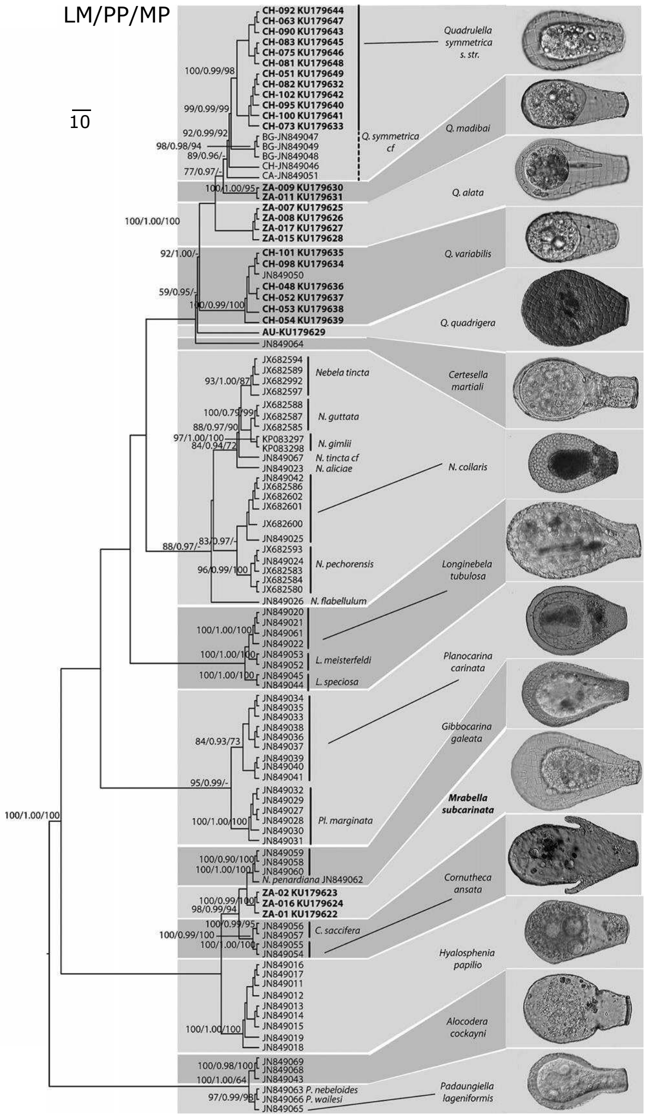

Figure 1.3 – Molecular phylogeny of the Hyalospheniformes from Kosakyan et al. (2016b) based on a COI marker with bootstrap values obtained by maximum likelihood analysis, posterior probability values obtain by Bayesian analysis and parsimony scores obtain by maximum parsimony analysis.

Before the advent of molecular methods (see Methods section), the taxonomy of Hyalospheniformes was mainly based on morphotypes defined by direct microscopic observations. Like in other arcellinids (and arguably many protists), Hyalospheniformes tests vary from one cell to another. Therefore, sorting out morphological heterogeneity between individuals (due to phenotypic plasticity or intra-clonal variability (Medioli et al., 1987; Wanner and Meisterfeld, 1994; Wanner, 1999)) and genetically encoded variation can be challenging. The combination of morphometrics and molecular studies has solved this issue at least for genus *Nebela* (Kosakyan et al., 2013), showing a wide diversity of *pseudo-cryptic* species, sometimes varying by slight but pervasively found differences in shell size and shape. These findings eventually lead to a taxonomic revision of the genus which now includes eight species described with molecular methods (Singer et al., 2015). On the contrary, some groups like the *Hyalosphenia papilio* species complex display a wide phenotypic plasticity that probably mask genetically encoded morphological variation, i.e. *cryptic* species (Mulot et al., 2017).

The application of this combination of molecular and morphological approaches resulted in improvements in the knowledge of Hyalospheniformes systematics. Indeed, genus *Nebela* has been shown to be paraphyletic (Lara et al., 2008) and split into several genera that have been erected these last years, i.e. *Padaungiella* (Kosakyan et al., 2012), *Longinebela, Planocarina, Gibbocarina, Mrabella, Cornutheca* (Kosakyan et al., 2016b) and *Alabasta* (see Chapter 4 (Blandenier et al., 2018)). In addition, several species have recently been described in the different genera, e.g. *Longinebela* (*Nebela*) *meisterfeldi* (Kosakyan et al., 2012), *Nebela guttata* and *N. gimlii* (Kosakyan et al., 2013; Singer et al., 2015), *Quadrulella madibai* and *Q. texcalense* (Kosakyan et al., 2016b; Pérez-Juárez et al., 2017) and *Cornutheca* (*Nebela*) *jiuhuensis* (Qin et al., 2016). The current phylogenies (Fig 1.3) indicate that hyalospheniids with an elongated neck, i.e. genera *Padaungiella* and *Alocodera*, are forming a sister clade to the other genera (Kosakyan et al., 2016b) and Luketa (2015) has already proposed the erection of the family Padaungiellidae for these genera. However, erecting this clade to family level has been considered as too premature due to taxonomic undersampling (Kosakyan et al., 2016b) and therefore the Hyalospheniidae (Schultze, 1877) is currently the only family belonging to the recent infraorder Hyalospheniformes and suborder Glutinoconcha (Lahr et al., 2019).

Kosakyan and co-authors (in prep) are finalizing a monograph that reviews current taxonomical, systematical and ecological knowledge of the Hyalospheniformes. The infraorder is now composed of 93 valid species distributed in 13 genera (Kosakyan et al. in prep), but this number is expected to increase largely in the following years. Indeed, there are still many potential species remaining undetected, especially in under-explored ecosystems and regions of the world (Pérez-Juárez et al., 2017). The clade formed by genera *Alocodera* and *Apodera* is thought to be distributed preferentially in the Southern Hemisphere (Smith et al., 2008), an undersampled region of the world as compared to Europe and North America (Smith and Wilkinson, 2007).

These two genera are usually considered as monospecific (Zapata and Fernandez, 2008) while some authors suggested the existence of several species (Deflandre, 1936; Penard, 1911) and even described some species, i.e. *Apodera crenata* Jung (1942) and *Apodera wellingtonia* Decloitre (1964). An investigation of the diversity of these genera shows new perspectives that challenged preconceptions about the Southern Hemisphere as a less diverse region for testate amoebae (see Chapter 5).

1.2 Ecology

1.2.1 Ecology of Hyalospheniformes

Since the beginning of the 20th century, early authors noticed that testate amoebae species have different ecological preferences, e.g. water quality, chemical composition, type of vegetation and nature of the substrate (Penard, 1902). Indeed, Penard noticed that species composition varied between some specific habitats such as *Sphagnum* mosses and "brown mosses", both characterized with many Hyalospheniformes, whereas deep lake habitats are globally deprived of them. However, he admitted that most testate amoebae species could not be associated to a specific habitat. Some years later, Deflandre (1927) argued that inferring testate amoebae species solely on the type of mosses present in the habitat (e.g. *Sphagnum*) was not sufficiently accurate to characterize the whole Hyalospheniformes community. Indeed, *Sphagnum* mosses are present both in minerotrophic and ombrotrophic wetlands (respectively refered to as fen and bogs), which present important nutrient and acidic variation. Moreover, *Sphagnum*–dominated peatlands often present contrasted microenvironments in term of water proximity, i.e. *pools* that are usually flooded, *lawns* generally wet and *hummocks* that vary from humid to dry (Rydin et al., 2013). In some cases, these ecological habitat categories were even wrong as some hyalospheniids species supposedly restricted to *Sphagnum* species were also found in others mosses e.g. *Calliergon* and *Hypnum*. Therefore, Deflandre (1927, 1936) recommended performing deeper ecological descriptions focusing rather on the water and oxygen availability over the months (i.e. WTD fluctuations), as well as calcareous content and pH level instead of the mosses type.

Following the recommendations of Deflandre, several early attempts have been made to classify Hyalospheniformes taxa into different categories of humidity (Jung, 1936; Meisterfeld, 1979; Schönborn, 1962). Additionally, the pH and trophic status of peatlands (i.e. gradient from bog to fen) were shown to drive communities composition (Mitchell et al., 1999; Tolonen et al., 1992; Warner and Chmielewski, 1992), although their role remained relatively minor as compared with the influence of hydrology (Booth, 2001). Tolonen et al. (1994) based their study on larger number of environmental variables associated with multivariate analysis. These authors proposed moisture and WTD as the most important factors for testate amoebae communities, followed by dissolved organic carbon (DOC), trophic status and finally pH. In general, it has been considered that xerophilous species possess large ecological tolerances (i.e. eurytopic), while hygrophilous species possess comparatively narrow ecological tolerances (i.e. stenotopic) (Charman and Warner, 1997; Lamentowicz and Mitchell, 2005; Mitchell et al., 1999), which make the latter more useful indicators of ecological and paleoecological conditions. Indeed *Alabasta* (*Nebela*) *militaris* is typically found in dry peatlands but also in brown mosses and in forest litter (Krashevska et al., 2018), whereas *Padaungiella* (*Nebela*) *lageniformis* is found in very diverse environments including dry humic soil and corticolous mosses (Bonnet, 1964, 1973), both are therefore poor indicators (Mitchell et al., 1999). However, because most research on testate amoeba ecology has been developed on peat bogs in the Holarctic realm, we virtually lack information on the ecology of Hyalospheniformes in other ecosystems and biogeographic areas. Some autecological works have been published on forest litter testate amoebae in Europe and Asia (Bonnet, 1964, 1990; Todorov, 2002; Bobrov and Kosakyan, 2015) but most other regions have never been studied under this perspective.

Several morphological features have been also proposed to be associated with ecological preferences. For instance, Bonnet (1964) and Heal (1963) demonstrated a strong correlation between test morphology and the humidity content on the environment with larger forms mostly found in wet habitats and smaller one in drier environments. Moreover, several functional traits (e.g. decrease of pseudostome size, decrease of biovolume, ventral or hidden apertures position and test compression) have been proposed to be adaptations to drought in order to maintain activity in a thinner water film (Fournier et al., 2015; Lamentowicz et al., 2015). These assumptions were recently confirmed by an experimental manipulation of the water table depth on peatland mesocosms (Koenig et al., 2018).

However, these early ecological studies were limited by the previous difficulties to differentiate some species, e.g. *Padaungiella* (*Nebela*) *lageniformis* and *P. wailesi* (Deflandre, 1936). Therefore it is probable that these confusions have led to some misidentification which will have blurred ecological signals (Heger et al., 2009; Mitchell and Meisterfeld, 2005). Moreover several new species were described recently (Kosakyan et al., 2016a) and there is often no detailed description to know exactly which specimens were previously investigated. Finally, current ecological interpretations are complicated by previous approximate phylogeny organization, e.g. genus *Nebela* sensu Bonnet (1990) included species now related to genera *Alabasta, Argynia, Longinebela, Padaungiella, Physochila* and *Porosia*, some of them not even included in Hyalospheniformes (Kosakyan et al., 2016a). Since the recent development of molecular methods (see Methods section) and the resulting capacity to identify Hyalospheniformes specimens to species level (Kosakyan et al., 2016a) more robust ecological patterns have been demonstrated. Indeed, it has been evidenced that the community of *Nebela* is mainly correlated to nitrogen variable with some species such as *N. tincta* and *N. guttata*, restricted to the most oligotrophic micro-environments in peatland (i.e. lawns and hummocks), whereas *N. collaris* was rather found in the more nutrient riche areas (i.e. fen) (Singer et al., 2018). These first insights seem to indicate that some Hyalospheniformes at least express more complex ecological distribution patterns than previously expected without molecular methods. Many studies are now needed to describe the fine ecology of the new diversity detected in other genera and species of this infraorder (see Diversity section) and in other environments (e.g. forest litter). For instance, to test if ecological distribution observed with elevation or latitudinal gradients surveys of macroscopic taxa also applied in microbial organisms (see Chapter 6 and Appendix C (Seppey et al., 2019). Finally, a particular care will now be needed for the identification of morphospecies from different regions as we are now aware of the problematics of (pseudo)cryptic diversity patterns in some taxa (see Chapter 5).

1.2.2 Biogeography of Hyalospheniformes

The geographic distribution of microbial-sized organisms across the globe has always been a central question for environmental microbiologists and for biogeographers alike. Darwin, in his voyage on the *Beagle* collected diatoms in what is now Tierra del Fuego and sent them to Ehrenberg for identification. Ehrenberg found forms that were already described from other regions of the globe, and Darwin concluded, in substance, that these old forms had the time to disperse everywhere (Darwin, 1846; O'Malley, 2009). Penard, one of the forefathers of testate amoeba research, was convinced that because amoebozoan where very small and able to encyst, they could be dispersed passively through winds global circulation (Penard, 1902). In his perception of protist distribution, similar habitats everywhere on the globe will possess the same communities (a view he latter modified (Penard, 1938)), whereas in different habitats of a same region, testate amoebae communities will differ greatly. Some years later, Deflandre (1928) argued that it was difficult to observed restricted distribution because of the lack of comprehensive survey on some regions (e.g. Arctic), and also because some rare taxa may have been overlooked. Deflandre (1936) showed nevertheless that some frequent Hyalospheniformes species in the Southern Hemisphere had never been observed in Europe or North America, e.g. *Apodera vas, Alocodera cockayni* and *Certesella certesi* (all of them previously in genus *Nebela*), whereas others were observed in every latitude as long as the environment was suitable, e.g. *Padaungiella* (*Nebela*) *lageniformis.* On the following years the existence of species with geographically limited distributions was confirmed with the description of many species from undersampled parts of the world such as Africa and New Zealand (Gauthier-Liévre, 1953; Van Oye, 1956). Testate amoebae were then globally assumed to be composed of two distinctive associations with specimens from Laurasia (i.e. North America, Europe and Asia) and Gondwanaland (I.e. South America, Africa, Australia, New Zealand and Antarctica) origins, with Central America and Indonesia as transitional zones including varieties common to both groups (Hoogenraad and De Groot, 1979).

A sudden regain of interest on microbial biogeography occurred during the late 1990's-early 2000's with a debate that crystallized around the sentence "everything is everywhere, but, the environment selects". This tenet took its roots in early 20^{th} century (Baas Becking, 1934) but was later made popular by authors like Tom Fenchel and Bland J. Finlay (Fenchel et al., 1997; Finlay and Fenchel, 1999; Finlay, 2002). They stated that protists were present everywhere because of their huge populations and small size, which could induce great dispersal capabilities, but that only groups with favorable environmental conditions were act-

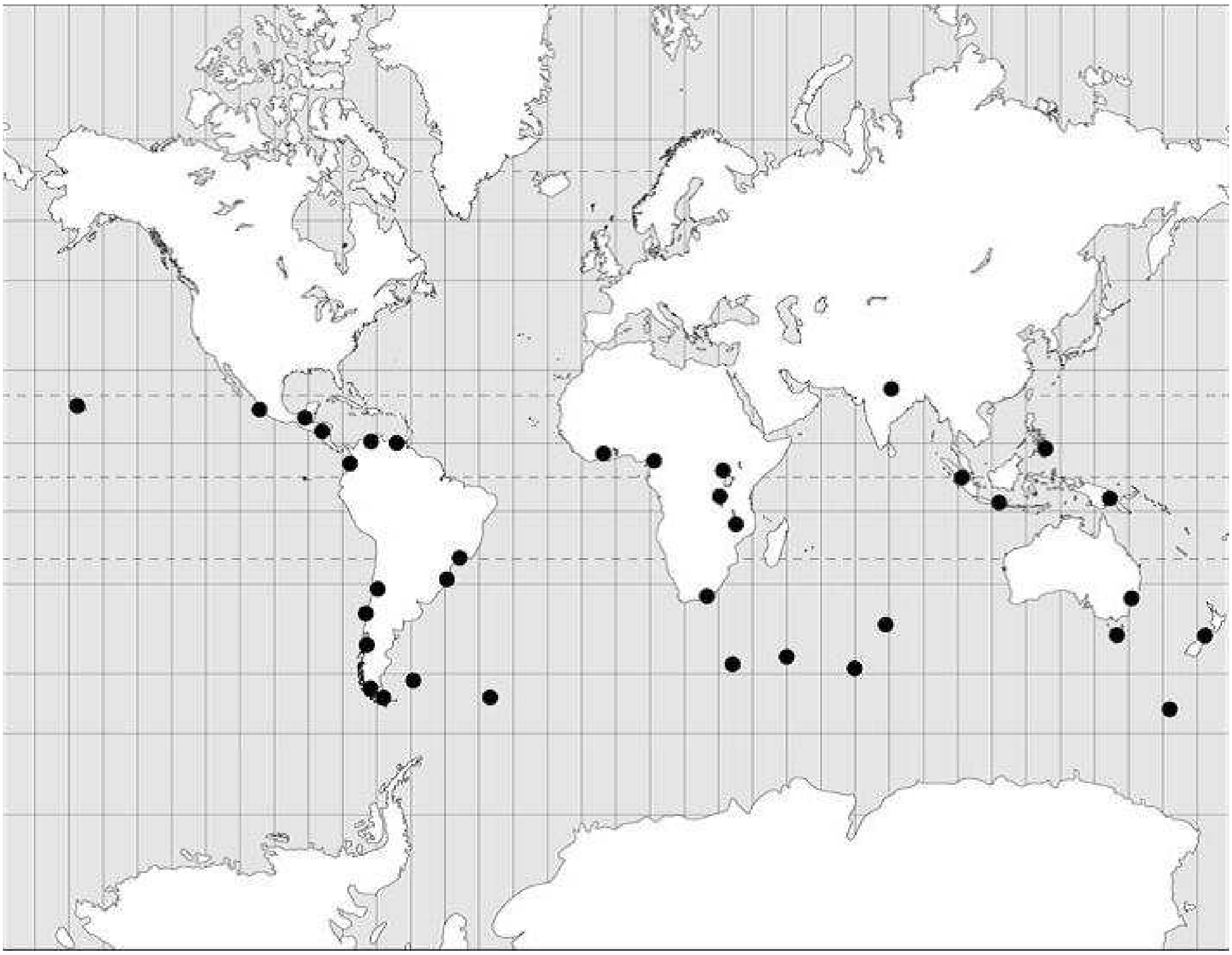

Figure 1.4 – Map of the recorded global distribution of *Apodera* (*Nebela*) *vas* from Smith and Wilkinson (2007).

ive and dominant, whereas the rest were maintained encysted and remained thus undetected (Fenchel et al., 1997). On the other hand, several protists have been shown to possess limited distribution and follow biogeographical patterns similar to animals and plants (Foissner, 1999). Therefore, Foissner (2006) proposed a *moderate endemism* concept based on a selection of *flagships* taxa, i.e. unmistakable morphological taxa (Foissner, 2004), that were not observed on similarly adequate environments in other regions of the world. Foissner included *Apodera vas*, a testate amoebae from former Gondwana origin (Deflandre, 1936), in his flagships list of endemic species based on its absence from Holarctic area despites intensive research (Smith and Wilkinson, 2007) (Fig 1.4). Further records of *A. vas* distribution confirmed Foissner argument and this species has now been frequently found in South America: e.g. in Brazil, Chile and Ecuador (Krashevska et al., 2007; Wailes, 1913), Central America: e.g. Costa Rica, Guadeloupe, Guatemala, Mexico, Nicaragua and Panama (Bonnet, 1977a; Heger et al., 2011c; Laminger, 1973), Africa: e.g. Congo, Madagascar and Tanzania (Mitchell and Meisterfeld, 2005; Todorov and Golemansky, 2014; Van Oye, 1956), Asia: e.g. Java and Nepal (Bonnet, 1977b; Van Oye, 1949), Oceania: e.g. Australia and New Zealand (Meisterfeld and Tan, 1998; Penard, 1911) and even in Kerguelen islands in Antarctica (Bonnet, 1981) (see Kosakyan Monograph (in prep) for an exhaustive records list). In contrast to Gondwanian Hyalospheniformes species (i.e. *A. vas, Alocodera cockayni* and *Certesella certesi* (Smith et al., 2008)), *Hyalosphenia papilio* is frequently recorded in almost all Holarctic peatlands (Amesbury et al., 2018; Heger et al., 2013) and is systematically absent from similar Southern Hemisphere habitats (Fernández et al., 2015; Swindles et al., 2018). Finally, *Cornutheca* (*Nebela*) *ansata* is restricted to a rather narrow region in the eastern North America (Heger et al., 2011b).

However, besides these biogeographic “flagship species”, most Arcellinida species are still considered as cosmopolitan, or at least having wide distribution areas. In other organisms, the limitations of the current morphology-based taxonomy prevent the evaluation of biogeographic patterns. As an example, the morphospecies *Nebela* (now *Argynnia*) *antarctica* was described as an endemism for Amsterdam Island, an isolated land mass in the Southern Indian Ocean, based on morphometrical comparisons with the close resembling *Argynnia dentistoma* (Grospietsch, 1971). However, the differences between the two species as observed by Grospietsch did not resist a further analysis including a wider sampling (Heger et al., 2009). Taxonomical issues, sometimes generated by a natural tendency of lumping taxa can also blur existing biogeographical patterns; grouping the Southern Hemisphere species *Apodera vas* with the assumed cosmopolitan *Lagenodifflugia vas* lead to the wrong conclusion that this species is distributed everywhere. However, this conclusion was rejected as, although some old drawings could lead to confusion, both species differed morphologically and ecologically (Mitchell and Meisterfeld, 2005).

However, the application of molecular methods have revealed a more complex story for the distribution of *Hyalosphenia papilio*, which is assumed to be composed of a complex of cryptic species (Mulot et al., 2017). Indeed, the different genetic lineages have been shown to possess different ranges of distribution, some of them largely distributed throughout Holarctic, whereas other are restricted to specific regions (Heger et al., 2013). The phylogeographic patterns of *Hyalosphenia papilio* have recently been evidenced to be related to the history of *Sphagnum*-peatland expansions and contractions in response to Quaternary glaciations (see Appendix D (Singer et al., 2019). Therefore, a particular care will now be needed for the identification of morphospecies from different regions in order to avoid lumping different cryptic morphospecies together and to highlight potential hidden biogeographical patterns (see Chapter 5).

1.3 Methods

1.3.1 What are traditional methods?

Observation, isolation and cultivation of Amoebozoa

Before the discovery of DNA, studies of protists were performed by microscopical observations either by direct observations of the samples or after cultivation of samples in enriched media. The first case apply mainly to freshwater and marine samples that contains relatively few mineral particles, whereas soils samples required to be diluted in water in order to limit opacity of the samples (Anderson, 2002). Direct observations are typically well suited to large organisms, like testate amoebae (Foissner, 1987), while minute heterotrophic flagellates and naked amoebae are usually elusive due to their transparent bodies and lack of obvious distinctive traits.

As these organisms are often encysted in soil under normal conditions (Anderson, 2000), isolation and cultivation steps are a good approach to study them (Lara et al., 2007). One of the most efficient methods to isolates soil microbial organisms is the most probable number (MPN) approach (Rønn et al., 1995) which was originally developed to estimate the abundance of soil protists (Darbyshire et al., 1974). The principle of this method consist on a serial of gradually more diluted soil suspension into a suitable growth medium (Ekelund and Rønn, 1994). After a growth period of several weeks, the determination of the maximal dilution before no protists can be observed allow estimating the original abundance of cultivable soil protists. In addition, an examination with an inverted microscope may also reveal a bloom of single organism in some wells of the lowest dilution that can then be transfer into culture flasks (Cowling, 1991). Alternatively, isolations can also be more selective by isolating individual cells from enrichment cultures with a micro pipette under an inverted microscope (Ekelund and Rønn, 1994). In both case, it is challenging to determine which type of media (e.g. (Page, 1976; Prescott and James, 1955)), which food preferences (e.g. bacteria alone, algae, fungi or flagellates (Ekelund and Rønn, 1994)), and which culture conditions (e.g. temperature, presence of light) are optimal for the target species. Therefore, soil protists are often not able to grow in cultures under controlled conditions (Foissner, 1987).

Nevertheless and despite being a laborious task, obtaining pure cultures of amoeboid taxa is often mandatory to analyze their life cycle ((Brown and Silberman, 2013) and chapter 3 (Blandenier et al., 2017b)) or their autecology, e.g. feeding behavior (Dumack et al., 2016) or salinity tolerances (Kudryavtsev et al., 2019). Some species of testate amoebae like *Euglypha rotunda* or *Assulina muscorum* can be cultivated and are available in culture collection such as CCAP (https://www.ccap.ac.uk/). In Arcellinida, successful cultures have been obtained mostly for generalist aquatic species such as *Arcella* cf. *intermedia* and *Phryganella paradoxa* (Dumack et al., 2019; Porfírio-Sousa et al., 2017), but also few terrestrial species such as *Cyclopyxis kahli* and *Difflugia lucida* (Ogden and Meisterfeld, 1991; Wanner and Meisterfeld, 1994). Nevertheless, most arcellinids, including Hyalospheniformes, are considered as uncultivable (Lahr et al., 2011c), probably due to their higher trophic requirements and prey selectivity.

Amoebozoan species identification and characterization

Naked amoebozoans identifications to species level with light microscope is in most case complicated or impossible due to the lack of stable morphological traits (Bovee, 1953; Ekelund and Rønn, 1994; Smirnov, 2008). According to Smirnov (2008) the main criteria to identify a naked amoeba species are 1) the movement of locomotive and floating forms, 2) the shape of subpseudopodia, 3) the type of uroidal structures, 4) the number and morphology of nuclei, 5) the glycocalyx (i.e. cell coat) composition, and 6) the cytological morphology of organelle, e.g. mitochondria and Golgi complex. Arcellinida identification is more precise due to the presence of a rigid test allowing more morphometrical traits (Luketa, 2017a,b). In addition, the test composition and ultrastructure can also be used as taxonomical traits. Indeed the presence and shapes of self-secreted plates or rods (i.e. idiosomes) or foreign agglutinated particles (i.e. xenosomes) are indicative features allowing to discriminate different genera and are evolutionary conserved (e.g. *Quadrulella* and *Difflugia*) (Meisterfeld, 2002).

Various approaches have been used to analyze the different features previously mentioned. For instance, the development of microscopic contrast enhancement like differential interference contrast (DIC) and confocal microscopy, and the use of staining, such as Carmine Borax (Penard, 1902) highlighted the structure of the nucleus and other organelles. Transmission Electron Microscopy (TEM) has been widely used has a taxonomical tool to investigates the external coating structures and evidence different morphological scales types displayed by some naked amoeba (Schuler and Brown, 2019; Udalov, 2016) and food strategy (Smirnov et al., 2011a). TEM has also allowed the distinction of the chromatin pattern of the nucleolus (Smirnov and Goodkov, 2004; Tyml and Dyková, 2018), the rough endoplasmic reticulum structure (Volkova and Smirnov, 2016), the shape of mitochondrial cristae and the position of microtubule-organizing centers (MTOCs) (Melton III et al., 2019). Finally, scanning electron microscopy (SEM) has permitted to describe the coating of some naked amoebae, e.g. *Armaparvus languidus* (Schuler and Brown, 2019), but has mostly been used for the characterization of Arcellinida test ultrastructure. Indeed the type of cement (i.e. rough, fragile or strong) and its network structure has been shown to be a useful taxonomic and even potential phylogenetic character, at least for agglutinated taxa such as *Difflugia* (Lahr and Lopes, 2006; Ogden, 1979, 1983, 1991). Similarly the fine composition of the test, i.e. the choice of construction material, evidenced with SEM, has a direct influence on the structure and the appearance of the test (Bian et al., 2019) and adds therefore other morphological criteria for taxa differentiation (Ogden, 1980). All these methods have allowed reaching a high taxonomical potential but also to define several ecological and biogeographical patterns in diverse Amoebozoa (Anderson, 2018; Meisterfeld, 2002).

1.3.2 What are molecular methods?

Barcoding and molecular phylogeny

The emergence of molecular methods has transformed biodiversity understanding, especially for microbial organisms. Indeed, the barcoding approach, i.e. using a section of DNA from a specific gene or genes to identify taxa, is now commonly used as an efficient tool for taxonomy (Blaxter, 2004; Hebert et al., 2003). This method is especially adapted for the determination of uncultivable or limited morphological traits taxa (Bovee, 1953; Smirnov, 2008). This approach has revealed an important cryptic diversity in all Amoebozoa (Douglas et al., 2011; Kalla et al., 2011; Kosakyan et al., 2012). The main barcoding steps include 1) the isolation of either single or couple of cells from environmental sample or culture under an inverted microscope, 2) DNA extraction with a modified protocol (Blandenier et al., 2018) based on guanidine thiocyanate buffer (Chomczynski and Sacchi, 1987), 3) amplification of a portion of a selected gene using specific primers through polymerase chain reaction (PCR) and 4) sequencing of PCR product including a facultative extra cloning steps depending on the sequence quality (Fig. 1.5) (Kosakyan et al., 2012; Porfírio-Sousa et al., 2017).

Molecular phylogeny permits to illustrate relationships between taxa with statistical support. This approach allowed the placement of both Arcellinida and Hyalospheniidae in the eukaryotic tree of life (Lara et al., 2008; Nikolaev et al., 2005), but also to clarify relationship among Hyalospheniformes genera (Kosakyan et al., 2016b) (see Diversity section) and difflugids taxa (Gomaa et al., 2017). Molecular phylogeny has also improved the phylogenetic placement of many naked amoebae species to the

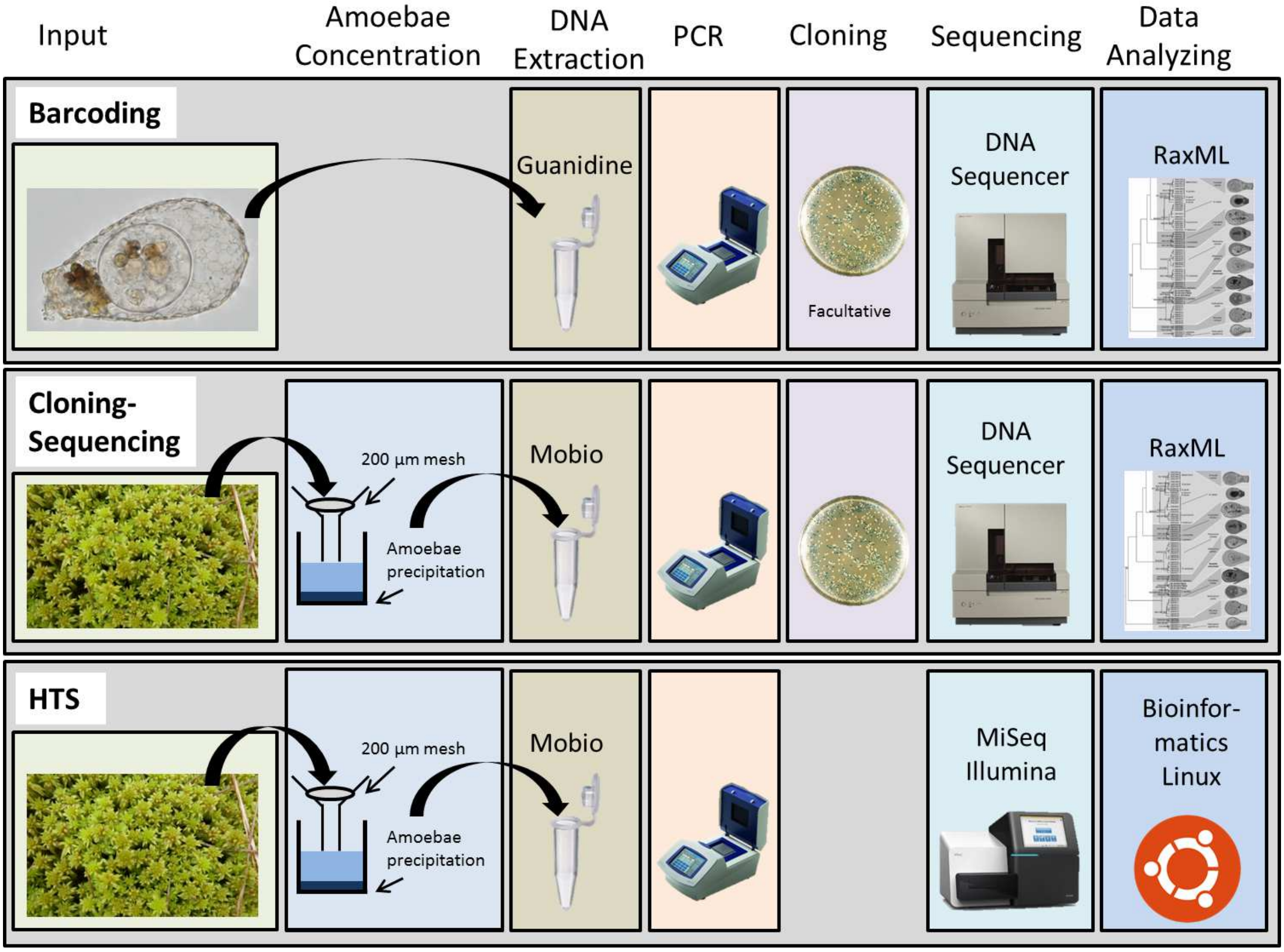

Figure 1.5 – Schematic summaries of the different steps required for the three molecular approaches used during my thesis, i.e. the barcoding, the cloning-sequencing and the High Throughput Sequencing (HTS).

correct genus, such as *Armaparvus languidus* that belong now to the genus *Pessonella* (Schuler and Brown, 2019). Phylogenetic reconstruction has also demonstrated the polyphyly of genus *Rhizamoeba* (Smirnov et al., 2017).

Metabarcoding

Metabarcoding is a common tool to evaluate environmental biodiversity using molecular methods. This approach consists in the extraction of DNA from an environmental sample (e.g. soil, water, sediments), then the amplification of environmental diversity with a genetic marker (i.e. a barcode) and finally sequencing the resulting amplicons. The first generation sequencing was originally performed by a cloning sequencing strategy with a Sanger methods (Blaxter et al., 2005), while high throughput sequencing (HTS), also referred to as next generation sequencing (NGS), is now carry out by Illumina sequencing technologies (Creer et al., 2016) (Fig. 1.5). The fast technological improvements of sequencing methods have permitted to drastically decrease the price and the time of analyses while increasing the yield and length of sequences (Gužvić, 2013), allowing the metabarcoding approach to become the gold standard of microbial communities survey. In order to analyze the huge amount of data produced (i.e. several millions of amplicon reads), several bioinformatics tools have been developed to clean the sequences and clusterize them, e.g. SWARM (Mahé et al., 2015), into operational taxonomic units (OTUs) that represent molecular biological entities (Valentini et al., 2009). These OTUs can then be assigned, e.g. with ggsearch (Pearson, 2000), to specific species with accurate databases obtained from previous barcoding studies of known taxa, e.g. Silva (Pruesse et al., 2007) and PR^2 (Guillou et al., 2012).

Metabarcoding studies have permitted to circumvent observational and cultivation biases inherent to microbial diversity, and revealed an unsuspected diversity of several environmental clades, such as LKM74 in freshwater (van Hannen et al., 1999), see also Chapter 3 (Blandenier et al., 2017b), or alveolates (Lopez-Garcia et al., 2001), stramenopiles (Massana et al., 2006) and diplonemids (Lara et al., 2009) in the oceans. The TARA oceans expedition on sunlit oceans protists confirmed the tremendous planktonic diversity (i.e. more than 150'000 OTUs), but also highlighted their heterotrophic dominant function, especially as parasites and symbionts (de Vargas et al., 2015). Finally, some applied metabarcoding studies have also been performed in order to screen the presence of bioindicators species such as Foraminifera or Diatoms (Pawlowski et al., 2016). Soils ecosystems have been comparatively less studied due the high prevalence of fungal, plant and metazoan sequences that concealed protists diversity (Lesaulnier et al., 2008). Nonetheless, several soils environmental clades were also discovered like the Opisthosporidia (Lara et al., 2010), Foraminifera (Lejzerowicz et al., 2010) and Dinophyceae (Bates et al., 2013). Similarly to oceanic ecosystem, several studies evidenced the importance of soil parasitic diversity (Bates et al., 2013; Geisen et al., 2015a; Mahé et al., 2017), but also unexpected importance of algivorous organisms in soils (Seppey et al., 2017). Metabarcoding has proved to be a successful approach for ecological studies such as elevational transects on microbial organisms (Dahl et al., 2018) and Chapter 6).

Limitations of the metabarcoding method

Due to the specificity of the primers and PCR biases, metabarcoding results are not homogeneous among groups and Amoebozoa are often under-represented or excluded (Geisen et al., 2019, 2015b; Stephenson et al., 2011; Voss et al., 2019). The main reasons are that amoebozoan are assumed to be too divergent (Fiore-Donno et al., 2010; Nikolaev et al., 2005) and to possess introns (i.e. portion of non-coding DNA present inside coding regions) (Stephenson et al., 2011). Only a few metabarcoding projects have focused on Amoebozoa with specific primers, as recommended by de Groot et al. (2016), and the target groups were only Myxomycetes (Dahl et al., 2018; Gao et al., 2019; Shchepin et al., 2019). Metabarcoding protocols for other amoebozoan groups, e.g. Arcellinida, remain therefore to be designed (Chapter 6).

In comparison with Sanger sequencing, the phylogenetic resolution obtained with metabarcoding using NGS is low due to the short length of the reads (2x300bp). Even if this size will increase gradually with technological innovation (Gužvić, 2013). In addition, an ideal marker should not contain any introns in the investigated groups of organisms, which would hamper their amplification. Primers targeting new regions and new genes are therefore requested to improve the range of amplification. PCR amplification will also induce semi-quantitative data which is a major issue for any abundance ecological study (Von Wintzingerode et al., 1997).

Another limitation of metabarcoding methods is caused by the tradeoff between *universality* and *high resolution* of primers. In other words, a deep phylogenetic marker, conserves enough resolution to retrieve sequences from all eukaryotic lineages but will be often too conserved to differentiate closely related species (Blaxter, 2004). The small subunit ribosomal RNA gene (18S rRNA) molecular marker is the best suited to obtain a good approximation of eukaryotic diversity (Mahé et al., 2017; de Vargas et al., 2015). The main reason is that the 18S rRNA gene function is extremely conserved across the tree of life, impeding rapid mutational changes. Therefore, this slow rate of evolution is problematic for the differentiation of closely related species. While some lineages have unusually fast-evolving ribosomal operons, like Foraminifera (Morard et al., 2016), it has been shown that 18S rRNA is not variable enough to resolve interspecies relationships within several amoebozoan groups like vannellids (Nassonova et al., 2010) and hyalospheniformes (Lara et al., 2008). Curiously, 18S rRNA can also form slightly divergent paralogs within single genomes, as evidenced for ciliates (Zhao et al., 2019), but also Myxomycetes (García-Martín et al., 2019) and Vannellida (Nassonova et al., 2010). This phenomenon may biased interpretations of environmental diversity in metabarcoding studies and rigorous investigations are needed to disentangle meaningful diversity from intra-individual polymorphism and methodological artefact (Decelle et al., 2014).

Other markers have been developed and applied in barcoding or metabarcoding studies (Pawlowski et al., 2012) in order to obtain a higher taxonomic resolution. For instance, the ribosomal internal transcribed spacers (ITS) is the most accurate fungal barcode (Schoch et al., 2012) and has been used in different protist lineages such as chlorarachniophytes (Gile et al., 2010) and oomycetes (Robideau et al., 2011). However, the use of ITS for barcoding Amoebozoa has been impeded by important paralogs in Vannellida (Nassonova et al., 2010), similarly to some plants and animals studies (Chase et al., 2005; Blaxter, 2004). Actin, a protein-coding gene has also been used for DNA taxonomy of Dactylopodida (Kudryavtsev and Pawlowski, 2015), but mostly as phylogenetic marker of Amoebozoa (Fahrni et al., 2003; Kudryavtsev et al., 2009). In both cases numerous paralogs were found which prevent their use as optimal taxonomic markers, at least for Arcellinida (Lahr et al., 2011b). In addition, mitochondrial genes have been expected to be promising alternatives to nuclear markers because they generally do not produce paralogs and are more accurate due to faster evolutional rates (Brown et al., 1979; Fučíková and Lahr, 2016; Stoneking, 2013). Indeed, the cytochrome oxidase subunits 1 (COI) gene has been frequently used as taxonomical barcode (Hebert et al., 2003) on several groups including Lepidoptera (Hebert et al., 2004a), birds (Hebert et al., 2004b) and spiders (Barrett and Hebert, 2005) due to its high level of species discrimination. This marker was then shown to be efficient for several protists lineages, for instance in red algae (Robba et al., 2006), ciliates (Barth et al., 2006) and diatoms (Evans et al., 2007). Nassonova et al. (2010) confirmed that COI was efficient to discriminate amoebozoan species such as *Vannella* and this marker has since been used for *Cochlopodium* (Tekle, 2014), *Rappella* (Kudryavtsev and Gladkikh, 2017) and is now considered as the canonical marker for Hyalospheniformes (Heger et al., 2013; Kosakyan et al., 2016b; Singer et al., 2015). However, the counterpart of this fast evolving marker is that "universal primer", e.g. Folmer's primer LCO1490-HCO2198 (Folmer et al., 1994), are not suitable for all groups and different specific primers have often to be developed (Kosakyan et al., 2013; Kudryavtsev, 2014). Chapter 2 (Blandenier et al., 2017a) focusses on the development and experimentation of a new mitochondrial marker for both deep phylogeny and taxonomy of Arcellinids (see also Appendix A (Macumber et al., 2019)).

1.4 PhD aims

The global aims of my thesis are to 1) improve and develop molecular methods to study amoebozoan diversity and ecology, 2) obtain a better knowledge of this diversity that plays important functions in nutrient cycling and, *in fine*, soil fertility, 3) improve the taxonomic and phylogenetic frame for this diversity in order to build a sound basis for further research and 4) characterize the ecological drivers which are likely to influence microbial diversity at local, continental and global scales. I used the infraorder Hyalospheniformes (Arcellinida; Amoebozoa) as my model group, but I also expanded my research to other amoebozoans and even to all protists in collaborative projects (Fig 1.6). The logical scheme of this thesis is to start by developing the molecular tools required for the investigation of Arcellinida, and especially of the Hyalospheniformes, to use these methods to assess and characterize their diversity and finally to analyze their ecological requirement and their distribution patterns. The Hyalospheniformes taxonomy has recently been considerably improved (Kosakyan et al., 2016a), allowing the evaluation of the ecological and biogeographical patterns with the highest taxonomical resolution. In this thesis, I will use a combination of morphological, genetic and ecological data as suggested for hyalospheniids (Singer et al., 2018) or Cyanobacteria (González-Resendiz et al., 2019). The final goal of my PhD will be to investigate how far the diversity of eukaryotic microorganisms follows the same ecological and distribution patterns compared to their multicellular counterparts.

I have led all chapters of my thesis from the design of the project to the publication or manuscript preparation. In addition, I also had the opportunity to participate to some collaborations that are presented in the appendices section. These side projects have allowed me to investigate others taxonomic groups and to develop alternative research interests, for instance spatial modeling.

1.5 Summary of the different chapters and appendices

1.5.1 Chapters

CH2 - NAD: Development of a new molecular marker Based on the observation that markers based on the 18S gene were not adapted neither for fine-level taxonomy, nor to resolve the most ancient nodes within this clade, we developed a protocol to amplify and sequence a new mitochondrial marker which includes two conserved genes and an intergenic variable region. We tested our new protocol on several key Arcellinida taxa and applied it to environmental DNA as a potential new marker for metabarcoding studies. This marker combines the qualities of a good (meta)barcode with a short size but a high taxonomic resolution and simultaneously possesses a conserved region that allows to reveal deep relationships between far-related taxa. *Published in European Journal of Protistology, DOI: 10.1016/j.ejop.2016.12.002.*

CH3 - *Mycamoeba gemmipara*: the first cultivated member of Clade LKM74 We discovered and described *Mycamoeba gemmipara*, the first species of a new clade of tiny naked amoeba (Amoebozoa, Dermamoebidea). We then characterized its morphology, phylogenetical position among Amoe-

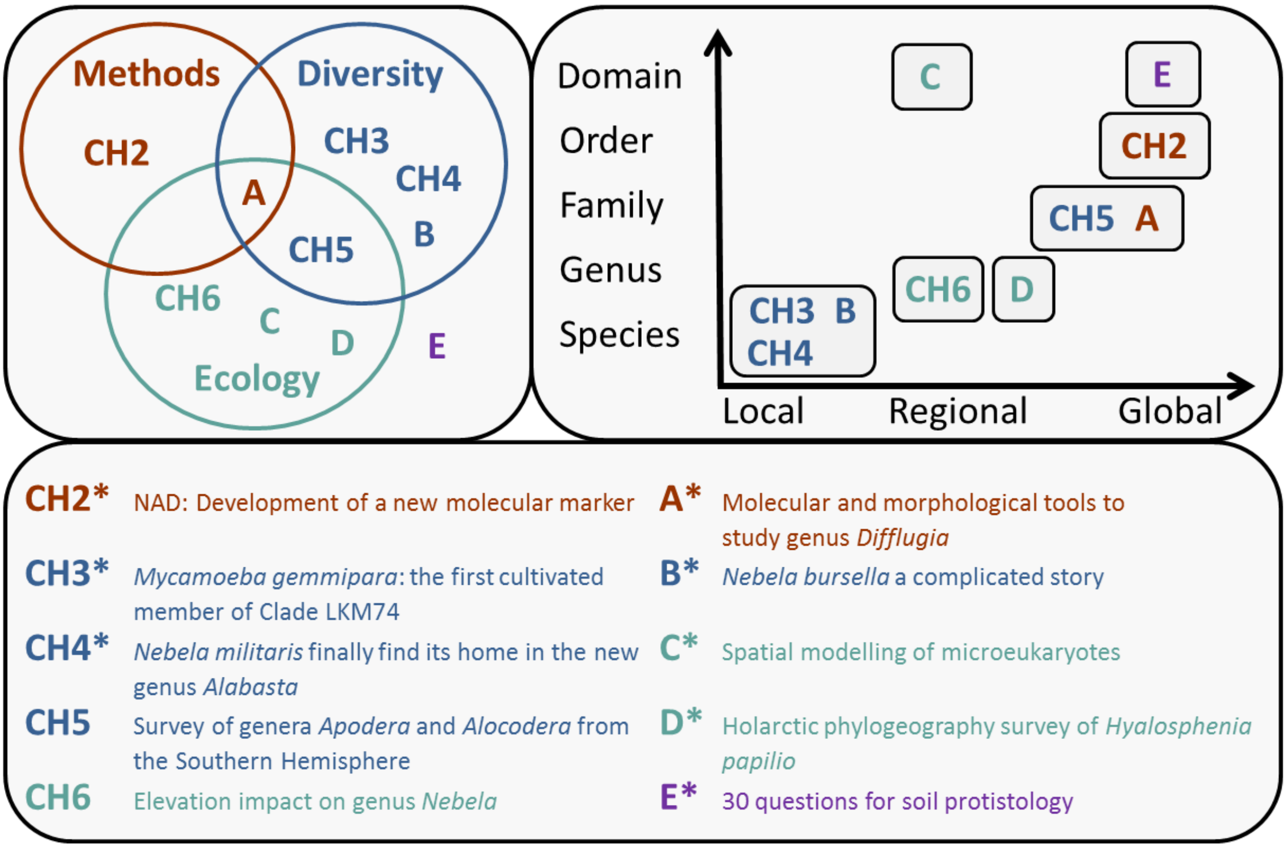

Figure 1.6 – Schematic outlook of my thesis highlighting the themes of the chapters (CH) and my collaborations (capital letters) in function of the three main foci of my thesis (methods, diversity and ecology). *: Articles already published.

bozoa and ecology. Finally, we emphasized its astonishing and unique life cycle based on an alternation of phagotrophic active trophozoites and osmotrophic fungi-like ramifications. *Published in Journal of Eukaryotic Microbiology, DOI: 10.1111/jeu.12357*

CH4 - *Nebela militaris* finally find its home in the new genus *Alabasta* We conducted a study on the peat bog species *Nebela militaris* with molecular and morphological methods to place phylogenetically the species among the Hyalospheniformes. Our analyses based on genetic and morphological criteria confirmed that *Nebela militaris* does not belong to genus *Nebela*. Therefore, we erected the genus *Alabasta* to accommodate it and other morphologically and ecologically similar species. *Published in European Journal of Protistology, DOI: 10.1016/j.ejop.2018.08.005.*

CH5 - Survey of the genera *Apodera* and *Alocodera* from the Southern Hemisphere We assessed the molecular and morphological diversity of genera *Apodera, Alocodera* and *Padaungiella* and discovered a vast diversity within each genus. The sequences obtained for *Alocodera* are biogeographically separated by continent, while European and Southern American *Padaungiella* isolates seemed phylogenetically divergent. Altogether, our results suggest that these Hyalospheniformes may contain more endemic taxa than previously thought.

CH6 - Elevation impact on genus *Nebela* We explored the diversity patterns of Hyalospheniformes along an elevation gradient and investigated how this diversity was related to elevation and other variables. We restricted our survey on a very homogenous habitat, i.e. forest *Sphagnum*-dominated peatlands, and focused on genus *Nebela*, which was already shown to present strong habitat selection. We detected that the abundance and diversity of species of this genus tend to decrease in high elevation and that the richness distribution seems to present a typical mid-domain effect. Moreover, we also discovered competitive exclusion between high elevation specialist species and dominant low elevation generalist ones.

1.5.2 Appendices

My contribution in the following side projects varied from different aspects of the work summarized in Table 1.1: experimental design, sampling, laboratory work, data analysis and writing. I differentiated my involvement as: not involved, slightly involved, significantly involved, and in charge of the

Table 1.1 – Table 1.1- Summary of my involvement in the different side projects realized during my PhD. NA = Not Available, - = Not involved, + = Slightly involved, ++ = Significantly involved, +++ = In charge of the whole section.

Appendices	Experimental design	Sampling	Laboratory work	Data analysis	Writing
A - Molecular and morphological tools to study genus *Difflugia*	++	-	+	++	++
B - *Alabasta militaris* zoological code	++	NA	NA	+++	++
C - Spatial modelling of microeukaryotes	-	-	++	-	+
D - Holarctic phylogeography survey of *Hyalosphenia papilio*	-	-	-	++	++
E - 30 questions for soil protistology	+	NA	+	-	+

whole section.

A - Molecular and morphological tools to study genus *Difflugia* We combined the high-resolution marker NAD9-NAD7 with advanced morphometrics to assess the phylogenetic signal of morphological traits of elongated *Difflugia* species. The phylogenetical analyses revealed two clades which could be reliably separated by test size and the presence/absence of mixotrophy. We noticed that mixotrophic taxa have generally larger tests characterized by wide flat bases, and hypothesized that these traits are beneficial for endosymbionts performance. *Published in European Journal of Protistology, DOI: 10.1016/j.ejop.2019.125645.*

B - *Nebela bursella* a complicated story While reviewing the history of *Nebela militaris*, we noticed that this species was a subjective junior synonym of *Nebela bursella*, and that the taxonomic relationship of both species was confusing. We made a request to the Bulletin of Zoological Nomenclature in order to be allowed to keep the specific epithet *militaris* for our new genus *Alabasta* and we reviewed the whole history of these species. *Published in Bulletin of Zoological Nomenclature.*

C - Spatial modelling of microeukaryotes We compared the power of topo-climatic versus edaphic variables for predicting the diversity of various soil protist groups in the Swiss western Alps. We generated the dataset through a metabarcoding approach of environmental soil DNA from 178 sites, where several key topo-climatic and edaphic variables were measured, and modelled the protists diversity patterns in Generalized Additive Models. We found that topo-climatic variables had a better predictive power than edaphic variables. However, predictive power varied among taxonomic groups. *Published in Journal of Biogeography, DOI: 10.1111/jbi.13755.*

D - Holarctic phylogeography survey of *Hyalosphenia papilio* We investigated the diversity of *Hyalosphenia papilio* genetic lineages and quantified their phylogeographical patterns to test if current distribution can be explained by historical factors or by the current distribution of suitable habitats. We discovered higher diversity in Western North America, which confirmed our inferred geographical origin for this species complex. Moreover, we proposed that the global distribution of *H. papilio* is tightly linked to the history of *Sphagnum*-peatland expansions and contractions in response to Quaternary glaciations. *Published in Molecular Ecology, DOI: 10.1111/mec.15117.*

E - 30 questions for soil protistology Soil protists diversity and their key ecological functions in ecosystems have recently received an increasing interest from the scientific community. This is mainly due to methodological and technical advances that have made their study more accessible. Based on a synthesis of experts opinion, we provided a synthesis of research gaps, in a broad range of topics, that should be prioritized in future studies of soil protistology to guide its ongoing revolution. *Published in Soil Biology and Biochemistry, DOI: 10.1016/j.soilbio.2017.04.001.*

Chapter 2

NAD9/NAD7 (mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene) — A new “Holy Grail” phylogenetic and DNA-barcoding marker for Arcellinida (Amoebozoa)?

NAD9/NAD7 (mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene) — A new "Holy Grail" phylogenetic and DNA-barcoding marker for Arcellinida (Amoebozoa)?

QUENTIN BLANDENIER[a,b], Enrique Lara[a], Edward A. D. Mitchell[a,c], Daniel M. C. Alcântara[b], Ferry J. Siemensma[d], Milcho Todorov[e], Daniel J. G. Lahr[b]

Abstract

Molecular phylogeny is an indispensable tool for assessing evolutionary relationships among protists. The most commonly used marker is the small subunit ribosomal RNA gene, a conserved gene present in many copies in the nuclear genomes. However, this marker is not variable enough at a fine-level taxonomic scale, and intra-genomic polymorphism has already been reported. Finding a marker that could be useful at both deep and fine taxonomic resolution levels seemed like a utopic dream. We designed Amoebozoa-specific primers to amplify a region including partial sequences of two subunits of the mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene (NAD9/NAD7). We applied them to arcellinids belonging to distantly related genera (*Arcella*, *Difflugia*, *Netzelia* and *Hyalosphenia*) and to Arcellinid-rich environmental samples to obtain additional Amoebozoa sequences. Tree topology was congruent with previous phylogenies, all nodes being highly supported, suggesting that this marker is well-suited for deep phylogenies in Arcellinida and perhaps Amoebozoa. Furthermore, it enabled discrimination of close-related taxa. This short genetic marker (ca. 250 bp) can therefore be used at different taxonomic levels, due to a fast-varying intergenic region presenting either a small intergenic sequence or an overlap, depending on the species.

Keywords: *Arcella*; Environmental DNA survey; Intergenic region; Mitochondrion; Molecular barcoding

[a]Laboratory of Soil Biodiversity, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland
[b]Department of Zoology, University of Sao Paulo, 05508-090 Sao Paulo, Brazil
[c]Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland
[d]Julianaweg 10, 1241VW Kortenhoef, Netherlands
[e]Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, 2 Gagarin St., 1113 Sofia, Bulgaria

European Journal of Protistology, 2017, 58, 175-186
doi:10.1016/j.ejop.2016.12.002

2.1 Introduction

Molecular phylogeny is the golden standard for phylogenetic reconstruction as well as species-level taxonomy. However, at both levels, currently available markers have their limits. Universal single-gene markers for deep phylogeny are not readily applicable to several eukaryotic groups such as the entire supergroup Amoebozoa (Lahr et al., 2011b; Tekle et al., 2008, 2016). When available, universal primers generally target the ribosomal small subunit (18S or SSU rDNA gene), but present biases across the large phylogenetic diversity of eukaryotes (Pawlowski et al., 2012). Likewise, species-level markers such as cytochrome oxidase subunit 1 (COI or COX) are generally useful only within small phylogenetic groups and different primer sets need to be designed for each group (Kudryavtsev, 2014). Finding a marker that would be useful for both deep phylogenetic and species-level studies is a goal that few taxonomists dare dream about – a true Holy Grail. Here we report the finding of a marker that comes very close to such a dream.

The eukaryotic superclade Amoebozoa comprises mainly free-living, unicellular, heterotrophic protists. Most species are characterized by metabolic cell bodies, often highly variable in their shape, as well as a unique way of moving and feeding by means of cytoplasmic streaming or pseudopodia (Grebecki, 1994). One group of Amoebozoa, the Arcellinida, have a test (shell) and belong to a polyphyletic morphological group termed "testate amoebae" (Kosakyan et al., 2016a). There are currently about 2,400 described Amoebozoa species, mainly catalogued using morphological approaches (Gray et al., 2004). However, environmental DNA studies have shown that this diversity is greatly underestimated (Berney et al., 2015; Geisen et al., 2014). Underestimates are due to both a lack of universality with general markers as well as low taxonomic resolution at species level. Furthermore, Amoebozoa are usually underrepresented in environmental DNA studies based on 18S rDNA, owing to long 18S rDNA sequence, frequent mismatches in primer regions and presence of introns (Berney et al., 2004; Geisen et al., 2015a; Pawlowski et al., 2012). The few existing DNA barcoding studies based on COI have revealed a multitude of genotypes hidden within morphospecies, some of which are differentiable only based on ultrastructural features and are referred to as "pseudo-cryptic species" (Kosakyan et al., 2012; Nassonova et al., 2010; Singer et al., 2015).

The most commonly used eukaryotic barcoding marker is the nuclear gene coding for the 18S rDNA because it is found in the genomes of all eukaryote and possesses a mosaic of highly conserved and variable nucleotide sequences (Pawlowski et al., 2012). Yet, unlike in other groups (e.g. Foraminifera, (Morard et al., 2016; Pawlowski and Holzmann, 2014), the 18S rDNA gene is not sufficiently variable to resolve interspecies relationships in several taxa including Amoebozoa (Nassonova et al., 2010). Other nuclear markers such as actin have been used to a lesser extent, but this marker is of limited usefulness in the case of Arcellinida due to the presence of paralogs (Lahr et al., 2011b). There are also concerns about the possible existence of paralogs of the 18S rDNA in the nuclear genome itself (Oliverio et al., 2014). On the other hand, Amoebozoa 18S rDNA gene sequences often present high heterogeneity in rates of evolution across sites and between species, and both phenomena impair considerably the establishment of deep phylogenies (Tekle et al., 2008). This problem has also been reported in arcellinids, where fast-evolving taxa can be difficult to place in trees and distort the relationships among internal branches (Kudryavtsev et al., 2009).

Mitochondrial genes are a promising alternative as they normally do not duplicate and evolve generally faster than nuclear genomes (Brown et al., 1979; Fučíková and Lahr, 2016; Stoneking, 2013); on the other hand, genes coding for some conserved protein genes may be easy to align and conserved enough to infer ancient phylogenetic relationships (Fučíková and Lahr, 2016). The commonly used mitochondrial barcoding marker COI, which is now established as the canonical barcoding marker for animals (Hebert et al., 2003) has been used also for various Amoebozoa groups (Kudryavtsev, 2014; Nassonova et al., 2010) and in particular for arcellinid testate amoebae (Kosakyan et al., 2012, 2013, 2016a; Singer et al., 2015). Although this gene proved to be an excellent marker for fine-level taxonomy, it is too variable to assess ancient dichotomies.

Nicotinamide adenine dinucleotide dehydrogenases genes (NADH) have been hypothesized to be the most conserved group of Amoebozoa mitochondrial genes (Heidel and Glockner, 2008). Therefore, we hypothesized that potential barcode markers could be found in these genes. To date, ten complete amoebozoan species mitochondrial genomes from seven different genera are available in GenBank under "Amoebozoa complete mitochondrion" search (Table 2.1). These species belong to distantly-related organisms (Lahr et al., 2011a; Nikolaev et al., 2005; Smirnov et al., 2005). As mitochondrial genomes of Amoebozoa show high levels of synteny (Fučíková and Lahr, 2016), we designed a set of primers to amplify a presumably highly conserved region in the NADH gene complex, i.e., the end of NAD9 and the beginning of NAD7, in order to obtain a marker that would be conserved enough in evolution to resolve deep phylogenetic relationships but which could be used to discriminate closely related species (due to the redundancy of the genetic code and the putative variability of the intergenic region). We tested amplification of this region with a selection of arcellinid species covering the main phylogenetic lineages, in order to evaluate the efficiency of the proposed marker. In addition, we applied these primers to environmental DNA samples to test whether this marker could be used in environmental barcoding studies.

2.2 Material and Methods

2.2.1 Testate amoeba collection and DNA extraction

Testate amoebae were obtained from various environments (i.e., mosses, ponds, bog lakes and sediments) and different geographical origins (Table 2.2). Both *Arcella* lineages were collected from a pond in Sao Paulo (BR), identified in Utermöhl's plankton chambers (Hydro-Bios, Kiel, Germany) under inverted microscope (Olympus IX81 mounted with Plan Apo objectives (Mitutoyo, Japan) with 40x magnification), picked up with a stirred Pasteur pipette, cleaned several times with sterilized pond water and cultured in 25 ml flasks containing 1 ml of cereal grass media (Lahr and Katz, 2009). For the other species, two to 28 cells of similar morphology and originated from the same sample were picked up individually, as explained above, and placed together into Eppendorf tubes containing 100 µl of guanidine buffer (Chomczynski and Sacchi, 1987). *Arcella* cultures were spun in microcentrifuge tubes for 10 min at 5000x g and 800 µl of guanidine buffer was added to the resulting pellet. DNA extraction was performed in these tubes as described in Gomaa et al. (2012). DNA was finally re-suspended in sterile demineralized water. Environmental DNA was extracted from *Sphagnum* mosses collected in the peatland Le Cachot in the Swiss Jura Mountains (Table 2.2). Env_L_1 and Env_L_2 are two environmental clones belonging to a lawn and Env_B_1 to a peatland border as described in Singer et al. (2016). Testate amoeba cells were extracted from 20 g of fresh *Sphagnum* and concentrated by sieving at 150 µm. The resulting fraction was used for global DNA extraction with a MoBio Power Soil® DNA Isolation kit according to the manufacturer's instruction (Kosakyan et al., 2015).

Cells used for DNA extraction were documented using alternatively a Leitz Orthoplan and an Olympus IX81 microscope. Shells were prepared and documented for scanning electron microscopy (SEM) as described previously in Gomaa et al. (2012). The following reference papers were used for species identification: Chardez (1967), Deflandre (1928), Ogden (1979, 1980). Test size, shape and composition, as well as nucleus shape (especially in genus *Difflugia*) were used as criteria for species identification and discrimination.

2.2.2 Primer design, PCR amplification and sequencing

We designed primers comparing the entire mitochondrial genomes from the following amoebozoans: *Acanthamoeba castellanii* (NC 001637), *Dictyostelium citrinum* (NC 007787), *Dictyostelium discoideum* (NC 000895), *Polyspondylium pallidum* (NC 006862) and *Vermamoeba vermiformis* (NC 013986) (Table 2.1). *Phalansterium* (KC121006) and *Physarum polycephalum* (NC 002508) mitochondrial genomes were partly unannotated and therefore could not be used for primer design. *Acanthamoeba polyphaga* (KP054475) and *Balamuthia mandrillaris* (NC 027736) genomes sequences arise after the design of the primer. All the NADH dehydrogenases sequences were obtained by a single polymerase chain reaction (PCR) with two newly design primers NAD9_386F (5′-TGG TTA GAA CGA GAA GTT TGG GAT ATG T-3′) and NAD7_67R (5′-GTG CGC AGC AGG RTG TTG WGG WCC- 3′) (list of all primers designed in Supplementary material 1). The DNA was amplified in a total volume of 17 µl with an amplification profile consisting of a 3 min initial denaturation at 94°C followed by 35 or 40 cycles program of 30 s at 94°C, 30 s at 61°C, 1 min at 72°C and a final elongation of 10 min at 72°C. Except for *Netzelia tuberculata*, all positive products were cloned using a PCR2.1 Topo TA cloning vector and transformed into *E. coli* TOP10' One Shots cells (Invitrogen kit) according to the manufacturer's instructions. *N. tuberculata* was sequencing without any cloning step. The protocol used for amplifying environmental DNA was identical to the one used for isolated cells and *Arcella* cultures. Sequencing of five to ten clones for each amplicon (34 for each environmental clones) was performed with single direction by an initial amplification using a BigDye197 Ter-

Table 2.1 – List of the ten complete mitochondrial genomes.

Species		NCBI reference sequence	References
Acanthamoeba	*castellanii*	NC_001637	(Burger et al., 1995)
	polyphaga	KP054475	Karlyshev et al., unpublished
Balamuthia	*mandrillaris*	NC_027736	(Greninger et al., 2015)
Dictyostelium	*citrinum*	DQ336395	(Heidel and Glockner, 2008)
	discoideum	NC_000895	(Ogawa et al., 2000)
	fasciculatum	EU275727	(Heidel and Glockner, 2008)
Phalansterium	sp.	KC121006	(Pombert et al., 2013)
Physarum	*polycephalum*	NC_002508	(Takano et al., 2001)
Polysphondylium	*pallidum*	NC_006862	Burger et al unpublished
Vermamoeba	*vermiformis*	NC_013986	(Bullerwell et al., 2010)

Table 2.2 – List of sequenced species, sampling localization and sequences details.

Species	Environments	Localisations and countries	Co-ordinates	Number of cells used per extraction	Sequences length (bp)	AT%	Genbank number
Arcella intermedia	pond	Sao Paulo (BR)	23◦57’ S, 46◦73’ W	40k	253 (+ 5*)	80.7	KY499885
Arcella intermedia laevis	pond	Sao Paulo (BR)	23◦57’ S, 46◦73’ W	40k	253 (+ 5*)	79.5	KY499884
Difflugia nodosa	Bog lake	Ljulin (BG)	42◦38’ N, 23◦09’ E	5	246 (+1*)	71.3	KY499878
Difflugia nodosa	pond	Groeneveld (NL)	52◦21’ N, 5◦25’ E	21	245 (+1*)	70.7	KY499879
Difflugia pyriformis	pond	Groeneveld (NL)	52◦21’ N, 5◦25’ E	28	267 (+1*)	70.9	KY499876
Difflugia pyriformis	ditch	Laegieskamp (NL)	52◦28’ N, 5◦14’ E	20	267 (+1*)	70.5	KY499877
Env_L_1	*Sphagnum* sp	Le Cachot (CH)	47◦00’ N, 6◦39’ E	-	251	77.7	KY499887
Env_L_2	*Sphagnum* sp	Le Cachot (CH)	47◦00’ N, 6◦39’ E	-	260	70.8	KY499880
Env_B_1	*Sphagnum* sp	Le Cachot (CH)	47◦00’ N, 6◦39’ E	-	320 (+1*)	72.8	KY499888
Hyalosphenia papilio	*Sphagnum* sp	Le Cachot (CH)	47◦00’ N, 6◦39’ E	3	261 (+1*)	77.9	KY499886
Netzelia oviformis	pond	Neuchâtel (CH)	47◦00’ N, 6◦95’ E	2	256	77.0	KY499881
Netzelia tuberculata	lake	le Loclat (CH)	47◦02’ N, 6◦99’ E	4	253 (+1*)	79.8	KY499882
Netzelia wailesi	sediments	Sima-Moor (AT)	47◦58’ N, 12◦62’ E	3	262 (+1*)	78.6	KY499883

minator Cycle Sequencing Ready Reaction Kit (Applied Biosystems), further analysed with an ABI-3130XL DNA sequencer (Applied Biosystems). Sequences were deposited in GenBank with the following accession number: KY499876–KY499888.

2.2.3 Alignment and phylogenetic analysis

Obtained sequences (environmental, cultures and isolates-derived) were aligned against all Amoebozoan sequences found in GenBank that included the fragment of interest, plus the oomycete *Phytophthora ramorum* NC 009384 as outgroup. The sequences were aligned manually using BioEdit software (Hall, 1999). We built a Maximum Likelihood phylogenetic tree with the software MEGA6 (Tamura et al., 2013). T92+G was selected as the best model using the Akaike criteria as implemented in the MEGA package. Robustness of the nodes was calculated by performing 500 bootstraps replicates.

2.3 Results

2.3.1 Structure of the sequences

We used the primers combination which worked the best (i.e., NAD9_386F–NAD7_67R) among all possible (Supplementary material 1) on the collected species (Figure 2.1). We obtained 13 sequences from at least ten different arcellinid testate amoeba taxa (Table 2.2) from four distantly related families. Four sequences belonged to genus *Difflugia* (two isolates of *D. nodosa* and two of *D. pyriformis*), two from clonal cultures of genus *Arcella* (*A. intermedia* and *A. intermedia laevis*), three to genus *Netzelia* (*N. oviformis*, *N. tuberculata* and *N. wailesi*), one to *Hyalosphenia papilio* and three environmental clones sequence type (Env_L_1, Env_L_2 and Env_B_1) from *Sphagnum* samples.

The sequences ranged from 245 to 320 bp in length. NAD9 formed the main part of the sequences with a length between 210 bp (both *Arcella* lineages) and 226 bp (both *D. nodosa*) whereas NAD7 was much shorter, between 42 bp (both *Difflugia* and *Netzelia* species) and 61 bp (*Hyalosphenia papilio*). The AT ratio was high, ranging from 70.5% to 80.7% (Table 2.2) which is similar to *Acanthamoeba castellanii* (70.6%), *Dictyostelium citrinum* (73%) and *Poly-*

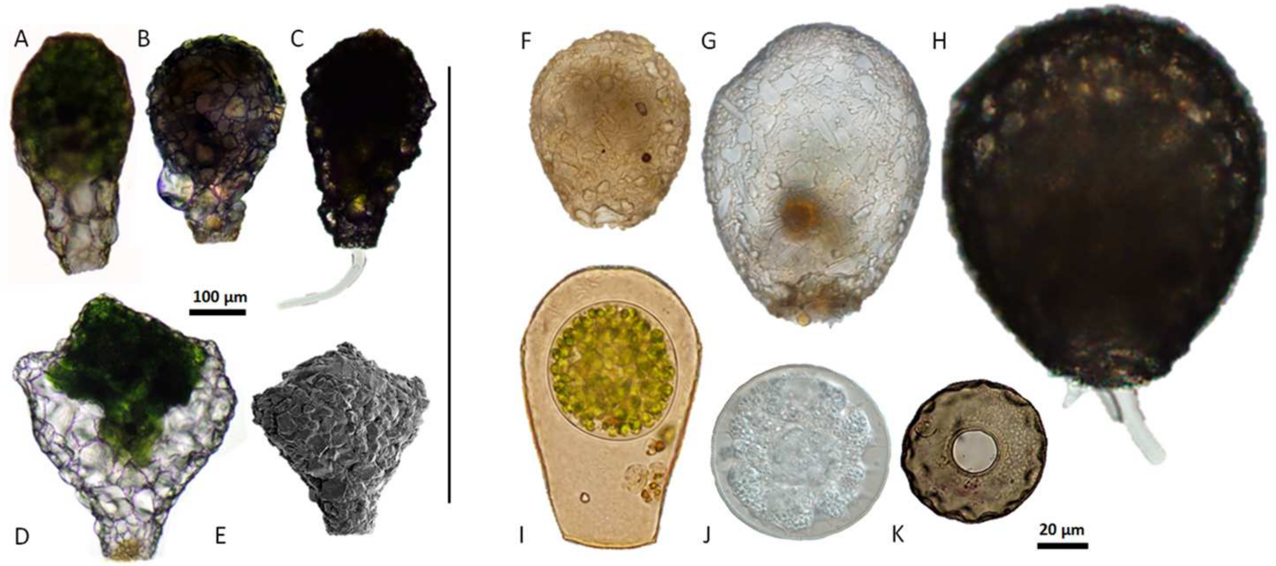

Figure 2.1 – Light (**AK**) and scanning electron (**E**) micrographs of the ten analysed species: **A.** *Difflugia pyriformis* from Groeneveld (NL); **B–C.** *D. pyriformis* from Laegieskamp (NL), in **C** an active cell showing a pseudopod; **D.** *D. nodosa* from Groeneveld (NL); **E.** *D. nodosa* from Ljulin (BG); **F.** *Netzelia wailesi* from Sima-Moor (AT); **G.** *N. oviformis* from Neuchâtel (CH); **H.** *N. tuberculata* from le Loclat (CH) showing pseudopods; **I.** *Hyalosphenia papilio* from Le Cachot (CH); **J.** *Arcella intermedia laevis* from Sao Paulo (BR); **K.** *A. intermedia* from Sao Paulo (BR) with characteristic depressions on the test. **J** and **K** *Arcella* pictures were taken by A. Porfirio at the University of Sao Paulo. Scale bar: **A–E** = 100 µm, **F–K**=20 µm.

sphondylium pallidum (75.8%). However, the AT:GC ratio on the priming region was not too unbalanced; furthermore, the high melting temperature of the primers allowed high annealing temperatures in the PCR cycling program, which reduced self-annealing problems.

The relative configurations of the NAD9 and NAD7 genes fell into three different categories (Figure 2.2). In the first case (both *D. pyriformis* isolates, *N. oviformis*, *N. wailesi* and Env_B_1) there was an intergenic region with specific sequences. This nucleotide indel is very short (i.e., 2–6 bp) and is made solely of A and T except for the Env_B_1 with a particularly long intergenic region (i.e., 46 bp) composed of all four nucleotides. In the second case (*N. tuberculata* and Env_L_2) the two genes are perfectly contiguous without any indels or overlaps. In the last case, the two genes overlapped, either only by 1 bp for *Arcella* (the A nucleotide at the end of TAA corresponded to the beginning of ATG), 7 bp for Env_L_1 as well as *Hyalosphenia papilio*, or 16 bp for both *D. nodosa*. All NAD9 sequences possessed a TGA codon (canonically a stop codon, Table 2.3). This codon was located in nucleotide position 195 for all species, except for the sequence of Env_L_2 where it was present in position 204 and for both *Arcella* and *Netzelia* which had two TGA codons in position 195 and 204. The insertion of single nucleotides in the sequences, corresponding to several species or groups of species, suggests frequent mitochondrial RNA editing, as demonstrated in other Amoebozoa (Traphagen et al., 2010).

2.3.2 Phylogenetic relationships among species

The phylogenetic tree (Figure 2.3) showed that all our sequences branched together, forming a robust arcellinid clade with 85% bootstrap support (B). Another robust clade (94% B) included all pyriform *Difflugia.* Likewise, both *Arcella* branch together robustly as well as both *Netzelia* (Table 2.5), all five sequences forming another robust clade (88% B). *Hyalosphenia papilio* and Env_L_1 were very similar (differing by 14 bp), and branch together along with Env_B_1 which is more basal (Figure 2.3). Finally, Env_L_2 position is weakly supported (<50% B) and is sometimes basal to the whole arcellinid tree (data not show).

The sequences we obtained using the NADH dehydrogenases primers allowed us to discriminate accurately all species, even in the case of closely-related species. The two isolates identified as *D. nodosa* differed by 14% in their NAD9/NAD7 sequences and the two *D. pyriformis* isolates by 6%. Similarly, the three *Netzelia* species diverged by, respectively, 9, 16 and 17% (Table 2.5).

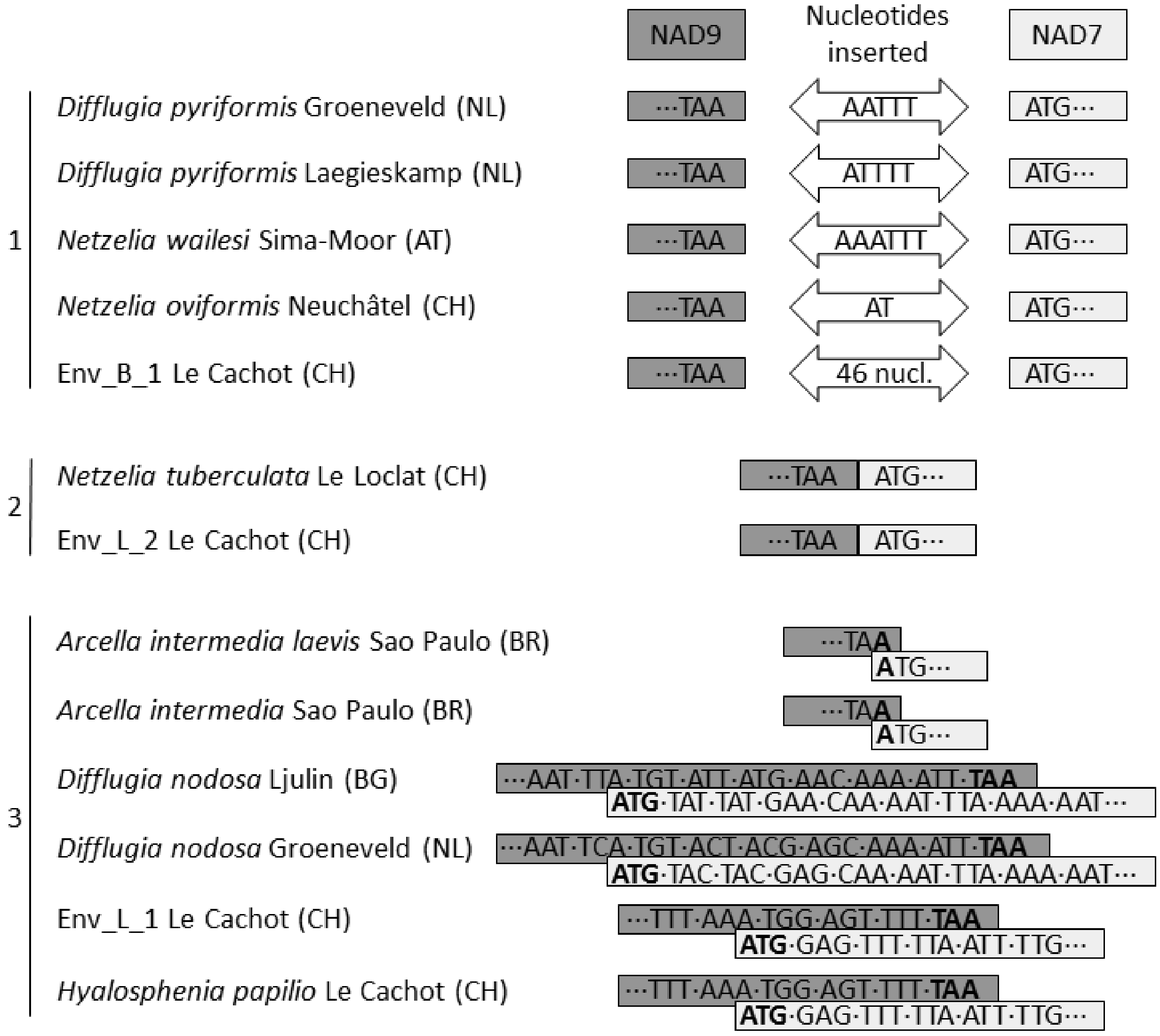

Figure 2.2 – Configuration of the three types of intergenic region. The dark grey box represents the NAD9 gene and the light grey one the NAD7. The three dots (···) express a continuation of the gene, whereas the single dot (·) expresses a fictive space between two codons. The first case represents the nucleotides insertion model, the second the contiguously genes alignment and finally the overlapping one.

2.4 Discussion

2.4.1 NAD9/NAD7 is an accurate fragment for species identification and phylogenetic reconstructions

We describe here a pair of primers that work reliably across all Arcellinida. The designed primers are, at first sight, incongruent with “canonical” practices, particularly GC content and length (Dieffenbach et al., 1993). This is because the amoebozoan mitochondrial genome is GC-poor (see Table 2.2), stable in terms of coding capacity and gene order, but present highly divergent nucleotide sequences, impossible to align outside coding regions (Glockner, 2013), resulting in lesser potential primer hybridization regions. Our two primers NAD9_386F and NAD7_67R successfully amplified the target genes for different taxa spread across the arcellinid tree.

The resulting fragment presents the advantage of being small (i.e., less than 320 bp) and is therefore easy to amplify. This short size may likewise allow their retrieval using high throughput sequencing (HTS) approaches, opening the way for the construction of a large Arcellinida sequence dataset, which has been a challenge so far owing to the difficulty to amplify the commonly used markers such as 18S rDNA. Indeed, environmental sequencing is very useful to strengthen the topology of phylogenetic trees and to discover unknown clades or new taxa within known clades. Here, we obtained an environmental sequence (Env_L_1) that

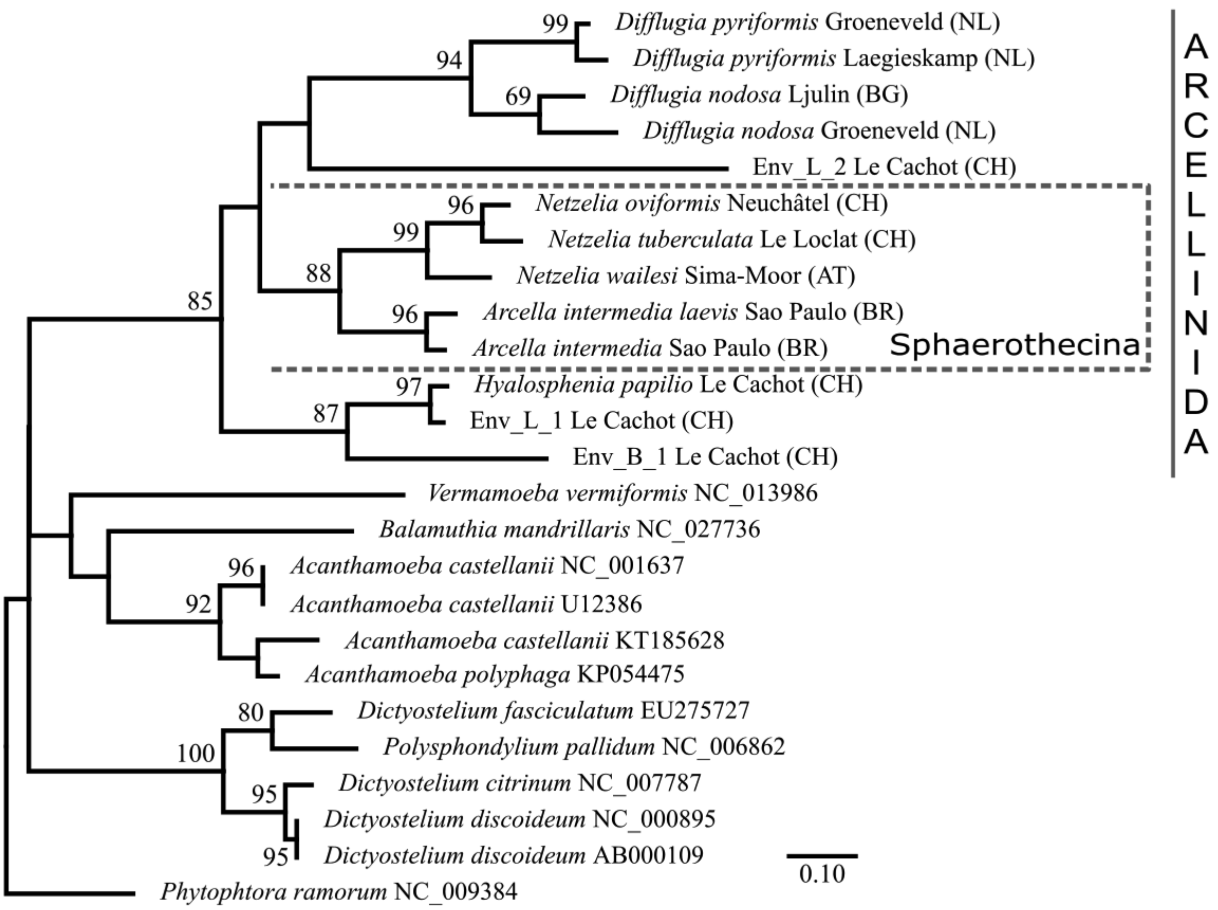

Figure 2.3 – Phylogenetic tree of Arcellinida based on NAD9-NAD7 genes sequences. The tree was rooted with the Oomycete *Phytophthora ramorum* and based on a maximum likelihood analysis. Numbers at the nodes indicate the resulting bootstraps (only values above 60% were kept). The scale bar indicates 0.1 sequence divergence.

was closely related to *Hyalosphenia papilio*. The *Sphagnum* sample from which it derived contained many cells of *Nebela guttata*, also belonging to Hyalospheniidae (Kosakyan et al., 2013). It is therefore most likely that the original organism from which the sequence derived belonged to this family. The other environmental clones encountered could not be related to any known sequence. Env_B_1, a clone from a peatland border, robustly branched with Hyalospheniidae. It belonged, likely, to the sister group of the clade; a good candidate could be the group of Bullinularia, which has been previously shown to robustly branch with that family (Gomaa et al., 2012). The third environmental clone Env_L_2 belongs also to Arcellinida but represents most likely a group unrelated to those represented in our study. After improving the method and develop a confident database, this marker might be used in environmental diversity and ecological studies in which Amoebozoa are typically poorly covered (Geisen et al., 2015a; Pawlowski et al., 2012).

On the other hand, NAD9/NAD7 may be used also for discriminating closely-related species at fine taxonomic levels. Notably, the two *Difflugia nodosa* isolates could be separated on the basis of their sequences (Table 2.5). These organisms were morphologically similar, with a large and robust pyriform test compressed in cross-section and vacuoles that were filled with numerous zoochlorellae in agreement with the description of Penard (1890). According toPenard (1890), the size ranges from 300 to 380 µm, yet the specimens from Groeneveld were considerably larger (483 µm in average), while the ones from Ljulin fitted the original description (353 µm). This taxon represents a case where molecular methods reveal hitherto unknown diversity. Similar instances are frequent in Arcellinida, as within the *Nebela collaris* complex (Kosakyan et al., 2012, 2013) or *Difflugia tuberspinifera* with or without spines (Gomaa et al., 2015). Thus, our study brings further support to the idea that the diversity of arcellinids has been underestimated by morphology-based taxonomy (Kosakyan et al., 2016a; Porfírio-Sousa et al., 2017).

Detailed study will most likely show that *D. nodosa* represents a complex of different molecular (and possibly morphological) species. This may be also valid for *D. pyriformis*, which belongs to a controversial group, the *Difflugia pyriformis* complex (Perty, 1849). Indeed, the morphological variation among this taxon is high and the descriptions are unclear (Mazei and Warren, 2014). A more detailed morphological and molecular study of this species complex would be needed to proceed to any taxonomic action. Our identifications of the analysed material were tentative as neither of our isolates corresponded in their dimensions to the original description of *D. pyriformis*. Indeed, cells from the Groeneveld isolate measured 302–410 µm and those from Laegieskamp 423 µm as compared to 140–200 µm for the original description (Perty, 1849).

The discrimination power of NAD9/NAD7 is due partly to the fast evolution of the third position of codons, but also to the varying size and sequences in the intergenic space. In summary, our new genetic marker contains both slow- and fast-evolving regions that are useful for phylogeny.

2.4.2 Three possibilities exist in the genetic frontier: insertion of an intergenic region, contiguously or overlap

All mitochondrial genes of Amoebozoa are closely packed and the overlap we observed is widespread in Amoebozoa (Burger et al., 1995; Ogawa et al., 2000). The intergenic space region can also possess a short sequence inserted (Figure 2.2) as shown in the two isolates of *Difflugia pyriformis*, *Netzelia wailesi*, *N. oviformis* and the Env_B_1 sequence as well as for *Dictyostelium discoideum*, *Polysphondylium pallidum* and *Vermamoeba vermiformis*. It is mostly solely composed of A and T repeats reaching up to six bp in *Netzelia wailesi*. Env_B_1 is the only current exception with a large nucleotides insertion (i.e., 46 bp) including two C and seven G base pairs. The region varies systematically even within one genus and each of the species surveyed in this study presented a different sequence. The case of *Netzelia tuberculata* shows a different configuration than the two others *Netzelia* as the two genes are perfectly contiguous. One could notice that the number of nucleotides inserted among this genus is diminishing from six in the case of *N. wailesi*, to two for *N. oviformis* and nothing for *N. tuberculata* which is more closely related to the previous species (Figure 2.3). However, in cases with overlap, the length of the "common" region is stable within genus; this is observed in both *D. nodosa* isolates, both *Arcella*, as well as Env_L_1 and *Hyalosphenia papilio*. It is noteworthy to mention that *Acanthamoeba castellanii* (NC 001637) and *A. polyphaga* (KP054475) possess also together an identical overlapping of 13 nucleotides (data not shown). This length can be proposed, therefore, as a deeper phylogenetic marker to be used for differentiating isolates at the genus level. This proximity suggests that these two genes need to be nearby to be efficient; this could be due to the existence of a single promoter for the regulation of the mitochondrial DNA genes transcription (Ogawa et al., 2000).

2.4.3 Arcellinids are using the same genetic code as *Acanthamoeba castellanii* and present RNA editing evidences

Both universal and alternative genetic codes have been identified in mitochondria (Heidel and Glockner, 2008). In Amoebozoa mitochondrial genomes, examples of both genetic codes have also been mentioned. Indeed, the social amoeba *Dictyostelium* (Cole and Williams, 1994), *Polysphondylium* and *Vermamoeba* use the universal code (i.e., TGA as a stop codon) whereas, *Acanthamoeba* uses the alternative option (i.e., TGA coding for tryptophan). All our Arcellinida mitochondrial sequences seem to employ an alternative genetic code. Indeed, the NAD9 genes are not supposed to end with a TGA codon (at position 195) given the presence of a tryptophan (W) located in the same position for Env_L_2 (Table 2.3), *Dictyostelium*, *Polysphondylium* and *Balamuthia*. The same case happens in position 204 for Env_L_2, both *Arcella* and *Netzelia* which possess also a TGA codon whereas a tryptophan is present at the homologous position in *Hyalosphenia* and Env_L_1, suggesting the use of a non universal code in their mitochondria. As *Vermamoeba* and Arcellinids are tubulinids and use different genetic codes, it seems that a transition took place within this clade. It is not yet clear where it occurred in the phylogeny of the group, as we do not yet have enough data to place this event; further data from more taxa will be needed to show if this can be considered as characteristic of Arcellinida. The same codon usage can be found in *Acanthamoeba*, likely resulting from convergence. As we have never found a TAG codon and all TGAs are coding for tryptophan, it is possible that arcellinids have only maintained the TAA stop codon in the mitochondria. This is not true for the nucleus, demonstrated by other arcellinid genes that apparently use TGA as stop (Lahr et al., 2011a,b, 2013). However, it is still necessary to verify if RNA editing is not being used in these

Table 2.3 – Amino acid sequence of NAD9 revealing the TGA codon (*) and the RNA editing sites (bold letter). The character Ω represent the TAA stop codon.

Amino acid position	10 20 30 40 50 60 70
Difflugia pyriformis Groeneveld (NL)	GLKFILHLDLRRTLTDYGFKGHPLCKDYPLCGYVELRYDDIHHLICS**A**PVEVSQGFRCFKFSNP*KTLCIFSΩ
Difflugia pyriformis Laegieskamp (NL)	GLKFILHLDLRRILTDYGFKGHPLCKDYPLCGYIELRYDDIHHLICL**A**PVEVSQGFRCFKFGNP*KTLCIFSΩ
Difflugia nodosa Ljulin (BG)	GLKFILHTDLRRILTDYGFKGFPLRKEFPLS**G**YVELHYDDIYQSISVAPLEVSQGFRNHKYKNP*RNLCIMNKIΩ
Difflugia nodosa Groeneveld (NL)	GLNFILHTDLRRILTDYGFKGHPLRKESPLNGYVELRYDDIYQSINVKPIEVSQVFRNY**K**YVNP*KNSCTTSKIΩ
Netzelia oviformis Neuchâtel (CH)	CI**K**FILHNDLRRILTDYGFHGHPLRKDYPLIGYVELYYNDLTQSIVFTSVEQMQGFRFFKFDNP*FN*KΩ
Netzelia tuberculata Le Loclat (CH)	CLKFILHNDLRRILTDYGFHGHPLRKDYPLIGYVELFYNDLIQSIVFTNVEQMQGFRFFKFDNP*FN*KΩ
Netzelia wailesi Sima-Moor (AT)	CIKFLLNNDLRRILTDYGFHGYPLRKDYPLIGYVELHYNDLMQSVIF**Y**CVEQMQGFRFFKFENP*FN*KLΩ
Arcella intermedia laevis Sao Paulo (BR)	GLKF**Y**LHSDLRRILNDY**G**VKGFPLRKDFP**L**IGFIELFYDDVNQSIII**E**CVEQTQSFRFFKFENP***Q**N*KVΩ
Arcella intermedia Sao Paulo (BR)	GLKF**Y**LHNDLRRILNDY**G**FKGFPLRKDFP**L**IGFIELFYDDVNQSIII**E**CVEQMQSYRFFKFENP***Q**N*KVΩ
Hyalosphenia papilio Le Cachot (CH)	GVKFLLHLDLRRLLTDYGFKGHPLRKDFPLIGYLEIRYDDILKIIISE**S**VETSQSYRVYRFINP*FKWSFΩ
Env_L_1 Le Cachot (CH)	GVKFLLHSDLRRLLTDYGFKGHPLRKDFPLIGYLEIRYDDILKIIISESVETSQSYRVYRFINP*FKWSFΩ
Env_L_2 Le Cachot (CH)	GCRFLLHEDLRRILTDYGFEGFPLRKDFPLVGYFEIRYENTTSFISRDPLEASQALRIFFFDNPWLL*KYΩ
Env_B_1 Le Cachot (CH)	GIKFLLHGDLRRLLTDYGFRGHPLRKDFPLIGYFEARYDDILKGIFTE**S**VETAQVYRRYKFINP*AKLNLFΩ

non-canonical TGA positions.

RNA editing seems to occur in arcellinids as well as for other Amoebozoa (Table 2.4), as already inferred for COI (Oliverio et al., 2015). NAD RNA editing is thus consistent with previous knowledge. We propose a nucleotide addition hypothesis (Table 2.4) based on sequence similarities. However, the exact location of editing sites remains difficult to infer in Env_L_1 and *Hyalosphenia papilio*. We observe also a narrow phylogenetic distribution of this process, especially for the *Arcella* lineages which shared the six same nucleotides added to the same loci, but also for both *Difflugia pyriformis* isolates which possessed in common a G addition in position 145 (Table 2.4). We did not observe nucleotide indels or substitutions (Gray, 2012).

2.4.4 NAD9/NAD7 is a reliable phylogenetic marker for the Arcellinida

Our phylogenetic analysis with marker NAD9/NAD7 supports the monophyly of Arcellinida robustly (85% B) and separates genera congruently to the current literature (Gomaa et al., 2015; Lara et al., 2008). Such a result suggests that the new marker may be useful also for deep phylogeny. Our survey on different genera is encouraging and indicates that the addition of further sequences from an even broader phylogenetic sampling (e.g., *Nebela*, *Centropyxis*, *Heleopera*, etc.) will be useful to reconstruct the Arcellinida tree.

Our new primers were efficient to distinguish closely related species as expected with two well-conserved mitochondrial genes. It is generally believed that mitochondrial genomes evolve faster than nuclear genomes (Heidel and Glockner, 2008). COI has indeed proven useful to discriminate closely-related species in Hyalospheniids, but failed to resolve more basal nodes (Kosakyan et al., 2013). The nuclear ribosomal internal transcribed spacer (ITS1/ITS2), possibly suffers from the same disadvantages, and may be present in multiple copies in the genome as suspected for other amoebozoans (Nassonova et al., 2010). Actin, a commonly used protein-coding nuclear marker (e.g., (Kudryavtsev et al., 2011; Kudryavtsev and Pawlowski, 2015; Nikolaev et al., 2005) has been shown to present high paralogy in Arcellinida genomes (Lahr et al., 2011b). Finally, Arcellinids possess highly variable 18S rDNA genes generating often long branches in the trees (Kudryavtsev et al., 2009, 2011) and frequent mutations in the priming sites, precluding the amplification of full length sequences. Furthermore, the frequent presence of introns and insertions in many taxa (Gomaa et al., 2012, 2015; Lara et al., 2008) renders alignments even more difficult. For instance, the information provided for a comprehensive phylogeny of the group including sequences of more than 2 kb in length had to be reduced to 700 bp alignable molecular characters. In comparison, NAD9/NAD7, despite its short length, contains enough phylogenetic information to retrieve the relationships between groups that can be observed with 18S rDNA. Indeed, results recently obtained by other means such as the monophyly of the Sphaerothecina (Kosakyan et al., 2016a), as well as monophyly of the vase-shaped members in genus *Difflugia* are robustly supported. The node supporting the monophyly of Arcellinida is also unambiguous. In addition, branch lengths in the NAD9/NAD7 tree seem more even than in 18S rDNA phylogeny (compare to (Gomaa et al., 2012; Kudryavtsev et al., 2011)), suggesting that reconstructions inferred from this fragment should be less prone to long-branch attraction. From this viewpoint, NAD9/NAD7 appears as an attractive example of a gene that can be used both for barcoding and relatively deep phylogeny. However, as for any mitochondrial marker, the use of NAD9/NAD7 is limited to species living under aerobic conditions; parasitic taxa such as the amitochondrial amoebozoan *Entamoeba histolytica* lack a mitochondrial

Table 2.4 – List of the RNA editing sites.

Species	Nucleotide editing	Positions	Resulting codon
Difflugia pyriformis Groeneveld (NL)	G	144	A (**G**CT)
Difflugia pyriformis Laegieskamp (NL)	G	144	A (**G**CT)
Difflugia nodosa Ljulin (BG)	A	182	K (AAA)
Difflugia nodosa Groeneveld (NL)	G	97	G (GGT)
Netzelia tuberculata Le Loclat (CH)	A	11	A (AAA)
Netzelia wailesi Sima-Moor (AT)	C	146	Y (TA**C**)
Arcella intermedia laevis Sao Paulo (BR)	T;T;A;A;A;T	15;55;89;143;196;246	Y (**T**AT); G (GG**T**); L (TT**A**); E (GAA); Q (CAA); L (**T**AA)
Arcella intermedia Sao Paulo (BR)	T;T;A;A;A;T	15;55;89;143;196;246	Y (**T**AT); G (GG**T**); L (TT**A**); E (GAA); Q (CAA); L (**T**AA)
Hyalosphenia papilio Le Cachot (CH)	T	147	S (**T**CA)
Env_B_1 Le Cachot (CH)	T	147	S (**T**CA)

Table 2.5 – Nucleotide similarity for NAD9-NAD7 genes [%].

	D. pyriformis G.	*D. pyriformis* L.	*D. nodosa* (BG)	*D. nodosa* (NL)	*N. oviformis*	*N. tuberculata*	*N. wailesi*	*A. intermedia* l.	*A. intermedia*	*H. papilio*	Env_L_1	Env_L_2	Env_B_1
D. pyriformis Groeneveld (NL)	-												
D. pyriformis Laegieskamp (NL)	**94**	-											
D. nodosa Ljulin (BG)	78	77	-										
D. nodosa Groeneveld (NL)	75	75	**86**	-									
N. oviformis Neuchâtel (CH)	67	65	64	64	-								
N. tuberculata Le Loclat (CH)	70	67	67	67	**91**	-							
N. wailesi Sima-Moor (AT)	66	64	66	67	**84**	**83**	-						
A. intermedia laevis Sao Paulo (BR)	68	68	70	68	76	75	75	-					
A. intermedia Sao Paulo (BR)	67	67	70	67	76	77	76	**94**	-				
H. papilio Le Cachot (CH)	67	67	67	67	69	69	67	70	72	-			
Env_L_1 Le Cachot (CH)	67	67	66	67	67	69	67	70	73	**96**	-		
Env_L_2 Le Cachot (CH)	64	64	65	61	60	61	59	66	66	64	64	-	
Env_B_1 Le Cachot (CH)	64	63	64	66	67	67	67	69	69	76	76	61	-

genome (Loftus et al., 2005) and will therefore not be covered by our proposed mitochondrial marker, and this will be a limitation for environmental DNA surveys.

Finding a marker that is useful for both deep phylogenetic and species-level studies has long been all but an unrealistic dream – indeed a potential barcoding “Holy Grail”. NAD9/NAD7 is as close to such a dream as can be!

2.5 Acknowledgements

This study was supported by Funds Wüthrich et Mathey-Dupraz; EL was supported by a grant from the Swiss SNF (31003A_163254), and DJGL by a FAPESP Young Investigator Award (# 2013/04585-3). We thank A. Pillonel for technical help, D. Singer for support and fruitful discussions, and Angie Opitz for the samples from Sima-Moor which contained *Netzelia wailesi*. We are finally grateful to A. Porfirio and C. Duckert for their help concerning cultures, isolation, identification and micrographs.

2.6 Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ejop.2016.12.002.

Chapter 3

Mycamoeba gemmipara nov. gen., nov. sp., the first cultured member of the environmental Dermamoebidae clade LKM74 and its unusual life cycle

Mycamoeba gemmipara nov. gen., nov. sp., the first cultured member of the environmental Dermamoebidae clade LKM74 and its unusual life cycle

---

Quentin Blandenier[a], Christophe V.W. Seppey[a], David Singer[a], Michèle Vlimant[b], Anaële Simon[c], Clément Duckert[a], Enrique Lara[a]

Abstract

Since the first environmental DNA surveys, entire groups of sequences called “environmental clades” did not have any cultured representative. LKM74 is an amoebozoan clade affiliated to Dermamoebidae, whose presence is pervasively reported in soil and freshwater. We obtained an isolate from soil that we assigned to LKM74 by molecular phylogeny, close related to freshwater clones. We described *Mycamoeba gemmipara* based on observations made with light- and transmission electron microscopy. It is an extremely small amoeba with typical lingulate shape. Unlike other Dermamoebidae, it lacked ornamentation on its cell membrane, and condensed chromatin formed characteristic patterns in the nucleus. *M. gemmipara* displayed a unique life cycle: trophozoites formed walled coccoid stages which grew through successive buddings and developed into branched structures holding cysts. These structures, measuring hundreds of micrometres, are built as the exclusive product of osmotrophic feeding. To demonstrate that *M. gemmipara* is a genuine soil inhabitant, we screened its presence in an environmental soil DNA diversity survey performed on an experimental setup where pig cadavers were left to decompose in soils to follow changes in eukaryotic communities. *Mycamoeba gemmipara* was present in all samples, although related reads were uncommon underneath the cadaver.

Keywords: Budding; Discosea; eukaryotic diversity; fungus; high throughput sequencing; Longamoebia; ribosomal genes; serial dilution; yeast

[a]Laboratory of Soil Biodiversity, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland
[b]Laboratory of Ecology and Evolution of Parasites, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland
[c]Laboratory of Microbiology, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

Journal of Eukaryotic Microbiology, 2016, 0, 1-9
doi:10.1111/jeu.12357

3.1 Introduction

Our vision of protist diversity has been radically challenged since the introduction of observation-independent environmental DNA surveys. Large and deep-branching groups of eukaryotes have been discovered with classical cloning/sequencing strategies, thus overtaking our estimations on eukaryotic environmental diversity. Unsuspected alveolate clades appeared to be extremely diverse in marine systems (Lopez-Garcia et al., 2001), and were found later to be exclusively composed by parasitoids (Guillou et al., 2008). Likewise, many new stramenopile lineages were discovered in the early 2000s in marine systems (Massana et al., 2004). These organisms, which include nowadays 25 different lineages spread all across the tree of stramenopiles, were found to be the most diverse and numerous bacterivores in the sunlit part of oceans (Massana et al., 2014), and thus vesting them with a prominent role in the oceanic microbial loop (Azam et al., 1983). In deeper waters, a particular group of excavates, the diplonemids revealed an immense diversity (Lara et al., 2009; Lukes et al., 2015). New massive sequencing technologies also revealed a large diversity in opisthokonts (del Campo et al., 2015).

Other systems like soils, however, have been by far not as deeply studied as the ocean. The high prevalence of fungal, plant, and metazoan sequences has been for a long time a major hindrance for studies on soil protist diversity (Lesaulnier et al., 2008). Still, previously unsuspected deep branching clades have also been discovered in soils, like the Opisthosporidia (also known as Rozellomycota) (Karpov et al., 2013; Lara et al., 2010). Recently, the development of high throughput sequencing has allowed obtaining high numbers of phylotypes, showing promising results in terms of overall microeukaryotic diversity (Geisen et al., 2015a). Relationships among organisms have been inferred using sound experimental designs and approaches such as co-occurrence networks (Lentendu et al., 2014). However, strong conclusions on the organisms' morphology and function can only be provided by direct observation or, even better, culturing. Therefore, a current challenge in eukaryotic microbiology is to identify the organisms hiding behind these environmental clades and to infer their ecological function. For this purpose several approaches have been used. Fluorescence in situ hybridization (FISH) in combination with scanning electron microscopy has been recently applied to the characterization of Paulinellida (Euglyphida testate amoebae) living in forest litter (Tarnawski and Lara, 2015). Bulk soil protists remain, however, widely inaccessible to FISH probes because of the large amount of soil particles unless organisms are large enough to be isolated individually (e.g. ciliates, macroscopic mycetozoa, and testate amoebae). Typically, naked amoebozoans are numerous in soils (Geisen et al., 2015a), and harbour a wide array of lifestyles and morphologies (Shadwick et al., 2009).

The amoebozoan environmental clade LKM74, named after the first clone encountered in an environmental DNA survey (van Hannen et al., 1999) is quite abundant and well distributed in soils (Corsaro and Venditti, 2013), but also present in freshwaters (Di Filippo et al., 2015; Richards et al., 2005) and peat bogs (Lara et al., 2011b). This clade has been repeatedly placed in the vicinity of *Dermamoeba algensis* in small subunit ribosomal gene trees (18S rRNA) (Corsaro and Venditti, 2013; Kudryavtsev and Pawlowski, 2015), although this relationship remained weakly supported. Despite its pervasive presence in many environments, the organisms have never been kept in culture and their morphology remains unknown. In this study, we describe a tiny naked amoeba isolated from the bulk soil of a coniferous forest by serial dilutions. We affiliated it to LKM74 based on 18S rRNA gene sequences, characterized its complex life cycle and feeding strategy, and documented its ultrastructure. Furthermore, we demonstrated that this species is a typical soil inhabitant by following related sequence reads in an environmental eukaryotic DNA survey of soils. This study was conducted in an experimental setup where pig cadavers were left to decompose and samples were taken at regular intervals to follow modifications of the microbial eukaryotic communities in the underlying soil during the process of decay.

3.2 Materials and Methods

3.2.1 Sample collection and identification

The original sample from which the species was described has been taken from a coniferous forest (dominated by *Picea abies*) near Neuchâtel, Switzerland (47°00.8530N; 6°55.9590E) in August 2011. Soil was suspended into phosphate buffer, and serially diluted into a low nutrient medium in 96 well plates as described in Lara et al. (2007). Active amoebae were transferred into culture flasks containing Page's Amoeba Saline medium amended with 1 g per 100 ml of Tryptone Soy Broth and *Escherichia coli* as food organism. Amoebae were subcultured regularly to obtain pure, monoprotistan strains and to lower the proportion of environmental bacteria. Cultures were kept at 12°C.

Cultures were observed using Utermöhl's plankton chambers with an inverted microscope (Olympus IX81). Cells were measured at different life stages, which were morphologically documented and photographed with light microscopy. Different life stages were also documented using Methyl blue ($C_{37}H_{27}N_3Na_2O_9S_3$) to stain cell walls and condensed cytoplasm.

We also used a full flask containing active and coccoid life stages for transmission electron microscopy (TEM). In that purpose, we pelleted cells from a thriving culture. Fixation, staining, and mounting were achieved as described in Lara et al. (2006). Observations were made on a Philips CM 100 transmission electron microscope.

3.2.2 Molecular analyses

In addition, a flask containing 10 ml of a thriving culture was used for DNA extraction. Cells were removed from the flask bottom with a cell scraper, and the resulting supernatant was placed into a Falcon tube and centrifuged at maximum speed during 20 min. The obtained pellet was placed into 200 µl of Guanidine thiocyanate buffer and nucleic acids were extracted following a protocol (Chomczynski and Sacchi, 1987) adapted after Lara et al. (2007). We amplified the 18S rRNA gene using primers EK 82F (5′-GAA ACT GCG AAT GGC TC-3′) and EK 1498R (5′-CAC CTA CGG AAA CCT TGT TA-3′) in a total volume of 30 µl with an amplification profile consisting of 4 min at 95°C followed by 35 cycles of 30 s at 94°C, 30 s at 55°C, and 1 min 30 s at 72°C with a final elongation of 10 min at 72°C. Sequencing was carried out using a BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit (PE Biosystems, Genève, Switzerland) and analysed with a ABI-3130XL DNA sequencer ABI PRISM 3700 DNA Analyzer (PE Biosystems). The sequence was deposited in GenBank with the following accession number: KX687875.

3.2.3 Phylogenetic analysis

The obtained sequence was placed in an alignment containing various sequences from Discosea (with an emphasis on Longamoebia as defined in Smirnov et al. (2011b)) derived either from isolated cells or cultures, or from environmental clone sequences. The root was placed on Vannellida. The alignment is available from the authors upon request. We build a maximum likelihood phylogenetic tree using the RAxML algorithm (Stamatakis et al., 2008) as implemented on the web server "http://embnet.vital-it.ch/raxml-bb", and evaluated the robustness of the nodes by bootstrapping.

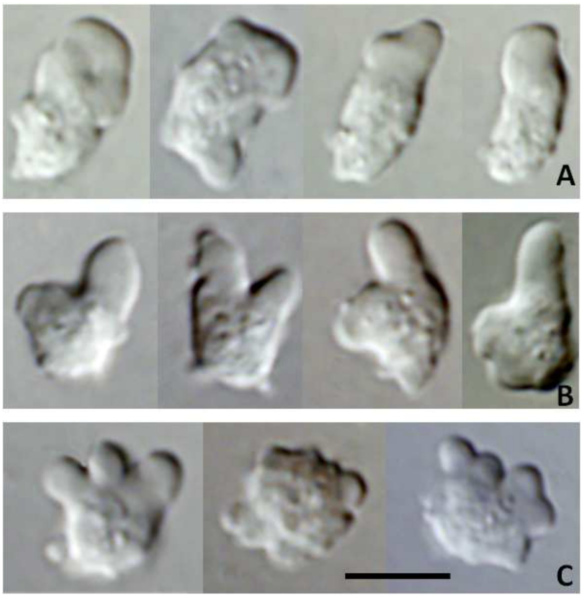

Figure 3.1 – Different morphologies of active trophozoites. (A) Trophozoites showing a typical lingulate shape. (B) Trophozoites when changing direction. (C) Stationary form. Scale bar: A–C = 5 µm.

3.2.4 Screening of high throughput sequencing data

The experimental setup from which the soil samples were retrieved has already been published in Szelecz et al. (2014). Briefly, three pig (*Sus scrofa*) carcasses were left to decompose during 1,051 d in a forest and soil was sampled at different time steps (21 samples in total). Three control soils (without pig) situated at a few meters of distance were sampled at the same time (24 samples). This experiment was performed in the context of a forensic study aiming at improving the estimation of the post mortem interval based on the observation of soil eukaryotic communities. DNA extraction, PCR of environmental DNA and sequencing of the V9 region of the 18S rRNA gene were performed as in Seppey et al. (2016). The obtained reads were pretreated as in de Vargas et al. (2015). The environmental sequences were aligned against a database constructed from publicly available V9 sequences of clones related to *Mycamoeba gemmipara* as determined with BLAST (GenBank: AY919786.1, AY919722.1, GQ861575.1, GQ861565.1, GQ861560.1), plus the sequence derived from the *M. gemmipara* culture. We determined empirically a threshold based on e-values to assess if a given environmental read belongs to the *M. gemmipara* group. The selected reads were then clustered into phylotypes using the SWARM v: 1.2.5 clustering algorithm (Mahé et al., 2014). To determine if the abundance of reads related to *Mycamoeba* changed significantly between control and cadaver plots, we performed Wilcoxon test on the data. All statistical analysis was done with R (Team, 2011).

3.3 Results

3.3.1 Light/electron microscopy observations

Active trophozoites were extremely small (maximum 7 µm in length when moving towards a single direction, and 2 µm width), had a flattened shape and lobose pseudopodia. When moving in a single direction, trophozoites had the typical lingulate shape observed in other Dermamoebidae (Fig. 3.1A), and conical-shaped pseudopodia can be observed when the amoeba changes direction (Fig. 3.1B). The stationary shape (Fig. 3.1C) showed small pseudopods radiating in several directions. The hyaline zone seemed to be generally only restricted on the extensive pseudopodia and was not visible laterally. No uroid or surface ridges were visible.

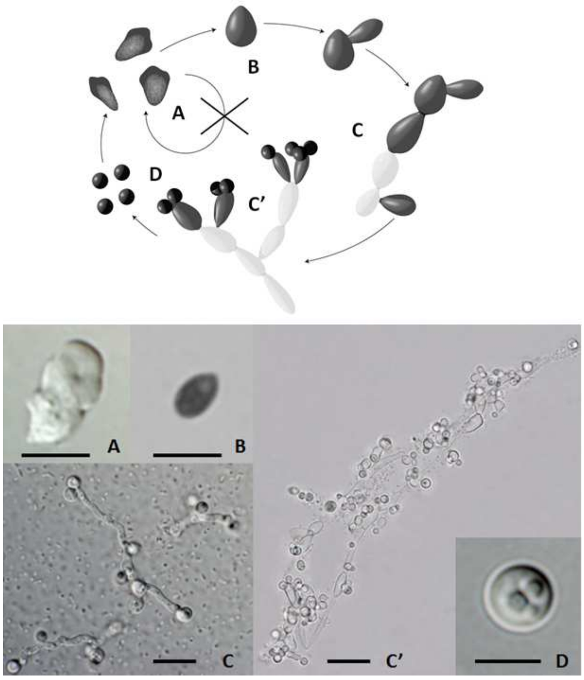

Figure 3.2 – Life cycle of *Mycamoeba gemmipara* including its four developmental stages. (A) Trophozoites. (B) Coccoid stage, stained with methyl blue. (C–C') bud like ramification. (D) Cysts released from branches of the pseudomycelium ramification. Scale bar: A–B, D = 5 µm, C–C' = 10 µm.

The life cycle of *M. gemmipara* comprised four stages (Fig. 3.2). (i) Active trophozoites were observed moving and feeding on *E. coli*. At this stage, no cell division could be observed. (ii) Cells then became rounded-elongated and smooth (coccoid phase), and stopped moving and feeding (Fig. 3.2B). Cytoplasm became highly condensed, and a cell wall was built. (iii) After about 24 h, the coccoid cells started budding several times successively, in a modular growth manner (Fig. 3.2C, 3.3A). The resulting structures grew, reaching up to 200 µm or more in length (Fig. 3.2C'). These structures in which cells cling together in chains correspond to a pseudomycelium as it has been described in several yeasts and bacteria. Old modules were devoid of cytoplasm, which probably migrated towards newly developed structures (Fig. 3.3B); still, cell walls without cytoplasm remained up to several months. Spherical dispersal cysts (Fig. 3.2D, 3.3B), were formed at the tip of the branched structures, the latter being finally degraded (iv). Dispersal cysts could then germinate into active amoebae immediately, but will do so most often when fresh medium is provided, thus closing the cycle. A movie (Movie S1) where the first steps of budding in coccoid cells can be observed is available on the journal website.

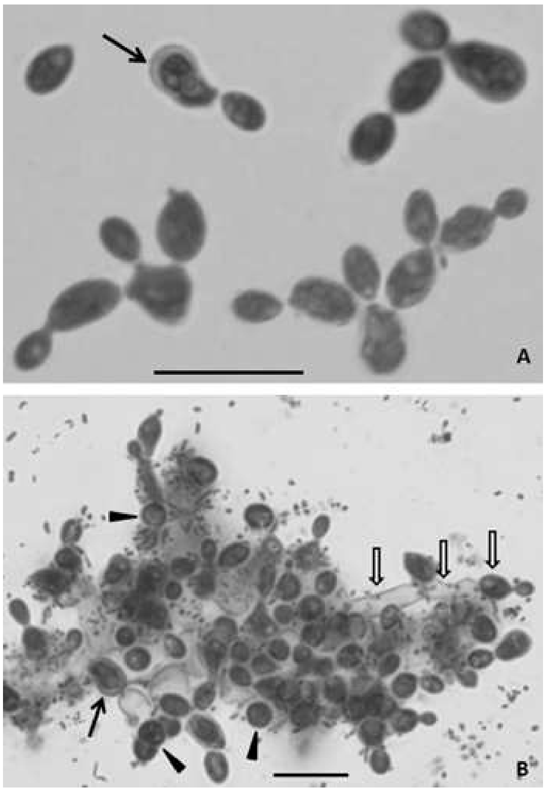

Figure 3.3 – Development of the coccoid cells. (A) Budding of coccoid (ovoid shaped) cells. Black arrow indicated cell wall. (B) Mature fructifications. Black triangles indicate mature cysts (spherical shape); a young coccoid cell filled with cytoplasm is indicated with a black arrow; old (empty) coccoid cells are shown with white arrows. Scale bar: A–B = 10 µm. All structures have been stained with methyl blue coloration.

Transmission electron microscopy pictures did not show any ornamentation on the cellular membrane (Fig. 3.4). Chromatin in the nucleus appeared inhomogeneous and presented several conspicuous zones with higher density within the nucleus. These zones were situated against the nuclear membrane, as well as in the centre of the nucleus, and a large nucleolus could be observed. Mitochondrial cristae were tubular. Several phagocytosis vacuoles could be seen containing bacteria at different stages of digestion. Small vesicles, probably containing digestive enzymes could be seen surrounding the bacterial cells.

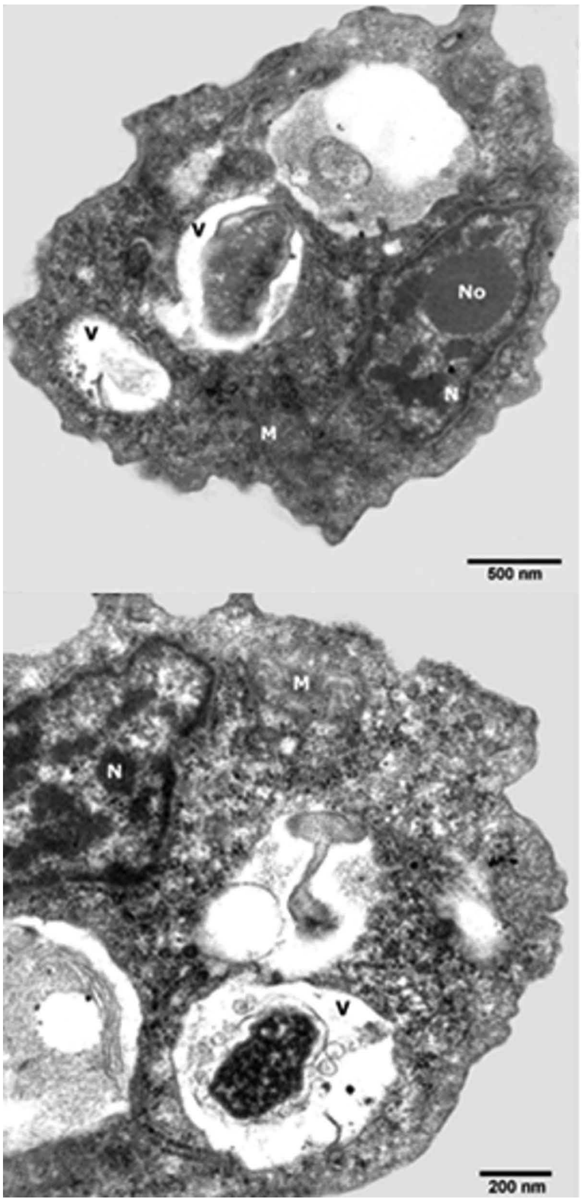

Figure 3.4 – TEM sections of *Mycamoeba gemmipara* showing: “M” the tubulicristate mitochondria; “N” the nucleus with the particular arrangement of the chromatin and “No” the nucleolus. Phagocytosed bacterial cells are visible in the vacuoles “V”, where lysosomes can be noticed in the process of releasing digestive enzymes.

3.3.2 Molecular phylogeny

Mycamoeba gemmipara branched within a group that comprised the original LKM74 environmental clone plus several other clones and unidentified isolates (Fig. 3.5). This clade, named genus *Mycamoeba* hereafter, received a moderate support (bootstrap value (BV) = 65). Furthermore, it branched robustly (BV = 87) at the base of a clade comprising only freshwater forms, the B1 clade sensu Corsaro and Venditti (2013). The monophyly of family Dermamoebidae was recovered with high statistical support (BV = 92), including genera *Dermamoeba*, *Paradermamoeba* and, now, *Mycamoeba*.

3.3.3 Presence of *Mycamoeba* reads in environmental DNA surveys

Our screen through the Illumina reads obtained from the soil from the forensic experiment allowed assigning 7,482 sequences to genus *Mycamoeba* (using an e-value alignment threshold below $1e^{-45}$) out of 25,579,257 environmental sequences, thus representing a total of 0.03% of all reads. These reads were present in all the 45 samples. The clustering of the environmental sequences resulted in a single Operational taxonomic unit (OTU) corresponding at 100% of identity with the original sequences of *M. gemmipara*. Figure 3.6 shows the sequences abundances distributions according to the control group and pig treatment. Number of reads is significantly lower in the pig treatment samples than in the control group (Wilcoxon test: p-value < 0.05).

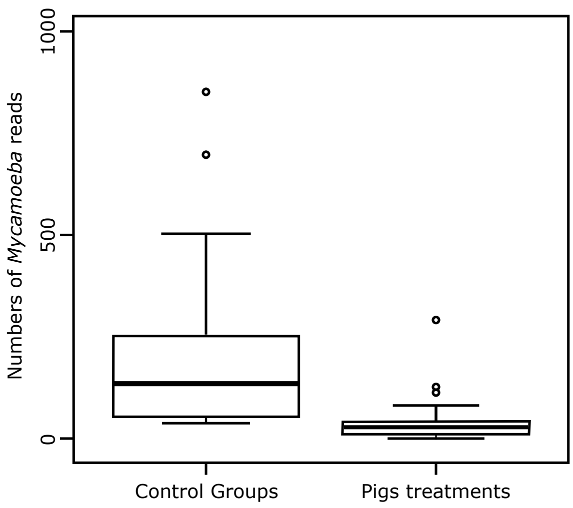

Figure 3.6 – Boxplot showing the number of V9 18S rRNA reads related to *Mycamoeba gemmipara* in control treatments and under the decomposing pig, suggesting a negative effect of cadaveric fluids on the amoebae.

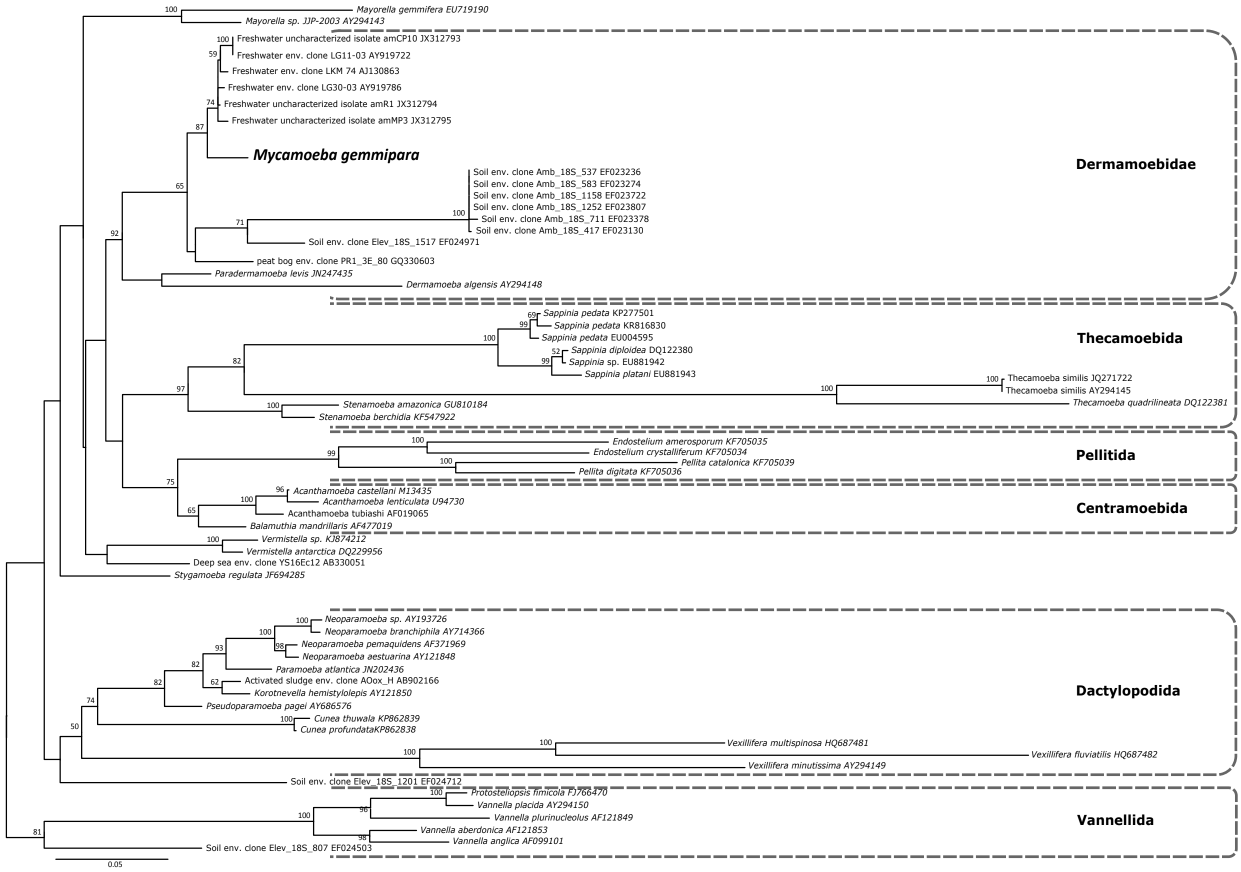

Figure 3.5 – Maximum likelihood phylogenetic tree based on complete 18S rRNA sequences of several Longamoebida, showing the position of *Mycamoeba gemmipara* within Dermamoebidae. The tree was built using the RAxML algorithm (Stamatakis et al., 2008).

3.4 Discussion

Mycamoeba gemmipara is the only described representative of environmental clade LKM74. Its size, clearly below 10 µm, places it among the smallest known amoebozoans. *Parvamoeba rugata* (Rogerson, 1993), *Paravanella minima* (Kudryavtsev, 2014) and *Sapocribrum chiconteaguense* (Lahr et al., 2015b) may be shorter in average but the thinner and flattened (lingulated) shape of *M. gemmipara* gives it a smaller biovolume. Coccoid forms measure a little above 3 µm. The size of *M. gemmipara* and its inconspicuous aspect, plus the fact that the amoeboid stage lasts a relatively short amount of time is probably the reason why it had never been detected previously. This small size may be characteristic of the whole LKM74 clade, as the isolates observed by Corsaro and Venditti (2013) were also smaller than 10 µm. Recently, there has been an increasing number of descriptions of nano-sized Amoebozoa, summarized in Kudryavtsev and Pawlowski (2015). As their 18S rRNA genes are often fast-evolving like in many other Amoebozoa, they remain undetected in DNA-based environmental diversity surveys, either because “general eukaryotic” primers fail in amplifying their 18S rRNA, or because their sequences cannot be assigned with confidence to a given group of eukaryotes. In that sense, application of systematic cultural approaches to soils and sediments are promising in revealing new lineages of Amoebozoa.

Our DNA survey demonstrated that *M. gemmipara* was a genuine forest soil inhabitant, as it has been found in all samples taken in the control group and under the pig carcasses. The site where *M. gemmipara* was isolated was located a few kilometres away from the experimental setup (Szelecz et al., 2014), and we can therefore suppose that this amoeba is widespread and probably abundant in the forests surrounding Neuchâtel (Switzerland). Nevertheless, it branched at the base of a clade comprising exclusively species from aquatic environments (Corsaro and Venditti, 2013). Other fruiting amoebae, such as mycetozoa and “protosteliales” were also recovered from aquatic environments, where trophozoites are probably active (Lindley et al., 2007). Therefore, it cannot be excluded that *M. gemmipara* lives in both freshwater sediments and soils. Its ability to build bud like ramifications in a liquid medium differs from most fruiting amoebae, although instances of myxogastrids completing their whole life underwater have been reported (Gottsberger and Nannenga-Bremekamp, 1971; Kappel and Anken, 1992). Moreover, this organism could be used as an indicator of certain environmental perturbation as it responded significantly to the effect of the cadaver decomposition (release of high concentration of nutrient, anoxic conditions; Fig. 3.6).

Order Dermamoebida has been first suggested by Cavalier-Smith et al. (2004). It hosted a single family, Thecamoebidae, with genera *Thecamoeba* and *Dermamoeba*. Further analyses based on 18S rRNA gene sequences suggested a separated grouping for *Thecamoeba* (Kudryavtsev et al., 2005), which formed the order Thecamoebida including also genera *Sappinia* (Brown et al., 2007), *Stenamoeba* (Smirnov et al., 2007), and *Vermistella* (Tekle et al., 2016). A new, emended order Dermamoebida comprised genera *Dermamoeba* and *Paradermamoeba* (Dermamoebidae), and *Mayorella* (Mayorellidae). To our knowledge, the monophyly of order Dermamoebida has never been recovered using 18S rRNA genes, as the grouping of genus *Mayorella* with Dermamoebidae remained either weakly supported (Pawlowski and Burki, 2009) or not at all (Corsaro and Venditti, 2013). Its inclusion within order Dermamoebida has been suggested on the base of the presence of a cell coat without wrinkles (Smirnov et al., 2011b); however, a recent multigene tree analysis suggested that *Mayorella* was rather related to the mainly marine amoebozoan clade Dactylopodida (Tekle et al., 2016).

The general lingulate shape of the locomotive form and the conical pseudopodia of *M. gemmipara* are typically found in all other members of Dermamoebidae (Kudryavtsev et al., 2011). Chromatin patterns in the nucleus of *M. gemmipara* are very distinctive, makes condensed granules all through the nucleus and especially near the borders of the nuclear membrane. All other known Dermamoebidae possess homogeneous chromatin in their nucleus with the exception of a large nucleolus. This includes *D. algensis* (Smirnov et al., 2011a) and *Paradermamoeba* (Smirnov and Goodkov, 2004). Genera placed within Dermamoebidae are characterized by a conspicuous cell coat. *Dermamoeba* and *Paradermamoeba* have a thick, highly structured cell coat (Smirnov et al., 2011a), either cuticle-like or consisting analogous glycostyle-like structures. In contrast, *M. gemmipara* does not possess any structure around its cell membrane, at least during its trophozoite life stage.

The peculiar life cycle of *M. gemmipara* has no known equivalents in Amoebozoa. While trophozoites actively ingest bacteria (as shown in TEM images; Fig. 3.4), the cell walls that are formed around coccoid cells preclude any phagocytosis. Nevertheless, the organisms undergo a considerable biovolume increase during pseudomycelium formation. Biomass incorporation can therefore only occur by osmotrophy. Its principle may remind the polyphyletic "protosteliales" (Shadwick et al., 2009), where a prespore stage (without cell wall) precede the formation of a stalk, which is used for spore dispersal (Olive, 1967). However, the simple shape of protostelids stalks differs considerably from the branched formations observed in *M. gemmipara* pseudomycelium (Fig. 3.2, 3.3). Moreover, in protostelids, stalks are formed from cysts when environmental conditions are degrading, whereas active amoebae (trophozoites) undergo cell division (Dykstra and Keller, 2000). In *M. gemmipara*, we did not observe any cell division at the trophozoite stage, which suggests that multiplication occurs only during coccoid budding process. Globally, the structures produced by *M. gemmipara* resemble those observed in Fungi. The reproductive mode by budding reminds strongly of yeasts such as *Saccharomyces cerevisiae*, and many fungi revert from yeasts to mycelial growth in a single organism (Rippon, 1982). As *M. gemmipara* original culture was regularly subcultured for about five years in the laboratory (thus generating dozens of replicates), we rule out the possibility that budding coccoid cells and pseudomycelia could be originated by fungal contaminants, which should have logically either disappeared or invaded our cultures. The main difference between the pseudomycelium observed in *M. gemmipara* and a classical mycelium is that its structures are not perennial and degrade once the cysts are released. *M. gemmipara* combines thus two life strategies: phagotrophy in a first stage of its life cycle (probably accumulating enough biomass to enter its next life stage) and osmotrophy afterwards. This strategy appears successful in freshwater and soil environments, and suggests a similar evolutionary pathway as in the Nucletmycea. These latter also evolved from amoeboid organisms, such as *Nuclearia*, towards mycelial growing organisms such as the true Fungi (Brown et al., 2009), which may have occurred also in soils or freshwater.

3.5 Classification summary

3.5.1 Taxonomic summary: *Mycamoeba gemmipara* nov. gen. nov. spec

Amoebozoa, Eudiscosea, Longamoebia, Dermamoebidae, *Mycamoeba, gemmipara*

3.5.2 Genus *Mycamoeba* gen. nov

Description: flattened amoebae with lobose or conical pseudopodia and a lingulate shape. Mitochondrial cristae are tubular. No glycocalyx or other ornaments on the cell surface. Possess a complex life cycle where active cells transform into coccoid stages, which undergo subsequent buddings, eventually turning into ramified structures (pseudomycelia) with spherical cysts in a terminal position on the ramifications. These structures disappear and cysts are released prior to germinating into active trophozoites. Etymology: From ancient Greek: myces; mould, fungus, in reference to the fungal-like structures appearing in the osmotrophic stage of its life-cycle, as well as its peculiar mode of reproduction, atypical for an amoeba, but reminding strongly yeasts. Sole species: *Mycamoeba gemmipara.* LSID numbers for the genus nomenclatural act urn:lsid:zoobank.org:act: 19381B2C-F0AA-43B5-AA5B-DED65E6F2B40 and for this publication urn:lsid:zoobank.org:pub: 20DF93D4-E26A-435F-AADB-10CA49CBB8AF.

3.5.3 *Mycamoeba gemmipara* sp. nov

Description: Small amoeboid cells up to 7 µm. Chromatin is not distributed in a homogeneous manner: there are conspicuous condensed regions pressed against the nuclear membrane and also in the centre of the nucleus. Ecology: has been detected to date only in forest soils. Seems to avoid important amounts of nitrogen (release of cadaveric fluids) or other perturbations generated by cadaver decomposition. Hapantotype: a culture has been deposited at the culture collection "Culture Collection of Algae and Protozoa". 18S rRNA gene sequence of *Mycamoeba gemmipara* (1,738 bp) was deposited in GenBank (KX687875). Etymology: *gemmipara*, as a reference to the reproduction mode (pario, giving birth in latin) through budding (gemma, a bud in latin). LSID numbers for the species nomenclatural act urn:lsid:zoobank.org:act: D0473730-785B-423A-808D-35DCBA38CE6D.

3.6 Acknowledgements

We thank the Swiss NSF for funding (project no. 31003A_163254). We are grateful to Daniel J.G. Lahr (University of Sao Paulo, Brazil) for his insightful comments, and also to an anonymous reviewer whose constructive comments helped us in improving the manuscript. We thank also Idlikó Szelecz for the soil samples. This work has been realized by Quentin Blandenier during his undergraduate studies.

3.7 Supplementary data

Additional Supporting Information may be found online in the supporting information tab for this article: Movie S1. Time-lapse showing the development of a ramification of *Mycamoeba gemmipara.*

Chapter 4

En garde! Redefinition of *Nebela militaris* (Arcellinida, Hyalospheniidae) and erection of *Alabasta* gen. nov.

En garde! Redefinition of *Nebela militaris* (Arcellinida, Hyalospheniidae) and erection of *Alabasta* gen. nov.

Quentin Blandenier[a,b,f], Clément Duckert[a,f], Fanny A. L. Kupferschmid[a], Anush Kosakyan[c], Edward A. D. Mitchell[a,d], Enrique Lara[b], David Singer[a,e]

Abstract

Molecular data have considerably contributed to building the taxonomy of protists. Recently, the systematics of Hyalospheniidae (Amoebozoa; Tubulinea; Arcellinida) has been widely revised, with implications extending to ecological, biogeographical and evolutionary investigations. Certain taxa, however, still have an uncertain phylogenetic position, including the common and conspicuous species *Nebela militaris*. A phylogenetic reconstruction of the Hyalospheniidae using partial sequences of the mitochondrial Cytochrome Oxidase Subunit 1 (COI) gene shows that *N. militaris* does not belong to genus *Nebela*, but should be placed in its own genus. The morphological singularities (strongly curved pseudostome and a marked notch in lateral view) and phylogenetic placement of our isolates motivated the creation of a new genus: *Alabasta* gen. nov. Based on their morphology, we include in this genus *Nebela kivuense* and *Nebela longicollis*. We discuss the position of genus *Alabasta* within Hyalospheniidae, and the species that could integrate this new genus based on their morphological characteristics.

Keywords: Amoebozoa; Biometry; DNA barcoding; Molecular phylogeny; Protist; Taxonomic revision

[a] Laboratory of Soil Biodiversity, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

[b] Real Jardín Botánico, CSIC, Plaza de Murillo 2, Madrid 28014, Spain

[c] Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branisovska 31, Ceske Budejovice 37005 Czech Republic

[d] Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland

[e] Department of Zoology, University of Sao Paulo, 05508-090 Sao Paulo, Brazil

[f] These authors contributed equally to the paper.

European Journal of Protistology, 2018, 66, 156-165

doi:10.1016/j.ejop.2018.08.005

4.1 Introduction

Testate amoebae are ubiquitous microorganisms in many terrestrial and freshwater to brackish aquatic environments (Meisterfeld, 2002; Tsyganov et al., 2016). Their ecological sensitivity, key functional role in microbial foodwebs (Jassey et al., 2012), and the good preservation of their test in sediments and peat designates them as excellent bioindicators for present and past environmental conditions (Amesbury et al., 2016; Mitchell et al., 2008), ecotoxicology (Amacker et al., 2018), and even forensic science (Seppey et al., 2016). Clear taxonomy is a prerequisite for sound ecological work. However, most original descriptions of testate amoeba species were based only on morphological characters (e.g. test composition, test and pseudostome dimensions) and the validity and phylogenetic position of many taxa are currently unclear. On one hand, several studies have revealed cases of cryptic or pseudo-cryptic diversity and proved that fine morphological differences of the shell indeed corresponded to distinct species (Dumack et al., 2016; Heger et al., 2011a; Kosakyan et al., 2013; Singer et al., 2015). On the other hand, phenotypic plasticity of the test has also been demonstrated experimentally (Mulot et al., 2017; Wanner, 1999; Wanner and Meisterfeld, 1994).

Hyalospheniidae (Amoebozoa, Arcellinida) are a widespread family of testate amoebae. They build their test from a proteinaceous matrix (*Hyalosphenia*) (Meisterfeld, 2002), which can be reinforced by self-secreted siliceous plates (*Quadrulella*) or by small silica elements taken from preys (e.g. *Nebela, Padaungiella*, etc.), a process referred to as kleptosquamy (Lahr et al., 2015a). They are mostly found in forest litter (Krashevska et al., 2018), fens and *Sphagnum* dominated peatlands (Singer et al., 2018), but also in more extreme environments like cryptogamic crusts in arid environments (Pérez-Juárez et al., 2017). They have been a key element in the debate over protist cosmopolitanism, showing several examples of biogeographic "flagship species", i.e. organisms with a conspicuous morphology only found in a limited part of the world (Heger et al., 2011a; Smith et al., 2008; Smith and Wilkinson, 2007). Molecular approaches on Hyalospheniidae opened new horizons in microbial biogeography; indeed, cryptic species showed non-overlapping geographic distribution areas (Heger et al., 2013), disproving the cosmopolitan hypothesis, for testate amoebae at least. Cryptic species also showed diverging ecological preferences for different micro-niches in *Sphagnum* peatlands (Singer et al., 2018), thus further demonstrating the need for improved taxonomy. Consequently, there is currently an active effort to do a systematic revision of testate amoeba taxonomy using combined molecular and morphological approaches

The first molecular surveys of arcellinid testate amoebae were focused on the Small Subunit Ribosomal RNA (18S rRNA) gene and permitted to place the Hyalospheniidae among the Arcellinida (Nikolaev et al., 2005). As the 18S rRNA gene is very conservative (Pawlowski et al., 2012), it is only of limited use for species delineation in testate amoebae. Other markers based on mitochondrial genes (Blandenier et al., 2017a; Hebert et al., 2003) were developed to overcome this limitation. The mitochondrial Cytochrome Oxidase Subunit 1 (COI) gene was shown to be useful to explore the cryptic diversity of hyalospheniid species and to resolve phylogenetic relationships within the family (Kosakyan et al., 2013). This resulted in splitting genus *Nebela*, the most species-rich genus of the family, into genera *Padaungiella* (Kosakyan et al., 2012), *Gibbocarina, Planocarina, Cornutheca, Longinebela* and *Nebela* (*sensu stricto*) (Kosakyan et al., 2016b).

However, the taxonomic validity and phylogenetic position of many hyalospheniids, including several common taxa, remain unclear. Our focus here is on *Nebela militaris*, an emblematic species in peatlands, where it typically occurs in relatively dry microhabitats (*Sphagnum* hummocks) with generally low pH (Diaconu et al., 2017; Väliranta et al., 2012). Its characteristic morphology and good preservation in peat designates it as a valuable bioindicator in ecological and palaeoecological studies of peatlands (Amesbury et al., 2016; Mitchell et al., 2008). Despite its unmistakable morphology, several morphologically similar taxa have been described, but as these descriptions are lacunar (e.g. poor biometry or illustrations, lack of differential diagnosis) their validity is unclear. Although *N. militaris* clearly belongs to the Hyalospheniidae, its morphological characters do not allow a clear assignation to one of the genera as defined by Kosakyan et al. (2016b). Its phylogenetic position among the Hyalospheniidae thus remains mysterious. Here, we characterized isolates of *Nebela militaris* from a Swiss peat bog based on detailed morphological observation and a single cell barcoding approach and we clarify its phylogenetic position and propose the new genus *Alabasta*.

4.2 Material and Methods

4.2.1 Single cell isolation

Cells corresponding to the original description of "*Nebela militaris*" (Penard, 1890) were isolated from *Sphagnum* mosses in a wooded peatland (*Pinus mugo uncinata*) in the Swiss Jura Mountains (Le Cachot bog, 47° 00 15.23 N, 6° 39 52.83 E). Testate amoeba cells were extracted from ca. 30 g of fresh *Sphagnum* after filtration through a 200 µm mesh filter. Cells were isolated individually under an inverted microscope (Olympus IX81) with a narrow pipette. Light microscopy pictures were taken at 600× magnification (Fig. 4.1) and morphometric characteristics (length, breadth and width of the aperture) were measured on 20 isolated cells (Fig. 4.2).

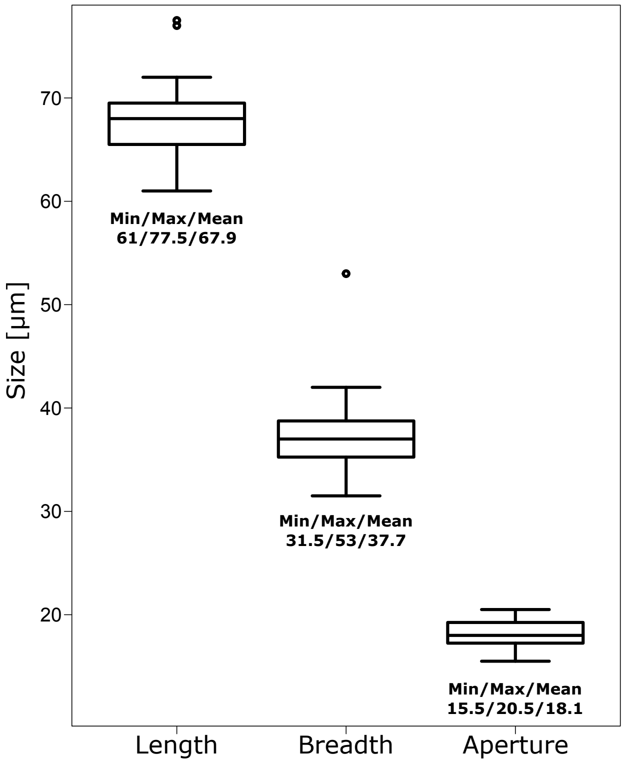

Figure 4.2 – Length, width and aperture of *Alabasta militaris* (n=20) taken from the same population as the four barcoded specimens (Le Cachot peatland, Swiss Jura Mountains).

4.2.2 DNA extraction

We used a specific procedure developed to extract DNA from a single cell of testate amoeba adapted after Chomczynski and Sacchi (1987). We prepared a thiocyanate guanidinium (TG) solution with the following protocol: We first dissolved 60 g of TG in 20 ml of EDTA (0.5M, pH 8) and 20 ml of H_2O under agitation at 65°C then when the solution reached room temperature, we added 5ml of Sarkosyl 10% (Na N-lauroylsarcisinate). Finally the solution was topped up to 100 ml with H_2O and filtered at 0.2 µm before being stored in the dark.

Each single cell was rinsed several times with distilled water to eliminate contaminants and then transferred into individual PCR tubes containing 50µl of TG solution. The tubes were heated at 65°C during 30 min, then 50µl of isopropanol was added and the tubes then left at 20°C during 12 h. The cleaning step of the DNA consisted of a first centrifugation at 15,000rpm during 20 min. The supernatant was removed before two additional washing steps with 180 µl of ethanol (70% and 99% respectively, 15,000rpm during 5 min). The residual ethanol is evaporated during 2 h under a fume hood.

4.2.3 PCR amplification and DNA sequencing

The PCR were processed by adding the components and reagents directly into the PCR tubes used for the extractions. Partial sequences of the mitochondrial COI gene were obtained using a nested PCR protocol. The first PCR was conducted using the Arcellinida-specific forward primer ArcelCox (5′-CAA AAT CAT AAA GAT ATT GGD AC-3′) (Kosakyan et al., 2012) and the eukaryote-general reverse primer HCO (5′-TAA ACT TCA GGG TGA CCA AAA AAT CA-3′) (Folmer et al., 1994). The PCR conditions were: Denaturation step at 95°C for 5 min, then 45 cycles with a denaturation step at 95°C for 15 s, an hybridation step at 43°C for 15 s, an elongation step at 72°C for 1 min, and a final elongation at 72°C for 10 min. The second PCR was done using hyalospheniid-specific primers HPCOIF (5′-GTT ATT GTT ACT GCT CAT GCC-3′) and HPCOIR (5′-ATA CAA AAT AGG ATC ACC TCC ACC-3′) (Gomaa et al., 2014) with the following conditions: Denaturation step at 95°C for 5 min, then 40 cycles with a denaturation step at 95°C for 15 s, an annealing step at 55°C for 15 s, an elongation step at 72°C for 1 min, and a final elongation at 72°C for 10 min. PCR products were purified using a Milipore kit and sequenced with an ABI3730XL DNA sequencer (Applied Biosystems) at Macrogen, Amsterdam NL. Sequences were deposited in GenBank with the following accession numbers: MH616621–MH616624. Light microscopy pictures of the DNA barcoded cells are shown in Fig. 4.1.

4.2.4 Phylogenetic analysis

The sequences obtained were 482 nucleotides long and were aligned using BioEdit (Hall, 1999) with

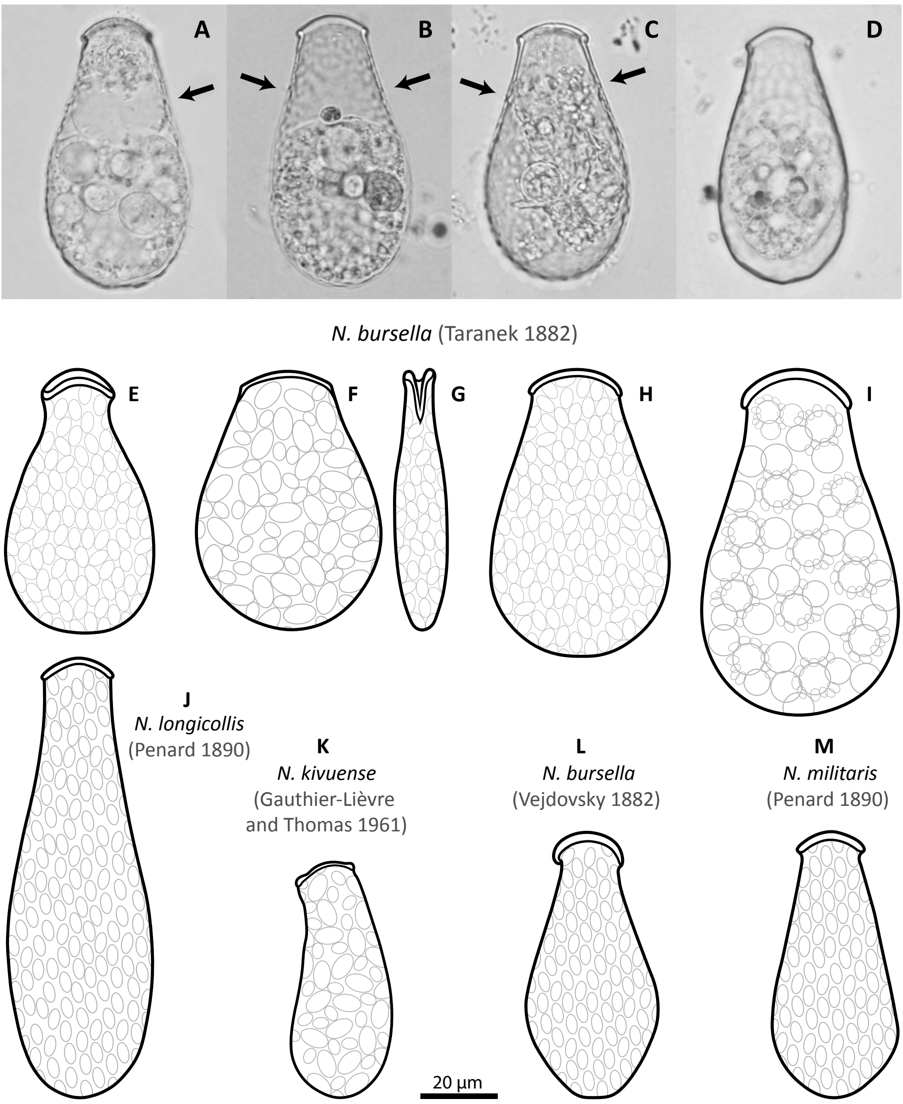

Figure 4.1 – A–D Light microscopy pictures of the barcoded specimens of *Alabasta militaris* (GenBank accession number A: MH616621; B: MH616622; C: MH616623; D: MH616624). The arrows point to the position of the lateral pores. E-I Reproduction of *Nebela bursella* sensu Taranek as it appears in Plate III Fig. 8, 9, 10, 7 and Plate IV Fig. 16 (Supplementary material), respectively, in Taranek (1882). J-M Reproduction of *N. longicollis, N. kivuense, N. bursella*, and *N. militaris*, respectively, as they appear in their original description. The size of E-I and L was approximated as no formal scale was given in the original drawing, but the proportions of E-I were kept. Scale bar=20µm.

an exhaustive reference database composed of COI sequences belonging to the Hyalospheniidae family in the GenBank database (Heger et al., 2013; Kosakyan et al., 2012, 2013, 2016b; Pérez-Juárez et al., 2017; Singer et al., 2015; Qin et al., 2016). Phylogenetic reconstruction was conducted using the CIPRES Portal (Miller et al., 2010). A maximum likelihood phylogenetic tree was built using the RAxML v.8.2.10 algorythm (Stamatakis, 2014) with the GTR+GAMMA+INVARIANT model and 1000 bootstraps. A Bayesian reconstruction was built with MrBayes (Ronquist and Huelsenbeck, 2003) using the GTR+GAMMA+INVARIANT model and ran on two independent chains for 200,000 generations sampled every 100 generations resulting in 4,000 trees, of which 25% were discarded as the burn-in. The trees were rooted using genera *Alocodera* and *Padaungiella* as outgroup based on previous Hyalospheniidae phylogenies (Kosakyan et al., 2016b).

4.3 Results and Discussion

4.3.1 The morphology of the studied cells is congruent with the original description of "*Nebela militaris*"

The cells described in this study correspond well to the original description of *Nebela militaris* (Penard, 1890). They have the same morphology (hyaline, yellowish or colourless and rigid test) and the shape of an ancient cannon (hence the species name given by Penard). The test is proteinaceous and can incorporate silica scales taken from prey. Lateral pores are frequently observed in the first anterior third of the test. The pseudostome is fan-shaped i.e. strongly curved in broad view, with a notch in narrow view, and shows a flare just before the pseudostome (Fig. 4.1A–D). We recorded the following measurements: Length: 61–77.5µm (mean 67.9µm), width: 31.5–53µm (mean 37.7µm), aperture width (pseudostome long axis): 15.5–20.5µm (mean 18.1µm) (Fig. 4.2, Table 4.1). However, our isolates differ slightly from the original description because of their wider test (Fig. 4.2, Table 4.1). Also, Penard (1890) originally described the pores on each side of the pseudostome to be infrequent but later stated that they are present in most cases

4.3.2 Clarification of the taxonomic status of *Nebela militaris*

The taxonomic history behind *Nebela militaris* is rather complicated as *N. militaris* is morphologically similar to two other taxa: *Nebela bursella* and *N. americana* var. *bryophila.* It is thus unclear whether *N. militaris, N. bursella* and *N. americana* var. *bryophila* represent distinct taxa and the confusion encompassing those names comforts the idea that they are merely the same morphospecies.

Nebela bursella was originally described by Taranek (1881) as *Nebela bursella* Vejdovský, described again by Vejdovský (1882) and was later reported in Taranek (1882) monograph. In this monograph, the illustrations that Taranek includes in the description of *N. bursella* are confusing, since some of them clearly correspond to the two first descriptions of *N. bursella* made by himself and Vejdovský while other depicted specimens are ambiguous. Indeed, while two illustrations (Plate III Fig. 8 and 12 by Taranek (1882) (Supplementary material 1), Fig. 8 here reproduced as Fig. 4.1E) resemble Vejdovský's and Taranek's original drawings of *N. bursella* with its characteristic curved pseudostome, the other illustrations correspond to specimens that are clearly broader. Based on modern knowledge of Hyalospheniidae systematics, these illustrated specimens should rather be assigned to another species within genus *Nebela* (Plate III Fig. 7 and Plate IV Fig. 16 by Taranek (1882) (Supplementary material 1), reproduced here as Fig. 4.1H, I respectively) or, possibly even, given the slit-like aperture, to genus *Heleopera* (Plate III Fig. 9–11 by Taranek (1882), reproduced here as Fig. 4.1F, G), resulting in one description possibly corresponding to three species. This situation led to the misinterpretation that *N. bursella* corresponded to the broader morphotype rather than to the cells illustrated by Vejdovský. This error was not corrected and *N. bursella* was ultimately considered as a synonym to *Nebela tincta* (Awerintzew, 1906). In his description of *N. militaris* Penard (1890) states that his new species could be considered by some as mere aberrant specimens of *N. bursella*, thus suggesting that he too was misled. However, unfortunately Penard did not explain what features could be considered as aberrant. He actually based his concept of *N. bursella* on Taranek's erroneous lumping and thus confused it with *Nebela tincta* (see Penard (1902)). Penard then described *N. militaris*, being certain that it was not only an aberrantly slender form of *N. tincta*, resulting in the redescription of *N. bursella* as *N. militaris.*

Nebela americana var. *bryophila* Van Oye (1933) was described as a smaller morphospecies of *Nebela americana* (Taranek, 1882). The dimensions and shape of the test are comparable to *N. militaris* but Van Oye (1933) noted the absence of lateral pores on his isolates. As the pores remain difficult to observe, we consider that the original description of *Nebela americana* var. *bryophila* is too lacunar to allow distinguishing it from *N. militaris* and we consider them as synonyms.

For these reasons, we propose to synonymize *N. militaris, N. bursella* and *N. americana* var. *bryophila.* Although we are aware that the name *Nebela bursella* precedes the others and thus should prevail, *N. militaris* has been consistently used since Penard's description and is well known to the scientific community working on testate amoebae. Moreover, we performed an extensive literature search and found no report for *N. bursella*, after 1964 (Sudzuki, 1964). For these practical reasons, we will invoke article 23.9.3 of the International Code of Zoological Nomenclature to keep *N. militaris* as the valid name even though it is technically a junior synonym.

4.3.3 Erection of the genus *Alabasta* and its phylogenetic position within the Hyalospheniidae

We obtained four identical partial COI sequences of 482 nucleotides from four different cells. The phylogenetic reconstruction does not place our sequences within the genus *Nebela*, but rather as a sister clade of the genus *Planocarina* (Fig. 4.3), this position being weakly supported (bs 48, pp 0.97). Moreover, the genetic distance between our sequences and the genetically closest species in genus *Nebela* (i.e. *N. flabellulum*, 18% of dissimilarity) is well above the barcoding gaps calculated for Hyalospheniidae and Amoebozoa in general (Kosakyan

Table 4.1 – Measurements and shape of the different species of *Alabasta* based on the literature.

Species	General shape of the test	Length (L) (µm)	Breadth (B) (µm)	L/B ratio	Pseudostome shape and width (µm)
Alabasta militaris	compressed, narrow piriform	50–82	25–46	1.8–2	deeply curved, 15-22
Alabasta kivuense	compressed, narrow piriform	49–64	18–26	2.4–2.7	very curved, 12-16
Alabasta longicollis	compressed, elongated piriform	120–140	40–45	3–3.1	deeply curved, n.a.

et al., 2012; Nassonova et al., 2010; Singer et al., 2018). Our isolates cannot be ascribed to *Planocarina* either due (1) to the genetic distance (ca. 20% of dissimilarity with *P. marginata*), and (2) by the fact that genus *Planocarina* is characterized by the presence of a keel on the posterior part of the test (Kosakyan et al., 2016b), whereas it is absent in *N. militaris*. This justifies the creation of a novel genus for *N. militaris* here-after named *Alabasta* that incorporates species with an elongated test and strongly curved pseudostome with a flare and a marked notch in narrow view.

4.3.4 Species to include in the genus *Alabasta* gen. nov.

Two species fit the morphological description of genus *Alabasta*: *Nebela kivuense* (Gauthier-Lièvre and Thomas, 1961) and *N. longicollis* (Penard, 1890). We thus formally move those species to genus *Alabasta*.

Alabasta (*Nebela*) *kivuense* (Fig. 4.1K) was described by Gauthier-Lièvre and Thomas (1961) from an area near Lake Edward in the Democratic Republic of Congo and was recently observed in two wetland-coniferous forest ecosystems in southern Ontario, Canada (Nicholls, 2015). *A. kivuense* may potentially be mistaken for *Alabasta* (*Nebela*) *militaris* as both species having overlapping length (Table 4.1), similar narrow piriform tests and deeply curved pseudostome, however the pseudostome of *A. kivuense* is less curved than *Alabasta militaris* (Fig. 4.1).

Alabasta (*Nebela*) *longicollis* (Fig. 4.1V) has been reported only on a few occasions, albeit from several geographical regions worldwide. *Alabasta longicollis* is larger and more elongated than *A. militaris* (almost twice as long and slightly wider, Table 4.1).

As the original descriptions present clear morphological distinctions between *A. militaris, A. kivuense* and *A. longicollis* we keep them as distinct species. However further investigations using molecular data and precise environmental descriptions are needed to investigate the relationships between these three morphologically similar species and clarify the true diversity within this genus.

4.3.5 Species morphologically related to the genus *Alabasta* gen. nov.

Besides *Alabasta* (*Nebela*) *longicollis* and *A.* (*Nebela*) *kivuense* there are several species that could be considered as morphologically related to *A. militaris*.

Hyalosphenia elegans (Leidy, 1879) and *H. insecta* (Harnisch, 1938) also share several common morphological characteristics with *Alabasta*. Furthermore, based on the 18S rRNA gene Lara et al. (2008) demonstrated that *H. elegans* does not branch with *H. papilio* as it is placed as a sister clade to genus *Nebela*. Thus, *H. elegans* and *H. insecta* could potentially also be included in genus *Alabasta*. However, two major differences between these species are 1) the inability of *H. elegans* and *H. insecta* to perform kleptosquamy as commonly observed in *Alabasta militaris*, *A. longicollis* and *A. kivuense* and 2) the presence of circular to oval depressions at the surface of the shell of *H. elegans* and *H. insecta* that are not observed in *Alabasta* species. Thus, we believe that *H. elegans* and *H. insecta* likely represent yet another clade (likely a new genus), but further molecular data are required to determine this.

4.4 Conclusion

Alabasta militaris is a cosmopolitan species that has been reported in the Northern and Southern Hemispheres and in high to low latitudes (e.g. Fernández et al. (2015); Golemansky (1962); Krashevska et al. (2017, 2018); Van Oye (1956)). However, these observations are based solely on rough morphology-based observation, and a study combining morphometry and single cell barcoding may well reveal an unsuspected diversity of cryptic species (Dumack et al., 2016; Kosakyan et al., 2012, 2013; Lara et al., 2011a; Singer et al., 2015). In parallel with findings in other Hyalospheniidae, these closely-related species may have different ecological optima (Singer et al., 2018) and may also have somewhat contrasted functional roles in microbial food webs (Geisen et al., 2018). As *Alabasta militaris* in its present definition has a narrow ecological tolerance, i.e. dry microhabitats in *Sphagnum* (Mazei and Bubnova, 2007), the distribution of its lineages is arguably less likely to be structured by geographic distance and barriers

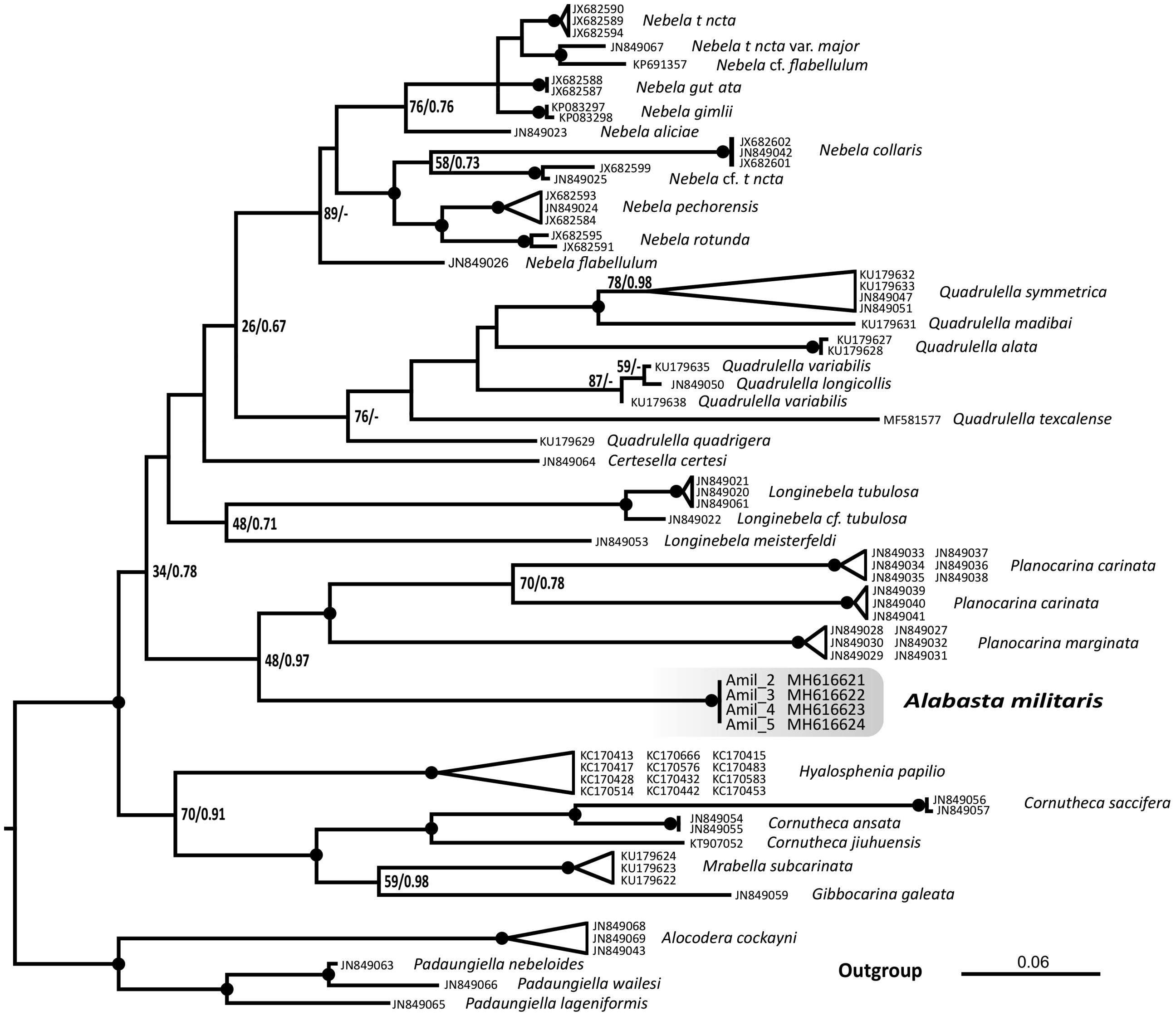

Figure 4.3 – Maximum likelihood phylogenetic tree of all COI gene sequences available on GenBank with the sequences of four single cells of *Alabasta militaris.* Numbers between branches correspond respectively to bootstrap values (bs) and posterior probabilities (p.p.) as calculated with Bayesian inference. The bs and p.p. values are indicated only for supported nodes (bs>50 and p.p.>70). Well supported nodes (bs>80 and p.p.>90) are marked with a dot.

to dispersal than larger species such as *Hyalosphenia papilio* (Heger et al., 2013). Finally, single cells transcriptomic studies of amoebae (Kang et al., 2017) may also provide new insight to the evolution of closely related species. The present study of *Alabasta* will pave the way to further studies in this iconic species of protist which may prove a useful model for future studies aiming to understand the general rules that shape ecological, biogeographical and evolutionary process in terrestrial protists.

4.5 Taxonomic actions

4.5.1 Description of new genus: *Alabasta* gen. nov. Duckert, Blandenier, Kosakyan and Singer

Taxonomic summary:

Arcellinida (Kent, 1880).

Hyalospheniidae (Schultze, 1877) Kosakyan et al. (2012).

Alabasta gen. nov. Duckert, Blandenier, Kosakyan and Singer.

Description: Test rigid, colourless or yellowish, elongated in broad view with a maximal width at about two thirds from the aperture and sides then tapering towards the aperture. Test proteinaceous often with incorporated silica scales taken from euglyphid preys. Pseudostome strongly convex with a flare (i.e. fan shaped) in broad view and a deep notch in profile. Lateral pores usually present at about one third of the distance from the pseudostome to the fundus.

Differential diagnosis: *Alabasta* can be distinguished from similar genera such as *Nebela* and *Longinebela* by its fan-shaped pseudostome (i.e. strongly curved with a flare) in broad view and a deep notch in profile.

Type species: *Alabasta militaris* comb. nov. (Penard, 1890) Duckert, Blandenier, Kosakyan and Singer.

Included taxa: *Nebela militaris* Penard (1890); *N. kivuense* Gauthier-Lièvre and Thomas (1961); *N. longicollis* Penard (1890).

Etymology: The name of the genus is derived from the greek word “” (alabaster), in connection with the diaphanous and yellowish aspect of the test of this genus. Also, it appears that *Alabasta* is the name of a fictitious city established in the middle of a desert in the manga One Piece (Oda, 1999), reminiscent of the ecological preferences of *Alabasta militaris* which is used as a dry indicator in peatlands monitoring.

LSID numbers for the nomenclatural act urn:lsid:zoobank.org:act: 7408C1E9-2DD6-4098-97DA-3CF743954B5F and for this publication urn:lsid:zoobank.org:pub:72E66220-2093-4765-8D79-413E9B52071D.

4.5.2 Key to the species

1. Smaller species <100µm, test narrow piriform in broad view

→ **2**

1*. Larger species: L=120–140µm, test elongated piriform in broad view, with an elongated neck.

A. longicollis

2. Wider species in broad view (L/B=1.8–2), pseudostome wide (15–22) µm

A. militaris

2*. Slender species, often laterally curved, in broad view (L/B=2.4–2.7) pseudostome narrow (12–16µm)

A. kivuense

4.5.3 Redefinition of *Nebela militaris* Penard 1890

Alabasta militaris comb. nov. Duckert, Blandenier, Kosakyan and Singer 1879 *Nebela collaris* (pars) in Leidy, Freshw. Rhiz. N. America, p. 147, pl. 22 Fig. 11, 12, 16 (misidentification). 1881 *Nebela bursella* Vejdovský in Taranek (1881) Sitzber. böhm. Ges. Wiss. p. 230, Fig. 2 1882 *Nebela bursella* Vejdovský in Thierische Organismen der Brunnenwässer von Prag 1882 *Nebela bursella* Vejdovský (pars) in Taranek (1882), 8, p. 36, pl. III, Figs. 8 and 12. 1890 *Nebela militaris* Penard, Mem. Soc. Phys. Hist. Nat. Genève 31: 164 (creation of *Alabasta* gen. nov. and transfer to the new genus) 1933 *Nebela americana* var. *bryophila* Van Oye, Arch. Naturgestch. Abt. B., N.F. 2 (overlapping morphology).

Updated description: Test rigid, colourless or yellowish, compressed in broad view, narrow piriform, reminding the shape of an ancient cannon (hence the species name) with a maximal width at about the first third from the posterior part and the sides of the test tapering towards the aperture. Test proteinaceous often incorporating silica scales taken from euglyphid preys. Lateral pores frequently observed at ca. one third of the distance from the pseudostome to the fundus. Pseudostome fan-shaped (i.e. strongly curved) in broad view, with a notch in profile, and a thick organic lip (Fig. 4.1.A–D). We recorded the following measurements (Fig. 4.2): Length: 61–77.5µm (mean 67.9µm), width: 31.5–53µm (mean 37.7µm), pseudostome (aperture) long axis: 15.5–20.5µm (mean 18.1µm). Dimensions based on previous observations: L=50–82µm, B=25–46µm. Pseudostome 15–22µm wide.

Differential diagnosis: *A. militaris* may be confused with *A. longicollis* and *A. kivuense*, from which it differs respectively by a shorter test or by a wider test and pseudostome (see Table 4.1). Also *A. kivuense* is often laterally curved.

Neotype: Due to the absence of a type, we declare the Specimen 516-2 isolated by Penard and mounted on a permanent slide, now deposited at the Natural History Museum of Geneva, Switzerland as the name-bearing type (https://commons.wikimedia.org/wiki/Category:Nebela_militaris#/media/File:Collection_Penard_MHNG_Specimen_516-2-1_Nebela_militaris.tif).

Type locality: Jura Mountains, Switzerland

Etymology: The name refers to its general shape which reminds of an ancient military cannon.

Habitat: *Sphagnum* mosses (relatively dry microhabitats such as hummocks), brown mosses, litter, sediments (rare). Geographical distribution: Apparently cosmopolitan

4.6 Acknowledgements

This work was funded by a SWISSBOL (Swiss Barcoding of Life initiative) project to D.S. D.S., Q.B. and E.L. were also supported by the Swiss NSF (310003A_143960 and 31003A_163254). EL is funded by a “Atracción de talentos de la Comunidad de Madrid” project (2017-T1/AMB-5210). We thank the anonymous reviewers for their comments on the manuscript, and the International Commission on Zoological Nomenclature for their help on taxonomic issues.

4.7 Supplementary data

Figures and descriptions of *Nebela bursella, Nebela militaris, Nebela kivuense, Nebela longicollis* and *Nebela americana* var. *bryophila* as they appear in Vejdovský (1882); Taranek (1882); Penard (1890); Gauthier-Lièvre and Thomas (1961); Van Oye (1933). The following is Supplementary data to this article: https://doi.org/10.1016/j.ejop.2018.08.005.

Chapter 5

Morphological and genetic diversity of genera *Apodera* and *Alocodera* from the Southern Hemisphere

Morphological and genetic diversity of genera *Apodera* and *Alocodera* from the Southern Hemisphere

---

QUENTIN BLANDENIER[a,b,h], CLÉMENT DUCKERT[a,h], LUCIANA BURDMAN[a], MILCHO TODOROV[c], ANUSH KOSAKYAN[d], DAVID SINGER[a,e], DANIEL J.G. LAHR[e], JANET WILMSHURST[f], MICHELLE MCKEOWN[f], MATT MCGLONE[f], EDWARD A. D. MITCHELL[a,g], ENRIQUE LARA[b]

Abstract

Testate amoebae are good model organisms for studying microbial biogeography. Several flagship taxa, among which the genera *Apodera* and *Alocodera*, are mostly limited to former Gondwanaland. Their true diversity has however not yet been assessed using molecular tools, but, given the evidence for widespread cryptic diversity within Hyalospheniformes it is likely that the diversity of these two genera is high and that some cryptic or pseudocryptic species have limited geographical distributions. We investigated the molecular diversity of genera *Apodera, Alocodera* and the closely related but cosmopolitan genus *Padaungiella* using single cell DNA COI barcoding. As anticipated we found a high genetic diversity within each genus, including a novel morphologically highly conspicuous *Apodera*-like species broadly distributed across New Zealand peatlands characterized by a keel. The genetic divergence among these species was considerably higher than the barcoding gaps reported for other Hyalospheniformes such as genus *Nebela*. The phylogeographical pattern of genus *Alocodera* showed a strong biogeographical signal with separate clades recorded in South America and New Zealand. Furthermore, all European *Padaungiella* isolates branched together in a robust clade while other members of this genus were scattered across the phylogenetic tree. As anticipated, our study revealed the existence of a considerable diversity within genera *Apodera, Alocodera* and *Padaungiella* as well as complex geographical distribution patterns. This diversity includes both cryptic and conspicuous species and calls for in-depth analyses combining morphological, ecological and molecular analyses of these fascinating model organisms for microbial biogeography.

[a] Laboratory of Soil Biodiversity, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland
[b] Real Jardín Botánico, CSIC, Plaza de Murillo 2, Madrid 28014, Spain
[c] Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, 2 Gagarin St., 1113 Sofia, Bulgaria
[d] Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branisovska 31, Ceske Budejovice 37005 Czech Republic
[e] Department of Zoology, University of Sao Paulo, 05508-090 Sao Paulo, Brazil
[f] Manaaki Whenua – Landcare Research, 54 Gerald Street, 7608 Lincoln, New Zealand
[g] Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland
[h] These authors contributed equally to the paper.

5.1 Introduction

The existence or not of geographic patterns in microorganisms is an old debate in biology, going back to the 19th century. The German protistologist Christian Gottfried Ehrenberg identified diatoms in dust samples collected on the Beagle by Charles Darwin while sailing in the tropical Atlantic and inferred a South American origin for some species (Darwin, 1846). Darwin was convinced that, just like macroscopic organisms, microbes presented distinctive distribution patterns, but also had a formidable capacity for dispersal (O'Malley, 2009). This capacity, together with their astronomic population sizes led to the idea that the distribution of microorganisms is only restricted by environmental conditions, summarized by the famous tenet "everything is everywhere, but, the environment selects" (Baas Becking, 1934). This premise was later used at the turn of the 21th century to explain the distribution of microbial diversity, putting protists at the centre of the debate (Finlay and Fenchel, 1999). A central corollary of this paradigm was that, in absence of allopatric speciation, microbial diversity should be low (Fenchel and Finlay, 2003; Finlay et al., 2004). Opposing to this viewpoint, some taxonomists and biogeographers provided several examples of species that had been found only in some regions of the globe, and whose conspicuous morphology could not be overlooked by researchers (Foissner, 2004, 2005). The validity of these "biogeographical flagship species" became a central element in the debate; amongst them, perhaps the most iconic species being the arcellinid testate amoeba *Apodera vas* (Foissner, 2006).

Apodera vas (Amoebozoa; Arcellinida; Hyalospheniformes) is a large species (typically between 120-250 µm), considered as monotypic (Zapata and Fernandez, 2008), characterized by having a test clearly separated into two chambers by a deep constriction around the base of the "neck". It is typically found in mosses, especially *Sphagnum*, and wet forest litter under subantarctic temperate to tropical climates, but has never been found beyond the Cancer tropic desert belt, despite the widespread presence of suitable environments further north (Smith and Wilkinson, 2007). Several other Hyalospheniformes taxa such as the monotypic genus *Alocodera* share a similar distribution (Smith et al., 2008), while the closely-related genus *Padaungiella* is considered cosmopolitan (Lara et al., 2008).

The application of DNA barcoding to study the diversity of testate amoebae, especially using the cytochrome oxidase subunits 1 (COI) marker (Hebert et al., 2003), has revealed the existence of high levels of cryptic diversity within nearly all investigated species of Hyalospheniformes (Kosakyan et al., 2012). Detailed analyses of the morphology of these barcoded organisms allowed the description of several overlooked pseudo-cryptic species (Kosakyan et al., 2013; Singer et al., 2015), which differed in their ecological preferences (Singer et al., 2018). A phylogeographic study of *Hyalosphenia papilio*, a common species in northern *Sphagnum*-dominated peatlands, revealed contrasted geographical distributions (Heger et al., 2013) that could be explained by historic events such as Pleistocene glaciations (Singer et al., 2019). An accurate taxonomy, backed by molecular data is thus key for understanding the biogeography of testate amoebae and protists in general (Heger et al., 2009). In particular, the systematics of *Apodera* and related genera needs to be better characterized by means of single cell barcoding in order to re-evaluate their "biogeographic flagship" status.

Here, we combined detailed morphological and molecular characterization of *Apodera, Alocodera* and *Padaungiella* specimens from different environments, substrates and locations mostly in New Zealand in order to search for, (1) a possible hidden diversity within species, notably testing the monotypy of the genera, (2) the adequation of present taxonomy to phylogenetic relationships in order to conduct future biogeographical studies on these Southern Hemisphere species and (3) the existence of probable biogeographical patterns.

5.2 Material and Methods

5.2.1 Cells isolation and morphological characterization

Cells morphologically assigned to genera *Apodera, Alocodera* and *Padaungiella* were retrieved from different environments and substrates from New Zealand, Brazil, Bulgaria and Spain (Table 5.1). Living cells were isolated using a narrow pipet under an inverted microscope (Blandenier et al., 2018). The cells were then photographed at 400x magnification with a light microscope (LM) and finally transferred into PCR tubes for DNA extraction. Empty tests were prepared and documented for scanning electron microscopy (SEM) as described previously by Todorov and Golemansky (2007). In addition, we also retrieved pictures from the literature (Kosakyan et al., 2012; Lahr et al., 2015a; Lara et al., 2008; Mitchell and Meisterfeld, 2005; Yeates

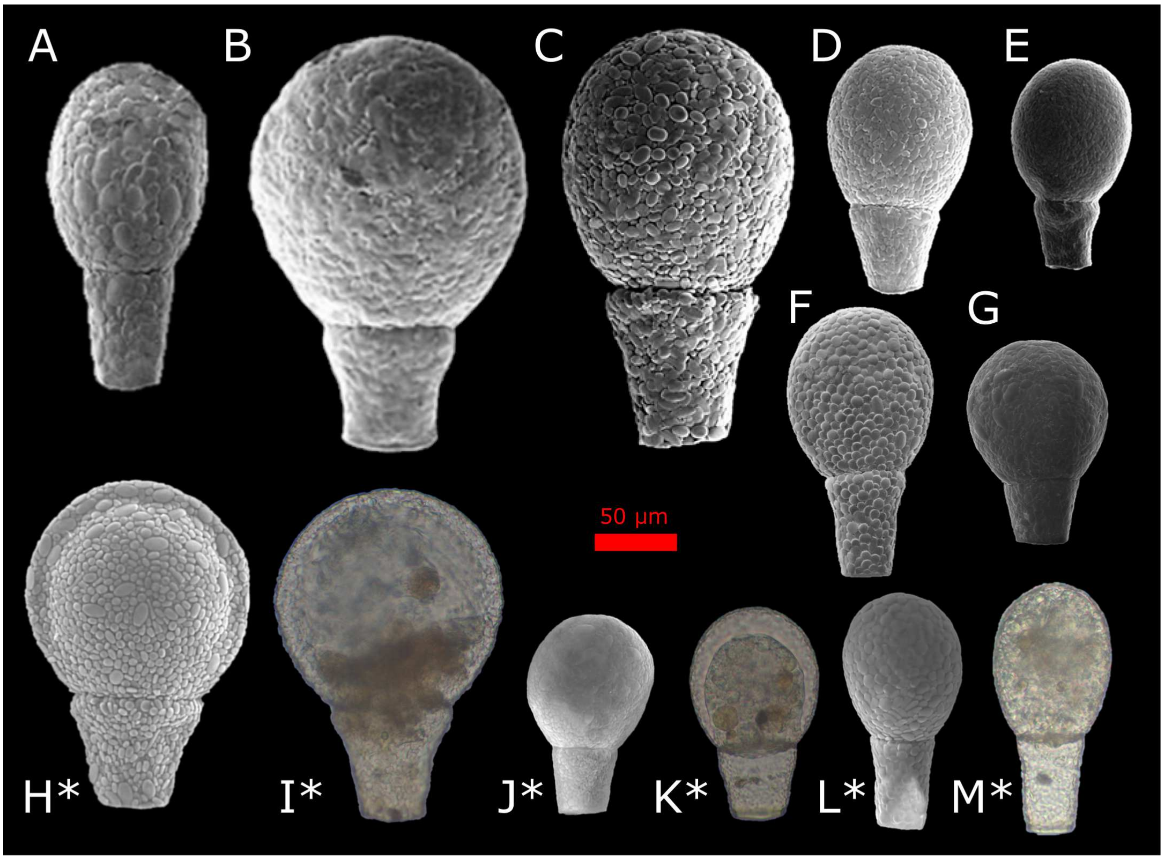

Figure 5.1 – *Apodera vas* SEM pictures from the literature (**A-G**): **A-C.** from Chile (Zapata and Fernandez, 2008), **D.** Tanzania (Mitchell and Meisterfeld, 2005), **E.** Marion Island (Lara et al., 2008), **F.** New Zealand (Yeates and Foissner, 1995) and **G.** Argentina (Lahr et al., 2015a). SEM and LM pictures of *Apodera* specimens from New Zealand from current study (**H-M***): **H-I.** *Apodera* sp 3 (EM2540 and EM2540-19), **J-K.** *Apodera* sp 1 (EM2543 and EM2543-22) and **L-M.** *Apodera* sp 2 (EM2581, EM2556-1) (Table 5.1).

and Foissner, 1995; Zapata and Fernandez, 2008) in order to compare cell dimensions and shapes with our observations. LM and SEM pictures of sufficient quality (Fig. 5.1-5.5) were used for measuring the length and breadth as well as the width of the aperture of each specimen.

5.2.2 DNA extraction, PCR amplification and DNA sequencing

DNA extraction was performed with a thiocyanate guanidinium (TG) solution following the protocol of Blandenier et al. (2018) adapted after Chomczynski and Sacchi (1987). PCR of partial mitochondrial cytochrome oxidase I (COI) gene were mostly processed using a nested PCR protocol. The first PCR was conducted with the eukaryote-general primers LCO (GGT CAA CAA ATC ATA AAG ATA TTG G) and HCO (TAA ACT TCA GGG TGA CCA AAA AAT CA) (Folmer et al., 1994). When the first PCR was not successful a second PCR was done using combinations of various primers including ArcelCOI and ApoCOI (Kosakyan et al., 2012), as well as APO2F (TGG AAT TAG CAT ATC MGG AAT T), a new forward primer specific to the group of *Apodera, Alocodera* and *Padaungiella*. All primers combinations and PCR conditions are detailed in Supplementary material 1 *(This table will be given after the acceptation of the manuscript).*

Two positives PCR products were cloned (Supplementary material 1) as described in Blandenier et al. (2017a), whereas the others were directly purified using a Milipore kit and sent for sequencing with an ABI3730XL DNA sequencer (Applied Biosystems) at Macrogen, Amsterdam NL. The sequences were deposited in GenBank with the following accession numbers: *(Numbers will be given after the acceptation of the manuscript).* Light microscopy pictures of the corresponding cells are shown in Supplementary material 2 *(All pictures will be given after the acceptation of the manuscript).*

5.2.3 Phylogenetic analysis

All sequences obtained were aligned and corrected manually using BioEdit (Hall, 1999) together with an exhaustive reference database composed of three sequences of genus *Alocodera* and three of

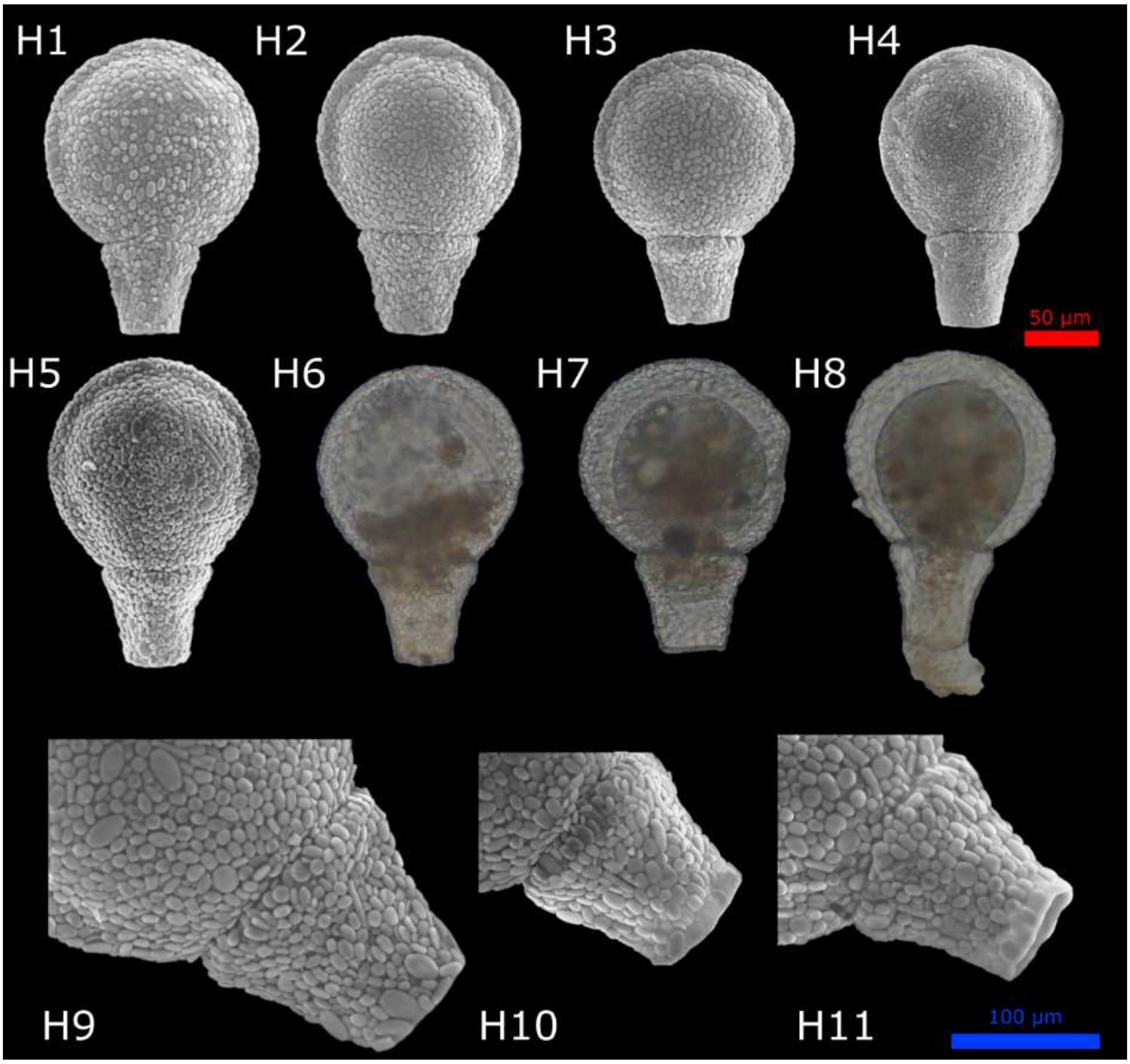

Figure 5.2 – SEM and LM pictures of several specimens of *Apodera* sp 3 presenting a characteristic keel expanding from the neck until the posterior extremity. Pictures **H6-H8.** correspond to barcoded individuals EM2540-19, EM2540-14 and EM2540-11. These specimens were all extracted from sample EM2540 (Mt. Taranaki, North Island, New Zealand, Table 5.1).

genus *Padaungiella* (Kosakyan et al., 2012) retrieved from the GenBank database (Table 5.1). We conducted the phylogenetic reconstruction on the complete LCO-HCO fragment (i.e. 655 bp) using the CIPRES Portal (Miller et al., 2010). A maximum likelihood phylogenetic tree was built using the RAxML v.8.2.10 algorithm (Stamatakis, 2014) with the GTR + GAMMA model and 1000 bootstraps.

5.3 Results

5.3.1 Morphology of specimens from genera *Apodera, Alocodera* and *Padaungiella*

We have compiled every available picture of *Apodera* (Fig 5.1), *Alocodera* (Fig 5.3) and *Padaungiella* (Fig 5.4) from the literature, with distribution and morphometrical data. We compared these data with the new pictures obtained from these three genera (Fig 5.1-5.5). The general morphology of all specimens of each of the three genera fits the original description, i.e. an elongated neck for *Padaungiella* taxa, associated with a septum (bi-cameral organization) or a more or less pronounced lateral constriction at the base of the neck for *Apodera* and *Alocodera* respectively.

Specimen in genus *Apodera* possessed variation in test size which can reach 100% with the largest specimen exceeding 250 µm in length (for instance in specimens B and C) and the smallest less than 120 µm (e.g. G and J) (Fig 5.1). Some specimens were particularly roundish (L/B ratio below 1.5 for specimen B and H), whereas others were much slender (L/B ratio of 2.2 for specimen L and M). The larger morphotypes from New Zealand were systematically char-

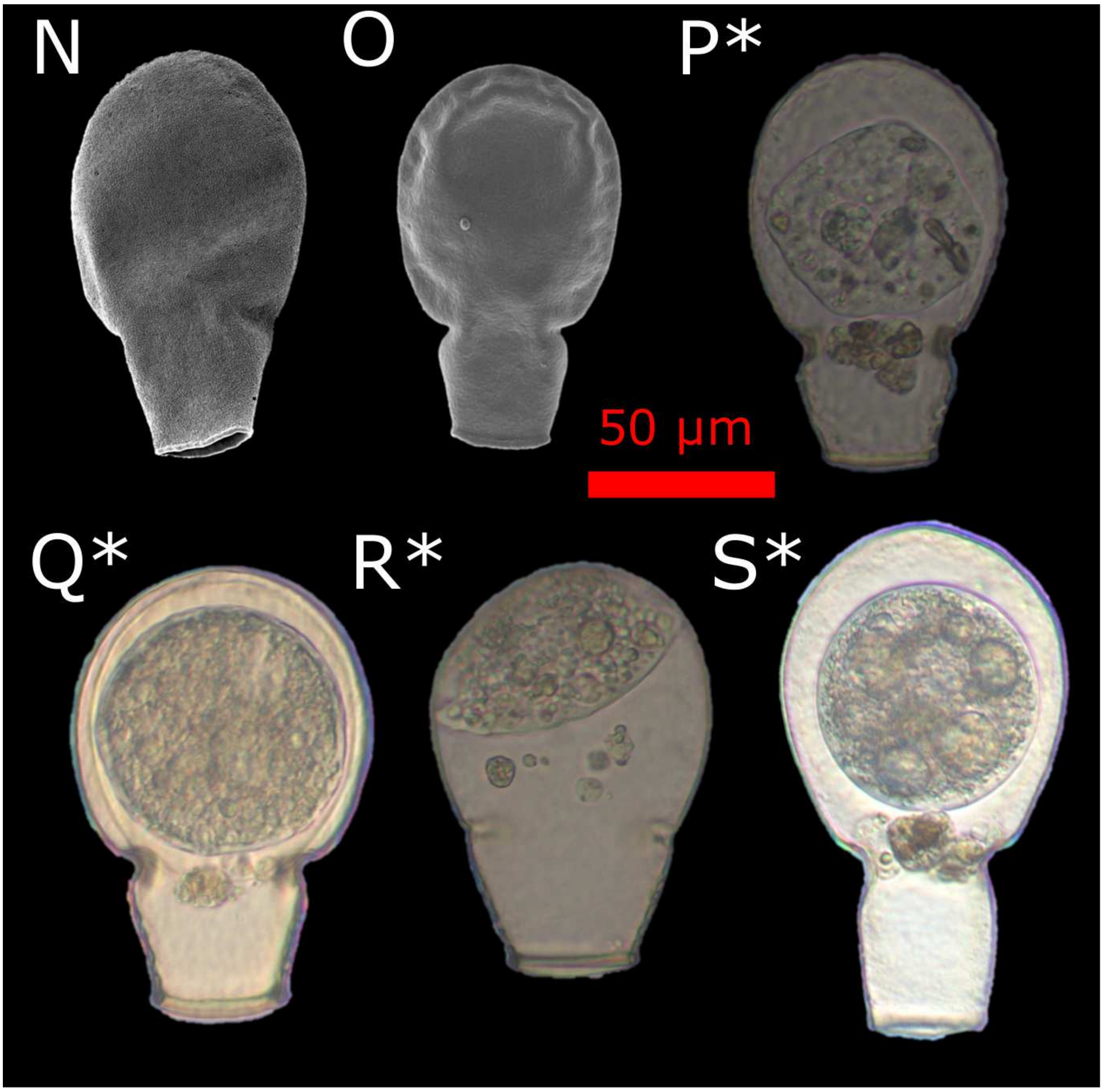

Figure 5.3 – *Alocodera* SEM pictures from Kosakyan et al. (2012) (**N-O**): **N.** *Alocodera* sp 1 from Chile and **O.** *Alocodera cockayni* from Argentina. New Zealand LM pictures from the current study (**P-S***): **P-Q.** *Alocodera* sp 2 (EM2539-3 and EM2638-7), **R.** *Alocodera* sp 3 (EM2583-6) and **S.** *Alocodera* sp 4 (EM2670-1) (Table 5.1).

acterized by a keel around the posterior part of the tests starting from the constriction (specimens H and I). Such a structure which is very similar to the keel found in genus *Gibbocarina* had to our knowledge, never been observed in this genus. Some keeled specimens extracted from the same sample are presented in figure 5.2 to illustrate the range of morphological variability within a population.

In contrast to genus *Apodera*, the size range of genus *Alocodera* was much narrower, i.e. between 100-120 μm of length, with the exception of one longer specimen of ca. 140 μm (specimen S). However we observed substantial morphological variability for the test breadth (ranging from roundish to slender; L/B ratio between 1.5 and 1.9, specimens Q and S respectively; Fig 5.3), the depth of the lateral constrictions (ranging from almost imperceptible, e.g. specimen R, to more pronounced, as in the type) and the shape of the neck (ranging from almost parallel sides, specimen S, to piriform, specimen R). Finally, we observed a granular texture in some specimen (*Alocodera* sp 1, specimen N; (Kosakyan et al., 2012)) while the shell of *Alocodera* is generally transparent and amber coloured similar to *Hyalosphenia papilio*.

All *Padaungiella* species share a similar general morphology, with the notable exception of the test with a circular cross section and presence of large agglutinated mineral particles in *P. nebeloides* (Todorov et al., 2010; Todorov and Bankov, 2019) (Fig 5.4). Some studied specimen were on the range size of *P. lageniformis* (i.e. 100-140 μm) (Penard, 1890), while the dimension of the others correspond to *P. wailesi* (i.e. 75-100 μm) (Deflandre, 1936). The dis-

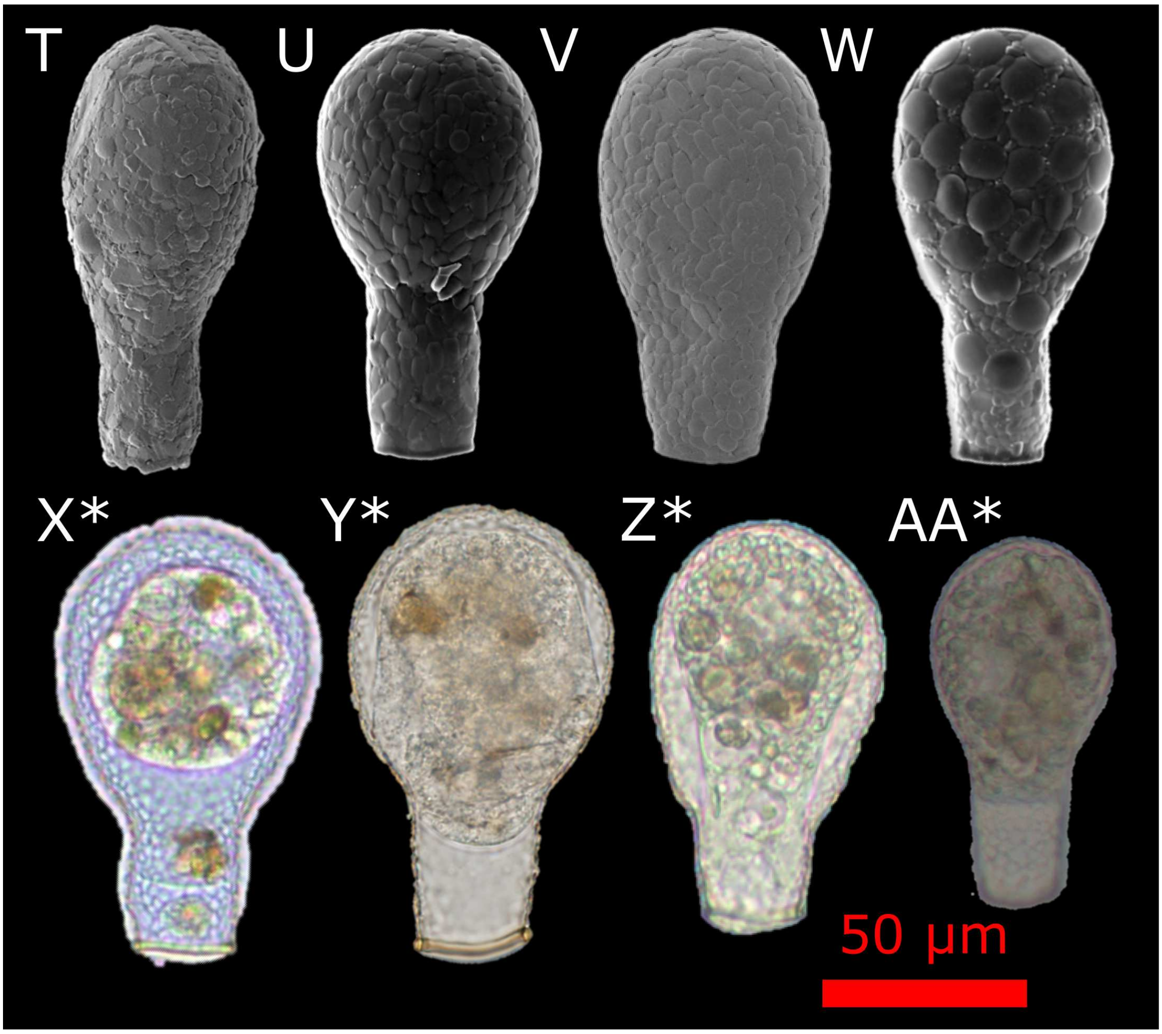

Figure 5.4 – *Padaungiella* SEM pictures from the literature (**T-W**): **T.** *Padaungiella nebeloides* from France, **U.** *Padaungiella lageniformis* from Bulgaria and **V.** *Padaungiella wailesi* from Switzerland (Kosakyan et al., 2012), and **W.** *Padaungiella wailesi* from Bulgaria (Todorov and Bankov, 2019) which has been barcoded in the present study as *Padaungiella wailesi* 1 BG. *Padaungiella* LM pictures from the current study (**X-AA***): **X.** *Padaungiella lageniformis* 1 ES (ESMA2-6), **Y.** *Padaungiella lageniformis* 2 BR (Pad4-10), **Z.** *Padaungiella lageniformis* 3 NZ (EM2534-7) and **AA.** *Padaungiella wailesi* 2 NZ (EM2543-1) (Table 5.1).

tinction between the neck and the main body form a gradient between almost absent in *P. wailesi* CH and *P. lageniformis* 3 NZ (i.e. V and Z) to a distinct subparallel neck in *P. lageniformis* BG, *P. wailesi* 1 BG and *P. lageniformis* 2 BR (i.e. U, W and Y). A lateral constriction separating the neck from the posterior part is the synapomorphy of the genus *Alocodera*, which separates it best from the genus *Padaungiella* (Jung, 1942). Specimens of *Padaungiella lageniformis* 2 BR observed in the same sample showed this constriction to different extents, from totally absent to clearly marked, suggesting a transition towards the morphology of *Alocodera*. Some specimens are presented in figure 5.5 to illustrate the range of variability observed in the shape of the neck in this species.

5.3.2 Characterization of the molecular diversity from genera *Apodera, Alocodera* and *Padaungiella*

We obtained partial COI DNA sequences from 38 specimens, including 15 from genus *Apodera* (barcoded here for the first time with this marker), but also 7 and 16 cells respectively for *Alocodera* and *Padaungiella*. All of these sequences form a robust monophyletic clade (79% bs) with respect to other Hyalospheniformes (Fig 5.6).

All *Apodera* are monophyletic (79% bs) and form two robust clades that grouped, respectively all specimens with a keel (99% bs) and those without a keel (95% bs). Both clades presented important internal genetical variation: The first clade is composed of two keeled species, referred to as *Apodera* sp 3 and 4, while the second clade includes specimens with reg-

Figure 5.5 – Pictures of several specimens of *Padaungiella lageniformis* 2 BR presenting variation in the shape of the neck. These specimens were all extracted from the same sample from Brazil (Table 5.1).

ular shape similar to the type (referred to as *Apodera* sp 1) and others which are much more slender (i.e. L/B ratio of 1.6 and 2.2 respectively, referred to as *Apodera* sp 2; Fig 5.6). Genetic dissimilarity between the two clades was measured around 18%.

The *Alocodera* sequences are mainly split into two clades; the first (100% bs) grouped the South American isolates *A. cockayni* and *Alocodera* sp 1, while the second (83% bs) is composed of New Zealand isolates *Alocodera* sp 2 and 3. The dissimilarities values among this second clade reach 4% between the most divergent sequences. Finally, the specimen *Alocodera* sp 4 is more divergent and groups with *Padaungiella wailesi* 2 NZ in a separate clade as a sister clade to genus *Apodera* (85% bs), thus making both genera *Padaungiella* and *Alocodera* paraphyletic (Fig 5.6). Genetic dissimilarity between the South American and New Zealand clades ranged between 8 and 9%.

Finally, all *Padaungiella* sequences from specimens collected from Europe branched robustly together (94% bs) in a homogenous clade, whereas *Padaungiella lageniformis* 2 BR (from Brazil), *P. lageniformis* 3 NZ and *P. wailesi* 2 NZ (from New Zealand) formed three independent groups (Fig 5.6). The first clade includes two new molecular species from Bulgaria and Spain, i.e. *P. wailesi* 1 BG and *P. lageniformis* 1 ES. The first species is basal and closely related to *P. wailesi* and *P. nebeloides* with a dissimilarity value of 2%, while the second is much divergent and forms a single clade near the base of the tree. Both *P. lageniformis* 2 BR and *P. lageniformis* 3 NZ form single clades at the base of genus *Alocodera*, while *P. wailesi* 2 NZ groups with *Alocodera* sp 4. The dissimilarity value between the former two species is 13%, while their respective genetic distance with the European clade is 12%.

Although most clades display robust phylogenetic support, the basal organization of the tree is weakly supported (Fig 5.6). Indeed, only genus *Apodera* formed a monophyletic clade while genera *Padaungiella* and *Alocodera* are currently paraphyletic.

5.4 Discussion

5.4.1 *Apodera* and *Alocodera* are not monotypic genera

Apodera and *Alocodera* are both traditionally considered as monospecific genera and the morphological variations observed were considered as within the normal variability of these species (Yeates and Foissner, 1995; Zapata and Fernandez, 2008). However, some of this morphological variability had been previously noted (Deflandre, 1936; Penard, 1911; Wailes, 1913; Van Oye, 1956), leading to the description of *Apodera crenata* (Jung, 1942) and *A.* (*Nebela*) *wellingtonia* (Decloitre, 1964), and two varieties of *Apodera vas*, i.e. *A. vas* var. *longicollis* and *A. vas* var. *obliqua* (Grospietsch, 1971). However, *A. crenata* and *A. wellingtonia* have not been reported since their description, and because varieties are not accepted by the ICZN, *A. vas* is now considered as monophyletic.

Our new data and synthesis of existing data clearly demonstrates the existence of several species within genera *Apodera* and *Alocodera* (Fig 5.1, 5.3). The most prominent discriminating trait is probably the keel observed in some New Zealand isolates (Fig 5.2). Hollow keels are a distinctive, taxonomically relevant trait that constitutes the synapomorphy of genus *Gibbocarina* and which can been observed in several species of Hyalospheniformes, e.g. *Nebela aliciae* and *Quadrulella alata* (Kosakyan et al., 2016b). Accordingly, the presence of a robust monophyletic

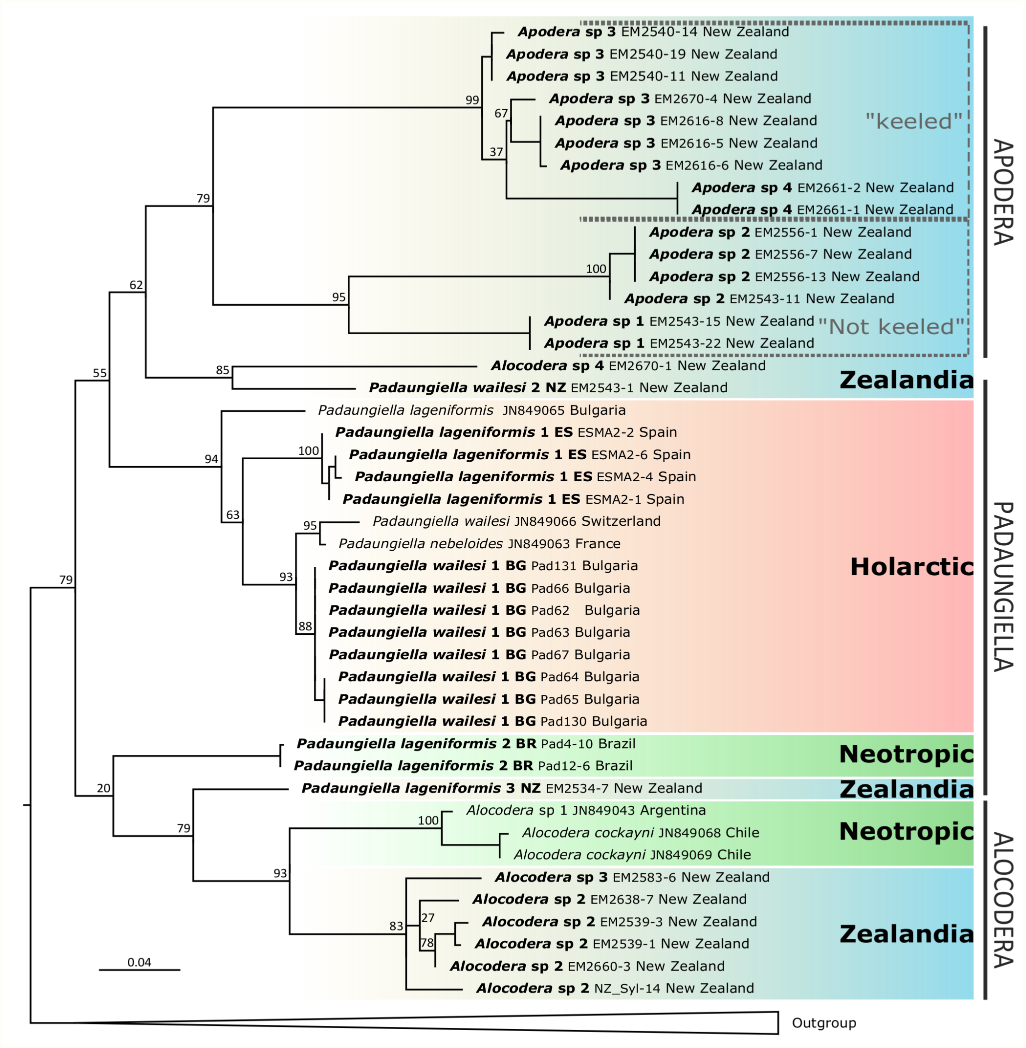

Figure 5.6 – Maximum likelihood phylogenetic tree of all COI genes sequences obtained during this study (bold font) as well as the six sequences retrieved from Genbank from Kosakyan et al. (2012). The tree was rooted with several hyalospheniid genera retrieved from Genbank (i.e. : *Longinebela speciosa* JN849044, *Certesella certesi* JN849064, *Alabasta militaris* MH616621, *Nebela flabellulum* JN849026, *Nebela gimlii* KP083297, *Nebela rotunda* JX682591, *Quadrulella symmetrica* JN849047, *Planocarina carinata* JN849038, *Cornutheca ansata* JN849054 and *Hyalosphenia papilio* JN849016). Numbers at the node correspond to bootstrap values (bs).

clade possessing a keel sister to the genus *Apodera* is a strong evidence of important overlooked diversity. Interestingly, the keeled specimens were all found on *Sphagnum* mosses (*Apodera* sp 3) or brown mosses collected in fens (*Apodera* sp 4), whereas non-keeled *Apodera* sp 1 and 2 were retrieved from forest litter (Table 5.1). Peatlands are well-known as a habitat for a high diversity of hyalospheniformes. *Sphagnum* mosses modify profoundly the abiotic conditions through physical (water holding capacity) and chemical mechanisms (acidification, high nutrient uptake capacity leading to nutrient depletion) (Clymo, 1973) with a clear impact on the associated microbial communities (Jassey et al., 2011), including testate amoebae. Such ecological filtering was suggested to be a driving force for the evolution of specific diversity (Fiz-Palacios et al., 2014). Overall, the clear genetic, morphological and ecological differences observed between *Apodera* sp 1-2 and *Apodera* sp 3-4 may even suggest the erection of a new genus. Further work including the characterization of more similar isolates will be necessary to justify such possible taxonomic action.

The genus *Apodera* shows a considerable range of test size. Zapata and Fernandez (2008) already demonstrated that *Apodera vas* was a polymorphic species composed of several discrete morphotypes in South America, however as the coefficients of variation were low (i.e. 1.4- 3.8%), these different morphotypes were considered as belonging to a single species. We discovered fortuitously that the reason of this erroneous interpretation is in fact due to a calculation mistake. Indeed, the authors calculated the ratio of the standard error (SE) to the mean ($\bar{x}$) instead of the standard deviation (SD) to the mean ($\bar{x}$). As a result, the true coefficient of variation (CV) value should be 10 X higher (i.e. 14-38%), which would strongly suggest the existence of more than one species. The morphological variability we observed in *Apodera* s.l. in New Zealand confirms the strong polymorphism observed by Zapata and Fernandez (2008) in Chile. Our data also suggests that morphology is genetically determined. Indeed, *Apodera* sp 1 and 2, which differ in size and shape (i.e. L/B ratio of 1.6 and 2.2 respectively) (Fig 5.1) are genetically distinct (i.e. 13%) (Fig 5.6), above 4% barcoding gap generally admitted for Hyalospheniformes (Singer et al., 2019).

The situation is even more complex (and rather confusing) in genus *Alocodera*. Our new results show that this genus is not monotypic and may be polyphyletic. *Alocodera* isolates from South America and New Zealand (*Alocodera* species 2 and 3) are clearly genetically distinct and branched robustly together. But *Alocodera* sp 4 branched totally outside the group together with a specimen of *Padaungiella* without constriction from New Zealand that was classified as *Padaungiella wailesi* 2 NZ. This node received a strong support (85% bs), showing that *Alocodera* is actually currently paraphyletic. However, as this result is very surprising and is based on a single sequence, we will not discuss it further until it is validated by additional data. The presence of a constriction in the neck, defined by Penard (1911) as the synapomorphy of the genus varies quite consistently between species. Indeed, *Alocodera* sp 3 (Fig 5.3 R) has almost no constriction and branched with New Zealand species (Fig 5.6). Interestingly, *P. lageniformis* 2 BR from Brazil showed an important variation with respect to the shape of its neck, some individuals showing a slight constriction, but others none at all (Fig 5.5). This variation did not correspond to any phylogenetic pattern, suggesting that it represents individual variation, or phenotypic plasticity (Mulot et al., 2017). A wider sampling including more specimens and other geographical regions would be needed to evaluate the validity of the existence and depth of the constriction as a valid taxonomical criterion for species identification.

Altogether, our results show the existence of a wide diversity of taxa in genera *Apodera, Alocodera* and also *Padaungiella. Padaungiella* and possibly *Alocodera* being currently paraphyletic, validity of any morphological trait to separate genera and species will need to be reassessed. Moreover, even though considerable genetic distances separate European and Southern Hemisphere *Padaungiella* species, we did not succeed in finding morphological traits that could represent synapomorphies for each clade (Fig 5.4). Here too, a wider sampling of individuals per species is needed to hopefully find discriminating traits that can distinguish species.

5.4.2 Implications for the distribution of microbial diversity, with particular emphasis on biogeography

Evidences provided for non-cosmopolitanism of protists used *Apodera vas* and *Alocodera cockayni* as examples of geographically-limited distributions, the argument being that these biogeographical flagships were commonly reported from the Southern Hemisphere and intertropical zone but were never found in the Northern Hemisphere, despite over 200 years of much more intensive investigation (Smith and Wilkinson, 2007). Here, we show that both Gondwanian flagships *Apodera vas* and *Alocodera cockayni* are composed of several, perhaps many species. Within *Alocodera*, isolates from South America and New Zealand were genetically dissimilar, which may suggest vicariant distributions. If this were the case, species distribution would be more limited than previously thought, adding to the evidence for narrow distribution in protists as shown for the *Hyalosphenia papilio* species complex (Singer et al., 2019) and *Cornutheca ansata* (Heger et al., 2011b). A systematic exploration of other regions of the Southern Hemisphere (e.g. Africa, Australia) and Northern Hemisphere (especially North America and Asia) would be needed to delineate the distribution of species and identify the processes that drove species present-day distributions.

According to the consensual “moderate endemicity” model, based mostly on ciliate morphological data, most protists are cosmopolitan and only 30% are endemic (Foissner, 2007). Under these assumptions, *Padaungiella wailesi* and *P. lageniformis* should be both cosmopolitan species, as no clear morphological difference separates European and Southern Hemisphere isolates (Fig 5.4). However, these two groups are genetically different, and while European isolates form a unique, robustly supported clade (94% bs), Southern Hemisphere isolates are scattered across the tree. A probable scenario would thus favour a Southern Hemisphere origin of the whole *Apodera-Alocodera-Padaungiella* clade, and a limited number of colonization events of the Northern Hemisphere. More data are necessary to confirm this hypothesis, and especially molecular data to infer biogeographical and phylogeographical patterns in Hyalospheniformes. When applying systematically barcoding, the distribution of species appears fragmented as compared with morphology approaches only (Singer et al., 2019). Although the consensual “moderate endemicity” model seemed appealing at a time of heated debate, the most recent findings in testate amoeba biogeography, and especially the results from DNA barcoding of variable markers providing species-level taxonomic resolution, rather point to the existence of a gigantic diversity of locally distributed species.

5.5 Acknowledgement

We thank the Swiss NSF for funding (project no. 31003A_163254). EL was also financed by the intramural project 201730E063 (CSIC) and the project PGC2018-094660-BI00 from the Spanish Government. EM’s collecting work was funded by a scientific leave to New Zealand as well as internal funding from Manaaki Whenua / Landcare Research and other sources funding the travel of colleagues who provided samples. We are also grateful to all the colleagues that sent us material from different location or that participate to sampling session.

Table 5.1 – List of sequenced species, sampling localizations and sequences details.

Species/Taxa	**Cells codes**	**Environments**	**Substrates**	**sampling location and countries**	**Co-ordinates**		**Genbank number**	**Published**
Alocodera sp 1	*A. cockayni AR*	Peatlands	*Sphagnum mosses*	Lapataya National Park, Tierra del Fuego (AR)	-54.510000	-68.340000	JN849043	Kosakyan et al. 2012
Alocodera cockayni	*A. cockayni CL-1*	NA	NA	Alerce Costero, Valdivia (CL)	-40.110000	-73.280000	JN849068	Kosakyan et al. 2012
Alocodera cockayni	*A. cockayni CL-2*	NA	NA	Alerce Costero, Valdivia (CL)	-40.110000	-73.280000	JN849069	Kosakyan et al. 2012
Alocodera sp 2	NZ_Syl-14	NA	NA	Lake Sylvan (NZ)	-44.708225	168.322814	-	-
Alocodera sp 2	EM2539-1	Peatlands	*Sphagnum hummock*	Edgemont National Park (NZ)	-39.255058	174.043106	-	-
Alocodera sp 2	EM2539-3	Peatlands	*Sphagnum hummock*	Edgemont National Park (NZ)	-39.255058	174.043106	-	-
Alocodera sp 2	EM2660-3	Lowland beech forest	*Sphagnum hummock*	Kepler Track (NZ)	-45.389282	167.576330	-	-
Alocodera sp 2	EM2638-7	Low Metrosideros forest	Moss tuft	Auckland Island (NZ)	-50.812636	166.065388	-	-
Alocodera sp 3	EM2583-6	Under Silver beech	*Sphagnum hummock*	Siberia (NZ)	-44.224333	169.033100	-	-
Alocodera sp 4	EM2670-1	Side of lake outlet	*Sphagnum hummock*	Roureburn great walk trail (NZ)	-44.727105	168.172889	-	-
Apodera sp1	EM2543-15	Forest	Soil litter and humus	Edgemont National Park (NZ)	-39.225050	174.122842	-	-
Apodera sp1	EM2543-22	Forest	Soil litter and humus	Edgemont National Park (NZ)	-39.225050	174.122842	-	-
Apodera sp2	EM2556-1	Beech forest	Soil litter and humus	Road to Lake Wakaremoana (NZ)	-38.629833	177.030778	-	-
Apodera sp2	EM2556-7	Beech forest	Soil litter and humus	Road to Lake Wakaremoana (NZ)	-38.629833	177.030778	-	-
Apodera sp2	EM2556-13	Beech forest	Soil litter and humus	Road to Lake Wakaremoana (NZ)	-38.629833	177.030778	-	-
Apodera sp2	EM2543-11	Forest	Soil litter and humus	Edgemont National Park (NZ)	-39.225050	174.122842	-	-
Apodera sp3	EM2540-11	Peatlands	*Sphagnum mosses*	Edgemont National Park (NZ)	-39.255058	174.043106	-	-
Apodera sp3	EM2540-14	Peatlands	*Sphagnum mosses*	Edgemont National Park (NZ)	-39.255058	174.043106	-	-
Apodera sp3	EM2540-19	Peatlands	*Sphagnum mosses*	Edgemont National Park (NZ)	-39.255058	174.043106	-	-
Apodera sp3	EM2670-4	Side of lake outlet	*Sphagnum mosses*	Roureburn great walk trail (NZ)	-44.727105	168.172889	-	-
Apodera sp3	EM2616-5	Peatland on lake shore	*Sphagnum mosses*	Old Ghost Road trail (NZ)	-41.698631	172.159492	-	-
Apodera sp3	EM2616-6	Peatland on lake shore	*Sphagnum mosses*	Old Ghost Road trail (NZ)	-41.698631	172.159492	-	-
Apodera sp3	EM2616-8	Peatland on lake shore	*Sphagnum mosses*	Old Ghost Road trail (NZ)	-41.698631	172.159492	-	-
Apodera sp4	EM2661-1	Alpine terrace with stream, ponds and fens	Brown mosses	Kepler Track (NZ)	-45.389282	167.576330	-	-
Apodera sp4	EM2661-2	Alpine terrace with stream, ponds and fens	Brown mosses	Kepler Track (NZ)	-45.389282	167.576330	-	-
Padaungiella lageniformis	*P. lageniformis BG*	NA	*Sphagnum mosses*	Vitosha Mountain (BG)	42.360000	23.170000	JN849065	Kosakyan et al. 2012
Padaungiella lageniformis 1 ES	ESMA2-1	Forested meadow	Forest litter	Canencia, Madrid (SP)	40.874406	-3.771439	-	-
Padaungiella lageniformis 1 ES	ESMA2-2	Forested meadow	Forest litter	Canencia, Madrid (SP)	40.874406	-3.771439	-	-
Padaungiella lageniformis 1 ES	ESMA2-4	Forested meadow	Forest litter	Canencia, Madrid (SP)	40.874406	-3.771439	-	-
Padaungiella lageniformis 1 ES	ESMA2-6	Forested meadow	Forest litter	Canencia, Madrid (SP)	40.874406	-3.771439	-	-
Padaungiella lageniformis 2 BR	Pad4-10	Rock on forest	*Sphagnum mosses*	Pico do Baepi, Ilhabella, Sao Paulo (BR)	-23.791842	-45.332375	-	-
Padaungiella lageniformis 2 BR	Pad12-6	Rock on forest	*Sphagnum mosses*	Pico do Baepi, Ilhabella, Sao Paulo (BR)	-23.791842	-45.332375	-	-
Padaungiella lageniformis 3 NZ	EM2534-7	Forest	Soil litter and humus	Edgemont National Park, Mt. Taranaki (NZ)	-39.408561	174.154592	-	-
Padaungiella nebeloides	*P. nebeloides FR*	Lake	Floating Sphagnum mire	Lac de Bellefontaine, Jura (FR)	46.572089	6.087791	JN849063	Kosakyan et al. 2012
Padaungiella wailesi	*P. wailesi CH*	Forest	Forest litter	Bois du Jorat, Vaud (CH)	46.582306	6.698465	JN849066	Kosakyan et al. 2012
Padaungiella wailesi 1 BG	Pad62	Beech forest	Forest litter	Natural Park Vitosha (BG)	42.627361	23.266817	-	-
Padaungiella wailesi 1 BG	Pad63	Beech forest	Forest litter	Natural Park Vitosha (BG)	42.627361	23.266817	-	-
Padaungiella wailesi 1 BG	Pad64	Beech forest	Forest litter	Natural Park Vitosha (BG)	42.627361	23.266817	-	-
Padaungiella wailesi 1 BG	Pad65	Beech forest	Forest litter	Natural Park Vitosha (BG)	42.627361	23.266817	-	-
Padaungiella wailesi 1 BG	Pad66	Beech forest	Forest litter	Natural Park Vitosha (BG)	42.627361	23.266817	-	-
Padaungiella wailesi 1 BG	Pad67	Beech forest	Forest litter	Natural Park Vitosha (BG)	42.627361	23.266817	-	-
Padaungiella wailesi 1 BG	Pad130	Beech forest	Forest litter	Natural Park Vitosha (BG)	42.627361	23.266817	-	-
Padaungiella wailesi 1 BG	Pad131	Beech forest	Forest litter	Natural Park Vitosha (BG)	42.627361	23.266817	-	-
Padaungiella wailesi 2 NZ	EM2543-1	Forest	Soil litter and humus	Edgemont National Park (NZ)	-39.225050	174.122842	-	-

Chapter 6

Diversity patterns of genus *Nebela* (Arcellinida, Amoebozoa) in *Sphagnum*-dominated peatlands along an elevation gradient assessed by high throughput sequencing

Diversity patterns of genus *Nebela* (Arcellinida, Amoebozoa) in *Sphagnum*-dominated peatlands along an elevation gradient assessed by high throughput sequencing

---

QUENTIN BLANDENIER[a,b], MATTHIEU MULOT[a], CLÉMENT DUCKERT[a], DAVID SINGER[a,c], RAPHAEL JAUSLIN[d], AMANDINE PILLONEL[a], MOLLY BAUR[a], EDWARD A. D. MITCHELL[a,e], ENRIQUE LARA[b]

Abstract

The study of diversity patterns along elevation gradients is a classical question in ecology but data for soil protists are limited. High Throughput Sequencing of environmental DNA now allows addressing these questions with the appropriate power and precision. Many characteristics of ecosystems, such as soil and vegetation type vary along elevation gradients, potentially representing confounding factors. One way to circumvent this limitation is to focus on well-defined and highly specific ecosystems. Our choice is *Sphagnum* peatlands and focusing on genus *Nebela* (Amoebozoa; Arcellinida; Hyalospheniformes) comprising at least eight species with contrasted ecological preferences. Our aim was to assess to what extent climatic factors or basic soil characteristics (pH and nitrogen content) explained the species distribution within this genus. We assessed the diversity of genus *Nebela* in *Sphagnum*-dominated peatlands along a 1200m elevation gradient in the Swiss Pre-Alps. We collected eight samples in each of 12 peatlands ranging from 620 to 1820 m asl. We amplified the mitochondrial Cytochrome Oxidase subunit I (COI) gene of genus *Nebela* from environmental DNA samples using specific primers. The analysis of ca. 26 million sequence reads revealed 14 operational taxonomic units (OTUs) eight of which corresponded to known species and six representing unknown phylotypes. Three OTUs largely dominated the communities accounting for 94% of all sequences (i.e. *N. rotunda* (45%), *N. gimlii* (30%) and *N. collaris* (19%)). All known species are ubiquitous along the gradient, while four of the six unknown phylotypes occur preferentially at high elevation. Finally, some OTUs from high elevation are mutually exclusive with generalist OTUs. These results show that novel diversity remains to be described even in this well-studied group and suggest that Hyalospheniformes respond to climatic gradients but the precise mechanisms of this response remain to be determined.

[a] Laboratory of Soil Biodiversity, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland
[b] Real Jardín Botánico, CSIC, Plaza de Murillo 2, Madrid 28014, Spain
[c] Department of Zoology, University of Sao Paulo, 05508-090 Sao Paulo, Brazil
[d] Institute of Statistics, University of Neuchâtel, Rue Emile-Argand 11, CH-2000 Neuchâtel, Switzerland
[e] Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland

6.1 Introduction

Biodiversity has long been shown to be distributed following simple patterns across larger scales. For instance, species richness generally decreases along elevation or latitudinal gradients (Gaston, 2000) and is positively correlated with high summer temperature or habitat size. These global patterns have long been described in macro-organisms such as plants and animals, but only recently for protists (Soininen, 2012). Protist diversity was shown to vary in relation to latitude (Schiaffino et al., 2016), nutrient (Grossmann et al., 2016), water availability (Harder et al., 2016) and pH (Shen et al., 2014) but patterns along elevation gradients are less clear (Geisen et al., 2017). Indeed, while the specific richness of animals and plants generally decreases along an elevation gradient, sometimes peaking at mid elevation, studies focused on protists have often yielded contradictory patterns. Using DNA metabarcoding, Grossman and colleagues did not find any correlation between elevation and protist richness in alpine freshwater lakes but attributed changes in community structures to phosphorus concentration and lake size (Grossmann et al., 2016). Similarly, Shen and co-authors reported that changes in the community structure of soil protists were explained more by variation of pH than by elevation (Shen et al., 2014). A study focusing on Myxomycetes combining metabarcoding and morphological approaches demonstrated that the main genera were present across the full length of the gradient, with some exception that were mainly found in the upper half of the transect (Dahl et al., 2018).

Contrary to these studies, morphological observation surveys on testate amoebae have demonstrated similar distribution patterns to macro-organisms. These studies report either a simple linear decrease of richness (Heger et al., 2016), a maximal diversity at mid-elevation (Krashevska et al., 2007; Lamentowicz et al., 2013; Todorov, 1998) or on the contrary a “U-shaped” distribution with the lowest specific richness at mid elevation (Tsyganov et al., 2013).

Two major sources of discrepancies among these studies might be the heterogeneity of the sampled habitats across the elevation gradient and the taxonomic resolution. On the one hand, vegetation cover, litter type and soil physico-chemical characteristics generally vary with elevation (Odum and Barrett, 1971) such that the observed changes in community might reflect habitat differences instead of an impact of elevation-related (climatic) factors (Mitchell et al., 2004). While on the other hand, a coarse taxonomic resolution might mask patterns in species richness and community structure (Hanson et al., 2012; Heger et al., 2009).

Here, using DNA metabarcoding coupled with genus-specific primers targeting the Cytochrome Oxidase subunit I (COI) gene, we assessed the diversity patterns and community structures of genus *Nebela* in wooded *Sphagnum*-dominated peatlands in Switzerland. This particular environment is characterized by highly specialized vegetation (dominated by *Sphagnum* mosses, sedges and ericaceous shrubs), fauna (Borcard, 1991; Borcard and Matthey, 1996) and microbial assemblages (Dedysh et al., 2006) and is thus very homogeneous independently of elevation within a region such as the Swiss Alps. Species of genus *Nebela* are very common in peatlands and the distribution patterns of closely-related species have been shown to be strongly linked to environmental factors (Singer et al., 2018). Moreover, this genus is probably the most studied Hyalospheniformes taxon with molecular methods, and as a results, DNA-barcoding protocols are available as well as a curated COI database (Kosakyan et al., 2013, 2016a; Singer et al., 2015).

6.2 Material and Methods

6.2.1 Sampling and physico-chemical analysis

We selected 12 sites in homogenous well-preserved peatlands ranging from 620 to 1820 m asl. in Swiss Prealps, and performed all sampling during the same week of June 2017 (Table 6.1). At each site we selected eight sampling spots within homogeneous tall pine forest with dense *Sphagnum* cover and a water table depth below 25 cm. At each spot we collected one sample of ca. 200 g of fresh *Sphagnum* mosses. For the molecular analysis 50 g of fresh *Sphagnum* mosses was shaken in sterile distilled waters and filtered with a 200 µm mesh. The filtrate was left to sediment for 30 minutes and the sedimented particles, including the living testate amoebae, were collected and transferred to a vial for DNA extraction (Singer et al., 2018). Physico-chemical analyses were performed on ca. 100 g *Sphagnum mosses* oven-dried during two weeks at 40°C. Percentage C, H, and N and pH was measured following standard protocols (Gregorich and Carter, 2007).

Table 6.1 – List of sampling information.

Elevation [m]	Site number	National site identification*	Site	Locality	Canton	Sample labels	Sampling date	Coordinate WGS84	
620	12	N°75	Sewelimoos	Reutigen	Bern	89-96	28.06.2017	46.7059	7.59834
840	3	N°59	Les Mosses– Rosez	Sâles, Vaulruz	Fribourg	17-24	20.06.2017	46.6259	6.97324
1100	1	N°130	Dévin des Dailles	ChâtelSaintDenis	Fribourg	1-8	20.06.2017	46.5213	6.96073
1180	4	N°555	Muschenegg	Plasselb	Fribourg	25-32	20.06.2017	46.7158	7.22554
1270	5	N°312	Stächelegg / Ghack	Flühli	Lucern	33-40	26.06.2017	46.8358	7.99744
1330	2	N°126	Tourbière des Alpettes	Semsales	Fribourg	9-16	20.06.2017	46.5699	6.97041
1400	6	N°319	Husegg-Ochsenweid	Flühli	Lucern	41-48	26.06.2017	46.8178	8.01058
1450	9	N°263	Seeliwald	Sarnen	Obwald	65-72	28.06.2017	46.9066	8.14927
1460	7	N°317	Gross Gfäl	Flühli	Lucern	49-56	26.06.2017	46.8112	8.01242
1500	8	N°260	Trogenwald	Sarnen	Obwald	57-64	28.06.2017	46.8889	8.10803
1670	11	N°256	Talhubel/Siterenmoos	Sarnen	Obwald	81-88	28.06.2017	46.9004	8.09848
1820	10	N°265	Ober Sewen	Sarnen	Obwald	73-80	28.06.2017	46.8902	8.07577

*https://www.admin.ch/opc/fr/classified-compilation/19910009/index.html

6.2.2 Molecular analysis

Environmental DNA was extracted using a Mo-Bio PowerSoil extraction kit according to the manufacturer's instructions. We performed a first PCR with universal COI primers (Folmer et al., 1994) followed by a nested PCR with *Nebela*-specific primers as previously described (Kosakyan et al., 2013; Singer et al., 2018). Positive PCR products were then quantified with a Qubit® 2.0 Fluorometer (Invitrogen) and 20 ng of each DNA samples was pooled. A DNA library was prepared from the pools using the TruSeq Nano PCR-free Library Preparation kit and the paired-end 2 × 300 bp sequencing was done on an Illumina® MiSeq at the University of Geneva (Molecular Systematics and Environmental Genomics Laboratory) (Seppey et al., 2019).

6.2.3 Bioinformatic pipeline

We adapted the bioinformatical pipeline of Seppey et al. (2019) to remove chimeras and low-quality sequences. We used the Swarm 2.1.8 algorithm (Mahé et al., 2015) with the default options (d = 1) to cluster the sequence into operational taxonomic unit (OTU). The OTUs were assigned manually based on barcoded sequences retrieved from Genbank from previous studies (Kosakyan et al., 2013; Singer et al., 2015, 2018).

6.2.4 Phylogenetic reconstruction

We added all resulting OTUs to reference database with 20 COI sequences representing the main diversity within genus *Nebela* retrieved from Genbank together with sequences from *Padaungiella, Alocodera* and *Certesella* as outgroup. We aligned the sequence manually using BioEdit (Hall, 1999). Phylogenetic reconstruction was conducted using the CIPRES Portal (Miller et al., 2010). A maximum likelihood phylogenetic tree was built using the RAxML v.8.2.10 algorithm (Stamatakis, 2014) with the GTR + GAMMA model and 1000 bootstraps.

6.2.5 Statistical analysis

The original dataset was composed of 96 samples and associated environmental geographical data (pH, nitrogen, carbon, hydrogen, elevation, latitude and longitude). We first discarded nine samples for which data were missing, i.e. seven with unsuccessful PCR and two with missing pH and CHN data. We checked for extreme values for one of those variables, which could yield erroneous interpretation in the downstream analysis. We then removed 11 outliers by computing the Minimum Covariance Determinant (MCD), according to Leys et al. (2018) which resulted in a final dataset composed of 76 samples (Table 6.2). Species read abundances were standardized using log(x+1) to compensate the PCR read amplification bias. Sequence numbers are correlated to biovolume, and in order to obtain relat-

Table 6.2 – Summary of environmental and geographical variables.

	N	Mean	SD	Min	Q1	Median	Q3	Max
pH	76	4.56	0.22	4.08	4.4	4.54	4.71	5.05
nitrogen	76	0.96	0.17	0.49	0.84	0.94	1.06	1.36
carbon	76	43.44	0.60	42.36	43.02	43.45	43.78	44.93
hydrogen	76	6.08	0.13	5.79	5.97	6.09	6.16	6.33
elevation	76	1318	316	620	1180	1400	1480	1820
latitude N°	76	46.77	0.13	46.52	46.71	46.82	46.89	46.91
longitude E°	76	7.70	0.48	6.96	7.23	8.01	8.10	8.15

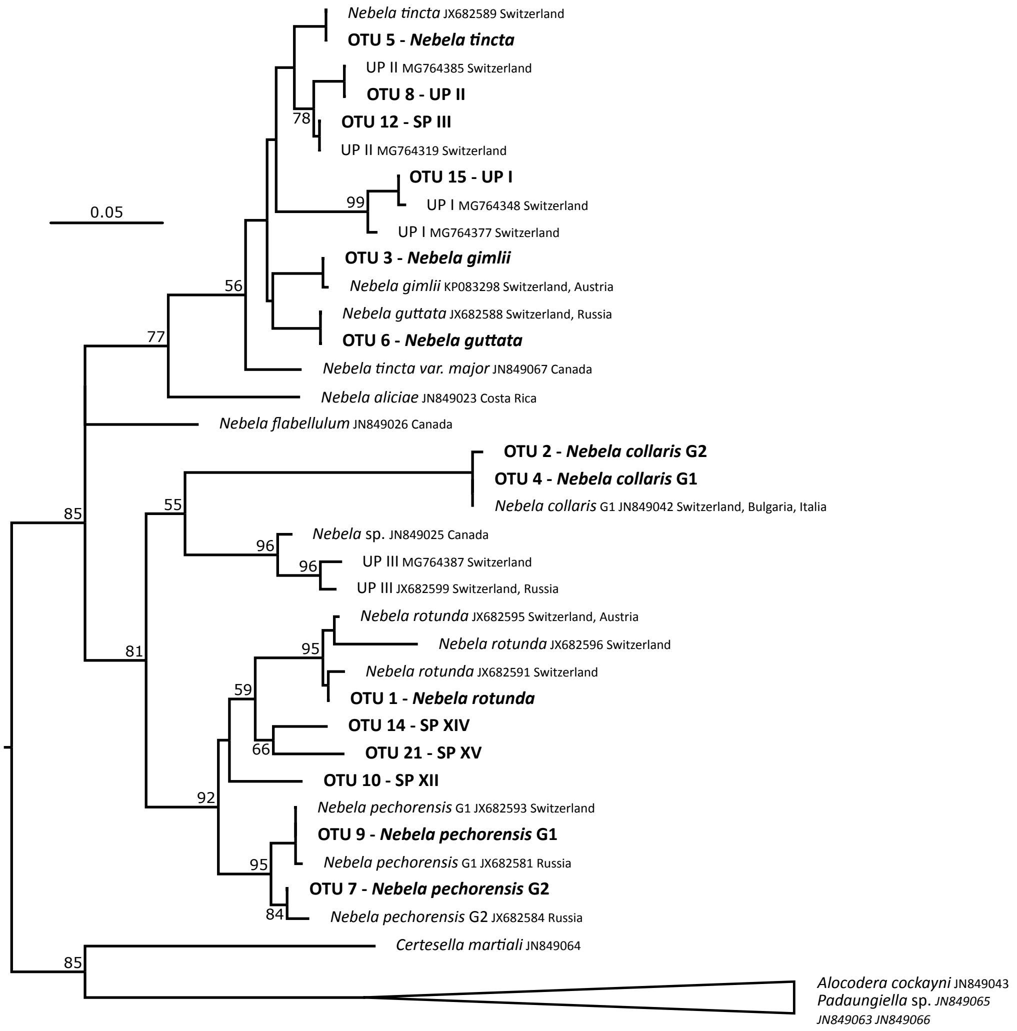

Figure 6.1 – Maximum likelihood phylogenetic tree of the 14 *Nebela* COI OTUs from *Sphagnum* peatlands along an elevation gradient in Switzerland and 20 COI sequences retrieved from Genbank representing the main diversity within genus *Nebela* (Kosakyan et al., 2013; Singer et al., 2015, 2018). The tree is rooted with sequences from *Padaungiella, Alocodera* and *Certesella* as outgroup. Numbers at the node correspond to bootstrap values (bs) and only values above 50% are shown.

ive abundances of individual cells, we corrected sequence numbers according to Kosakyan et al. (2015). The biovolume of unknown taxa was assigned to the largest testate amoeba species, i.e. *Nebela collaris*, to avoid overestimating their relative abundance in the community (Singer et al., 2018). Corrected biovolumes were then standardized using Hellinger transformation prior to ordinations.

6.2.6 Multivariate analysis

To assessed the relationships between the Hellinger-transformed *Nebela* community data and environmental variables, we performed a redundancy analysis (RDA) using the function rda of the vegan package v2.3-5 (Oksanen et al., 2016). We selected the best explaining model using forward selection with the ordistep function of the vegan package. Elevation being the sole variable with a significant correlation, we studied the distribution of each OTU along the elevation gradient. We then performed a principal component analysis (PCA) on the Hellinger-transformed *Nebela* community data to explore the patterns of community composition. As the PCA plot revealed the existence of four distinct clusters, we split the dataset into four groups using a simple kmeans computed on the first two principal components of the PCA. The clustering were computed using 76 random seed to prevent biases induced by odd initial centroid position. We then calculated the ecological optima and tolerance of each OTU for elevation, using the method describes in Cristóbal et al. (2014). Finally, to assess the possible existence of association or co-exclusion between pairs of *Nebela* species we performed a probabilistic co-occurrence analysis (Veech, 2013).

Table 6.3 – Composition, abundance, proportion and assignation of the OTUs.

OTU names	Number of haplotypes	Number of sequences*	Proportion of total sequences	Assignation
1	696	1752894	45.44%	*N. rotunda*
2	602	716380	18.57%	*N. collaris* G2
3	510	1166687	30.24%	*N. gimlii*
4	17	22103	0.57%	*N. collaris* G1
5	15	56888	1.47%	*N. tincta*
6	6	47587	1.23%	*N. guttata*
7	7	44947	1.17%	*N. pechorensis* G2
8	7	17384	0.45%	UPII
9	4	24364	0.63%	*N. pechorensis* G1
10	2	4035	0.10%	sp XII
12	2	2907	0.08%	SP III
14	1	713	0.02%	SP XIV
15	1	595	0.02%	UPI
21	1	128	0.00%	SP XV
Total	1871	3857611	100.00%	

*Corrected with biovolume

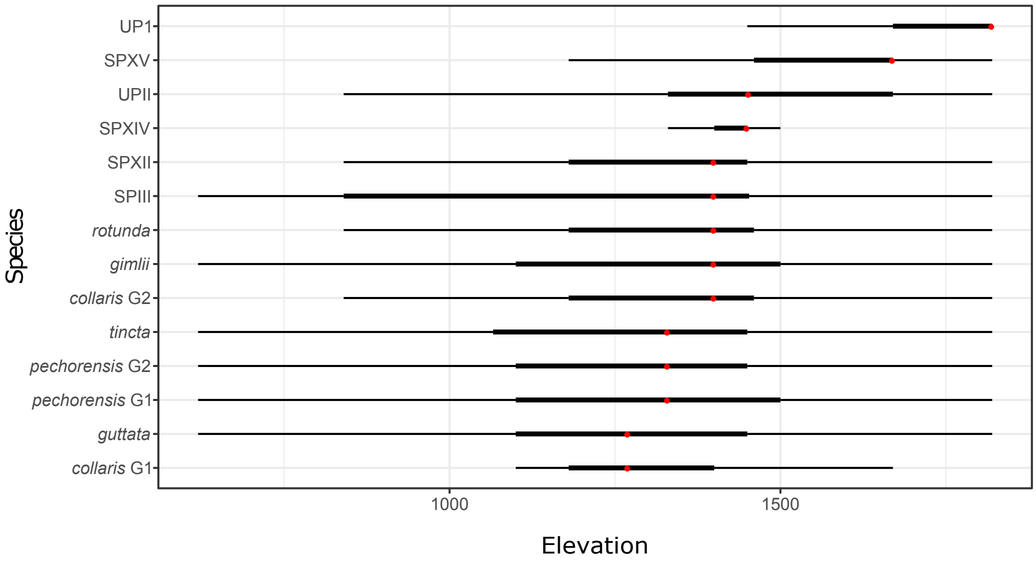

Figure 6.2 – Optima and tolerance of *Nebela* COI OTUs from *Sphagnum* peatlands along an elevation gradient in Switzerland. Red points: species' optima computed by weighted median; bold line: optimal ranges; narrow line: tolerance range.

6.3 Results and Discussion

6.3.1 Molecular diversity within genus *Nebela*

We obtained 1878 unique haplotypes with around 30'000 sequences on average per sample. Our clustering approach revealed 21 OTUs, seven of which were assigned to other genera (e.g. *Arcella*) or unidentified fragments of genomic DNA and were thus discarded. The three dominant OTUs were composed of 696, 602 and 510 haplotypes respectively, whereas all others included fewer than 20 haplotypes (Table 6.3). Similarly, the abundance of sequences among the different OTUs varied considerably. Indeed, the three most abundant OTUs accounted for 94% of the total sequences (i.e. 45%, 30% and 19% respectively), whereas the four rarest each accounted for <0.1%.

The 14 OTUs present along our elevation gradient were related and assigned either to species already described formally such as *N. collaris, N. guttata, N. pechorensis, N. rotunda, N. tincta* and *N. gimlii* (Kosakyan et al., 2013; Singer et al., 2015), to unknown phylotypes already recovered in former studies, i.e. UPI and UPII (Singer et al., 2018), or to unknown phylotypes recovered here for the first time, i.e. SP III, SP XII, SP XIV and SP XV. Four OTUs related to different haplotypes of *N. collaris* and *N. pechorensis* were kept separate and are referred to as group 1 and group 2 (G1, G2). Our OTUs represent the majority of the diversity usually present in peatlands from the Jura Mountains the only ex-

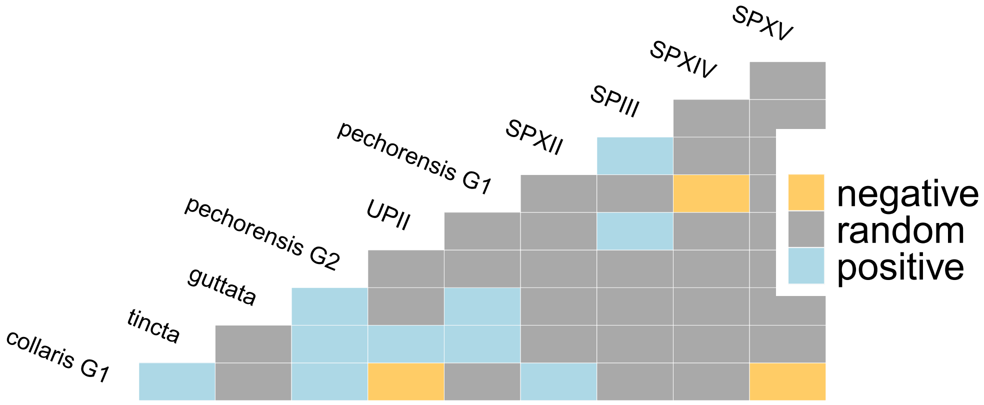

Figure 6.3 – Co-occurrence and co-exclusion of *Nebela* COI OTUs in *Sphagnum* peatlands along an elevation gradient in Switzerland computed according to (Veech, 2013).

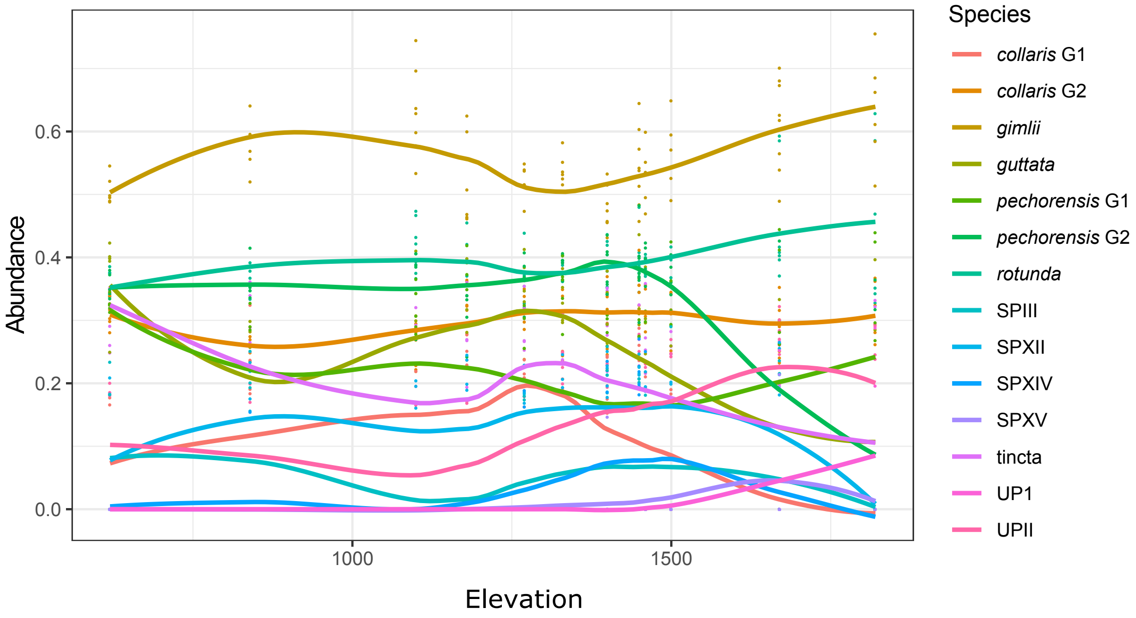

Figure 6.4 – Abundance patterns of *Nebela* OTUs from *Sphagnum* peatlands along an elevation gradient in Switzerland. The abundance values have been standardized using log (x+1) and Hellinger-transformed.

ception being UPIII, absent from our samples (Singer et al., 2018). The four new OTUs could correspond to at least two new species, forming a sister clade to *N. rotunda* (Fig 6.1). Further investigations, including morphological characterization as well as lineage delineation, e.g. using the general mixed Yule-coalescent (GMYC) model (Fontaneto et al., 2007) and the automatic barcode gap discovery (ABGD) method (Puillandre et al., 2012), are required to resolve this issue. The dominant OTUs belonged to formally described species, whereas the uncharacterized OTUs (UP/SP I-XV) belonged to rare species or species not dominant in this environment (Table 6.3).

6.3.2 Diversity patterns of genus *Nebela* along the elevational gradient

None of the variables investigated were correlated to elevation (Appendix 6.1). As elevation was the only significant variable in the RDA (F test, $p < 0.05$), we computed the elevation optimum and tolerance for each OTU. This analysis clearly shows that while “common” OTUs are quite ubiquitous along the elevation gradient, some newly discovered OTUs occur preferentially at high elevation, sometimes with a narrow range (Fig 6.2).

The probabilistic co-occurrence analysis revealed that the OTUs occurring preferably in high elevation range were mutually exclusive with more common OTUs (Fig 6.3). For instance, *Nebela* SPXIV, which occurs at a very narrow elevation range (1270-

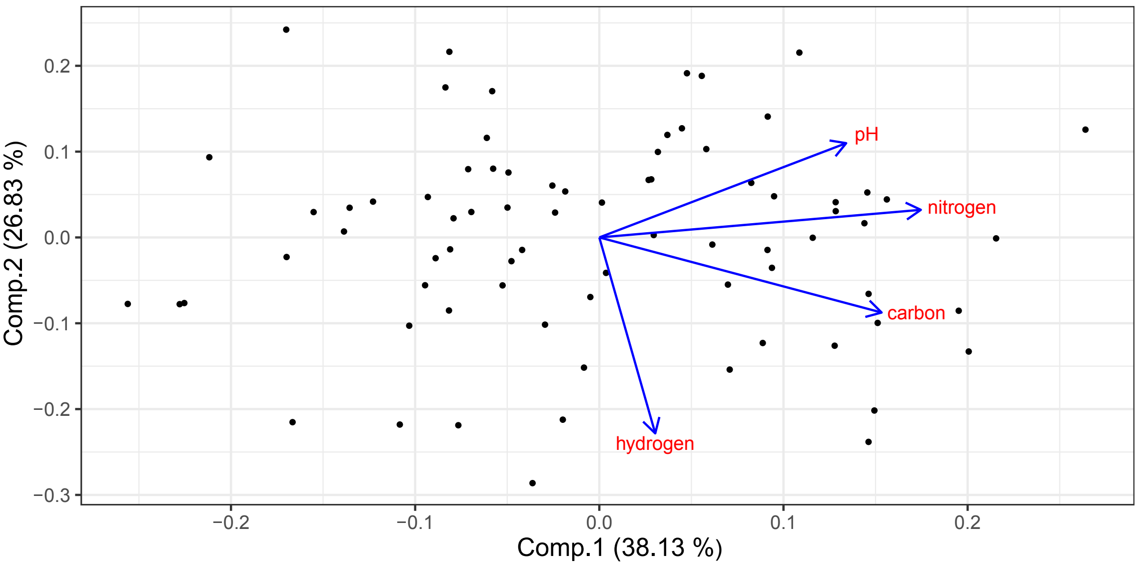

Figure 6.5 – Abundance patterns of *Nebela* OTUs from *Sphagnum* peatlands along an elevation gradient in Switzerland. The abundance values have been standardized using log (x+1) and Hellinger-transformed.

1460 m), showed a mutual exclusion with *N. pechorensis* G1, which it replaces in its optimal range. The same observation can be made for *Nebela* SPXV and the UPII which replace *N. collaris* G1 (Fig 6.4). Such analyses should be performed on more sites to further explore these patterns and their driving causes.

Elevation was the only of the measured variable to which OTUs assemblage were correlated (Fig 6.5). Several of the newly discovered OTUs had an optimum for high elevation, being absent from lower elevation sites. Our data does not allow us to determine the possible causes for such a pattern but it may be that these species have a low tolerance for warmer or drier conditions that are more likely to occur at the lower elevation sites. Longer snow cover at higher elevation results in a shorter vegetation period and a reduced risk of hydrological stress allowing specialized species to thrive, as exemplified with vegetation (Dvorský et al., 2015; Komac et al., 2015). This could be explored using transplantation experiments as done previously in the region (Bragazza et al., 2016). In the PCA, the OTUs assemblages clustered into four groups corresponding to elevation ranges, with little overlap (Fig 6.6). These assemblages are mainly determined by the newly discovered OTUs which are often absent below 1500m.

In a previous study on *Nebela* communities in *Sphagnum*-dominated peatlands of the Jura Mountains at ca. 1000m elevation, no competitive exclusion was observed and the Pianka index was low. These findings suggested that the distribution of these species among microhabitats of *Sphagnum*-dominated peatlands was determined by effective niche partitioning unrelated to competition (Singer et al., 2018). This is in contradiction with our co-occurrence analysis that revealed some cases of mutual exclusion. Such a pattern corresponds to tree succession along elevation gradients where generalists species are outcompeted at high elevation by specialists (Pavlović et al., 2019). However, with only two studies on such patterns it is clearly too early to reach general conclusions. But our study and the one of Singer et al. (2018) show the potential of this approach to address questions including alpha diversity, community ecology and macroecology.

6.4 Acknowledgement

We thank the Swiss NSF for funding (project no. 31003A_163254 to EL). EL was also funded by the intramural project 201730E063 (CSIC) and the project PGC2018-094660-BI00 from the Spanish Government. We are grateful to Fanny Kupferschmid and Pilar Fernandez-Conradi for their help in the sampling sessions and to Emanuela Reo for performing the sequencing.

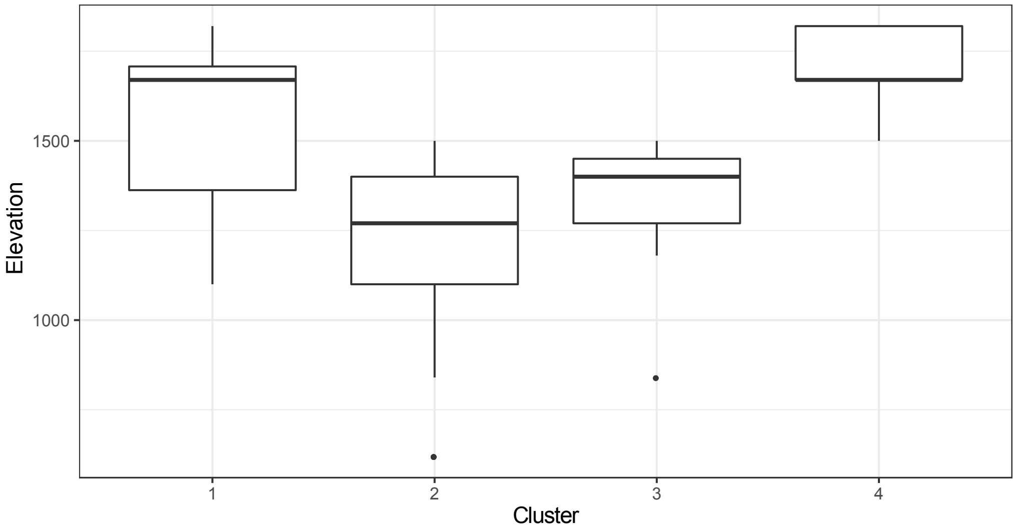

Figure 6.6 – Range of elevation covered by each cluster of *Nebela* COI OTUs from *Sphagnum* peatlands along an elevation gradient in Switzerland. The boxes include 95% of the cluster values, the transversal bars correspond to the medians and the dots are the outliers.

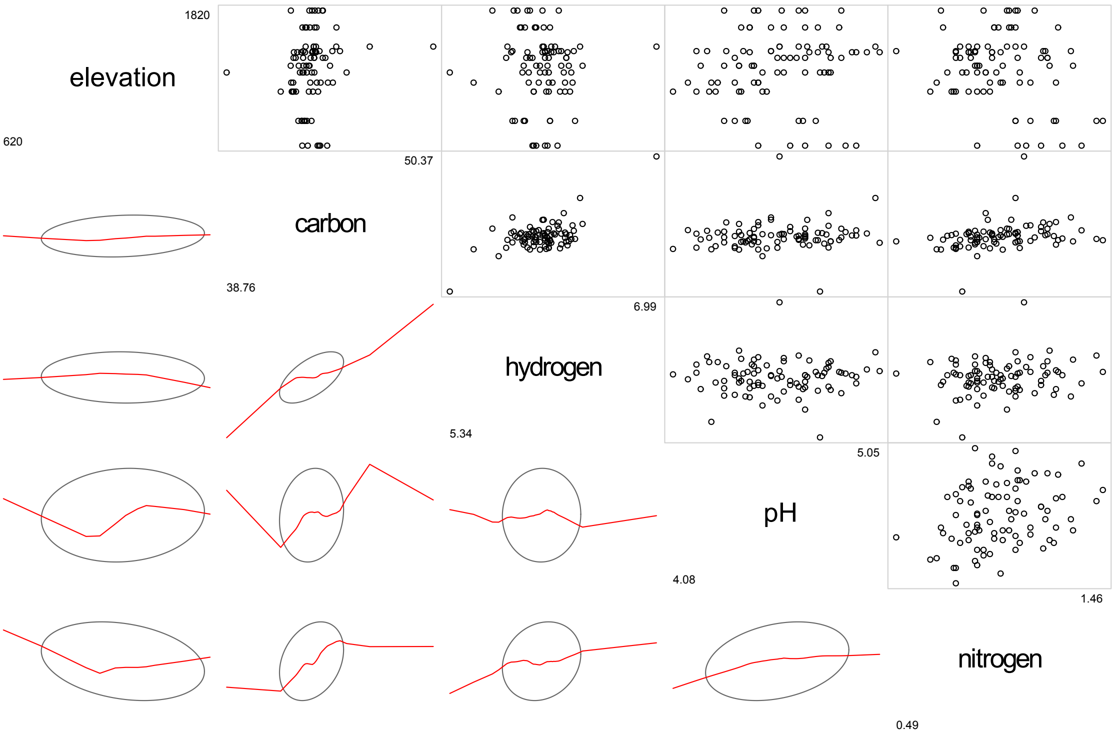

Figure 6.7 – Correlation between environmental variables in *Sphagnum* peatlands along an elevation gradient in Switzerland. Elevation is not significantly correlated with any other variable.

Chapter 7

Discussion and Conclusion

7.1 Context

Improvements in the estimation of microbial diversity have always been directly related to methodological developments. The onset of electron microscopy (both scanning and transmission) brought considerable insights that allowed a more accurate characterization of the organisms, allowing the access to ultrastructural features. In protists, the revolution brought by electron microscopy in the mid-20th century permitted the characterization of the smallest forms that could not be distinguished with light microscopy. The English botanist Irene Manton pioneered the use of electron microscopy in protistology with observations on the ultrastructure of phototrophic (Manton and Parke, 1960; Manton and Leedale, 1969) and heterotrophic (Manton et al., 1976) organisms. Her works on the internal cell structure of many algae permitted documenting the immense structural variety hiding behind apparently identical "little green balls" (Leadbeater, 2004). Electron microscopy contributed to expand knowledge of testate amoebae as well, and contributed in improving systematics, leading notably to the description of new species and setting the standards for the morphological characterization of new species (Coûteaux et al., 1979; Ogden and Meisterfeld, 1989).

The application of molecular techniques to the estimation of protist diversity came afterwards, as the molecular characterization of species became the new standard amongst protistologists (Pawlowski et al., 2012). However, due to the low phylogenetic resolution (Lara et al., 2008; Nassonova et al., 2010) and the possible existence of several diverging copies of the 18S rRNA gene (García-Martín et al., 2019), the search for new molecular markers adapted for Amoeboza has been a necessity. While actin proved equally unreliable because of high paralogy levels (Lahr et al., 2011b), the first subunit of the mitochondrial cytochrome oxidase gave the most satisfactory results (Nassonova et al., 2010) and was later used for Arcellinid testate amoebae. Barcoding, as defined in Hebert et al. (2003), has been applied to testate amoebae starting with the Hyalospheniformes (Kosakyan et al., 2012) and revealing high levels of pseudo-cryptic diversity. This has led to the description of many new species (Singer et al., 2015), and set the methodological ground for the discovery of many others more. This improved taxonomic frame is a prerequisite to delve into more precise research questions, such as those concerning community ecology, biogeography and macroecology, and to reassess previous findings.

The global aims of my thesis were to 1) improve and develop molecular methods to study the amoebozoan diversity and ecology, 2) estimate their taxonomic and functional diversity in the soil, 3) improve the taxonomic and phylogenic frame for this diversity in order to build a sound basis for further research and 4) characterize the ecological drivers which are likely to influence microbial diversity at local, continental and global scales.

7.2 Methodological achievements

7.2.1 Development of a new marker

Several studies demonstrated that common molecular markers such as 18S, ITS and actin are not efficient to discriminate closely related species (Lara et al., 2008; Nassonova et al., 2010) and do not allow an adequate detection of Amoebozoa in metabarcoding surveys (Geisen et al., 2019; Stephenson et al., 2011). We confirmed (Chapter 4, 5, 6 and Appendix D) previous evidences that COI marker was indeed adapted for molecular investigations of Hyalospheniformes (Kosakyan et al., 2012), however different specific primers often have to be developed in function of the group targeted (Kudryavtsev, 2014). Therefore we designed a new molecular marker, localized between two mitochondrial nicotinamide adenine dinucleotide hydrogenase genes (i.e. NAD9 and NAD7), in order to improve both taxonomic resolution and molecular phylogeny (Chapter 2). With this marker we successfully amplified and sequenced specimens from several Arcellinida taxa and confirmed its potential for high taxonomical resolution, due to a fast-varying intergenic region, as well as for deep phylogeny due to the slow evolving nature of these genes. Moreover, conserved syntenic loci in the primers annealing regions allowed good PCR amplification of distantly related organisms. In addition, we also confirmed the efficiency of this marker to disentangle "wastebin" genera such as *Difflugia* by examining members of two common morphotypes of this genus(i.e. the "Pyriform" and "Lanceolate" clades) and attest that they were phylogenetically clearly differentiated (Appendix A).

7.2.2 Investigation of genus *Nebela* with HTS metabarcoding

Then we used the genetic COI marker to reveal unknown diversity of genus *Nebela*. Protocols were initially based on a cloning-sequencing approach applied to *Sphagnum* mosses, with a preliminary filtration/sedimentation step to concentrate cells and specific COI primers for genus *Nebela* (Kosakyan et al., 2013; Singer et al., 2018). Here, we adapted these protocols further for the application of HTS metabarcoding (Illumina MiSeq), which allows large sampling designs for a limited cost and working time (Chapter 6). Furthermore, we adapted the bioinformatics pipeline from Seppey et al. (2019) to clean the sequences and build the OTU. Although HTS metabarcoding based on COI has been applied to monitor animal communities in several ecosystems, including peat bogs (Garcés-Pastor et al., 2019), we are not aware of any application of this methodology to protist diversity. Therefore, our approach is novel and very promising as it allows to combine the advantages of both high taxonomical resolution of COI marker and huge sequencing depth of HTS.

7.3 Methodological limitations and perspectives

7.3.1 DNA barcoding

DNA barcoding will certainly remain a valuable and accessible approach to characterize the many Hyalospheniformes still uninvestigated. However, several drawbacks will have to be improved to increase DNA barcoding efficiency and sequences quality, while decreasing the cost of analysis and time of manipulations. Indeed, the COI barcoding efficiency is currently still rather low for amoebozoans (at least for genera *Nebela, Apodera, Alocodera* and *Padaungiella*) and even lower for NAD. Moreover, PCR biases and sequencing errors are frequent, and have to be manually corrected by observing other sequences or inferred based on the genetic code. These corrections may lead to a wrong interpretation of the data and possibly overestimate richness. The use of high fidelity proofreading enzymes such as *Phusion* polymerases may improve both accuracy and yield. This polymerase is expensive, but its great efficiency could balance the numerous failures previously obtained with cheaper enzyme for a better quality. In addition, the Nanopore (Oxford Nanopore) technology may improve the sequencing process by yielding fast results at lower cost (compare to cloning library on Sanger). Finally, the design of new specific primers, adapted to various taxonomical levels (i.e. from genus to order), will certainly improve the success rate of barcoding and permit to investigate new taxa.

7.3.2 Metabarcoding

We demonstrated that the metabarcoding approach is adapted for large scale environmental diversity and ecological studies (Chapter 3 and 6, Apendixe C). However, I am convinced that the development and adaptation of new environmental methods and technologies, e.g. as reviewed by Geisen et al. (2019), will expand the range of future research possibilities and allow the improvement of microbial diversity understanding. For instance, Nanopore and PacBio (Pacific Biosciences of California) are promising technologies because they can extend the sequences length compare to other existing plateforms, which may improve taxonomical resolution, but also permit to amplify other markers (e.g. complete 18S or COI maker) (Jamy et al., 2019). Finally, the recent development of 'omics' and especially multi-'omics' approaches (i.e. metagenomics and metatranscriptomic) could solve the issue of bias of under-covered taxonomical groups of previous metabarcoding approaches by performing PCR-free amplification. Indeed, a metatranscriptomic survey of beech litter has already revealed unprecedented diversity of Amoebozoa (Voss et al., 2019), however one main inconvenient of the use of this methods at large scale is their important cost (Geisen et al., 2019).

7.4 Diversity achievements, limitations and perspectives

7.4.1 Characterization of new environmental clade

There is a need for characterizing morphologically and functionally the huge environmental diversity of microbial eukaryote revealed by environmental DNA surveys (de Vargas et al., 2015). This must be achieved by so-called traditional methods that rely on obtaining pure cultures or isolation of single cells for those organisms that cannot be cultured (such as many Arcellinida). This, of course, requires a good knowledge of protist morphological diversity and ecology, i.e. good naturalist skills (Heger et al., 2014). These organisms can be barcoded, which allow further interpretation of molecular diversity such as determining relationship with already described taxa or differentiating morphotypes in coherent units. Nowadays, in addition, new tools can help in this quest for the characterization of molecular diversity. On the one hand, single cell genomics coupled with microscopy gave interesting results for characterizing Diplonemida, a clade of elusive marine planktonic heterotrophic flagellates particularly difficult to isolate (Okamoto et al., 2019). On the other hand, organisms that have been detected molecularly can be documented *a posteriori* using fluorescent in situ hybridization (FISH), as exemplified for the testate amoeba *Micropyxidiella edaphonis* (Tarnawski and Lara, 2015).

In Chapter 3, we isolated, cultivated and described the first member of the environmental clade LKM74 (van Hannen et al., 1999) that revealed to be a tiny naked amoeba belonging to the Dermamoebidae. This genus has been frequently detected in several freshwater and soils environments (Di Filippo et al., 2015; Lesaulnier et al., 2008; Richards et al., 2005), but also in peat bogs (Lara et al., 2011b). Corsaro and Venditti (2013) discovered a small amoeba molecularly related to this clade in samples from freshwater sediments, but could neither maintain nor describe the organisms and its life cycle. Surprisingly, we discovered that this amoeba presented a unique life cycle based on an alternation of phagotrophic active trophozoites and osmotrophic fungi-like ramifications. We therefore erected a new genus, named *Mycamoeba* based on its similarity to mycelium, and described the species *Mycamoeba gemmipara* as type. This example shows that even deep branching amoeboid clades remain to be discovered, but also, and most importantly that novel life cycles and lifestyles still remain to be unveiled.

7.4.2 Hyalospheniformes diversity overtake all estimations

At finer taxonomic levels, we confirmed in most chapters that the diversity of Hyalospheniformes was indeed greatly underestimated by traditional methods and that barcoding was a very effective approach to discriminate objectively closely related species. Indeed, the species number in genus *Nebela* has been continuously increasing since the first barcoding studies with the description of *N. guttata, N. pechorensis, N. aliciae* (Kosakyan et al., 2013), *N. gimlii* and the three unknown phylotypes (UPI, UPII and UPIII) (Singer et al., 2015). Four new phylotypes (SPIII, SPXII, SPXIV and SPXV) were discovered in our elevation gradient (Chapter 6) which suggests that the genus *Nebela* is probably composed of more than 15 species, although only eight have been formally described so far with molecular and morphological methods.

Heger et al. (2013) identified 12 molecular lineages in the *Hyalosphenia papilio* species complex of cryptic species based on delineation statistics. Although it is possible that several lineages remain to be discovered, we retrieved only one additional species despite a globally extensive sampling, suggesting that we have now captured most of the group's diversity (Appendix D). Finally, our most striking findings of gaps in diversity estimates were certainly in *Apodera vas* and *Alocodera cockayni*. Indeed, despite some early suggestions of numerous morphotypes (Deflandre, 1936) and even some species description (Jung, 1942), these two taxa are still usually assumed to be monospecific (Foissner, 2006). In chapter 5, we show new morphological and molecular evidence of numerous potential species. In particular, we discovered several large *Apodera*-related morphotypes in possession of a large and obvious keel, surrounding the entire fundus of the test, in different sites in New Zealand. Similarly, we also investigated several new morphotypes of *Alocodera* and obtained clear molecular distinctions with the previous barcoded specimens from Tierra del Fuego and Chilean Patagonia respectively (Kosakyan et al., 2012). In these Southern Hemisphere genera, new diversity is not only cryptic or pseudocryptic as in the taxa discussed previously, but obvious morphological differences corroborated by consistent genetic distances do separate taxa. Southern Hemisphere testate amoeba diversity is still poorly known and more basic surveys are needed (Fernández et al., 2015).

The diversity of Hyalospheniformes is expected to be increasingly revealed by metabarcoding surveys, but also by exploring other regions and environments than Holarctic peatlands. In particular, temperate and tropical forests should possess important richness, at least in humid habitats, that will differs greatly from peatlands because of more nutrient and different mineral composition. These habitats have been shown to harbour for instance the genus *Porosia* (Bobrov and Kosakyan, 2015; Bonnet, 1990; Todorov, 2002) that has never been characterized with COI marker and that could reveal several cryptic or pseudo-cryptic species similarly to genera *Nebela* and *Hyalosphenia*. In addition, it remain to be demonstrate that specimens of *Nebela* inhabiting forest mosses and litter truly belong to this genus.

7.4.3 Hyalospheniformes diversity characterization

Besides the exploration of new diversity it is also important to redefine the diversity already described, especially species description that relied on approximate drawing and lacked type definition. Therefore, in chapter 4 we redescribed the iconic hyalospheniid species *Nebela militaris* (Penard, 1890), with up to date molecular and morphological methods, and observed that its placement among genus *Nebela* was incorrect as this species was shown to branch as a sister clade of genus *Planocarina*, albeit with low support. Therefore, we erected the genus *Alabasta* to accommodate it as well as two other species that shared its characteristic features, i.e. narrow or elongated piriform shape with a deeply curved pseudostome. During this project, we reviewed the taxonomic history of *A. militaris* and discovered that this species had already been described nine years before Penard (1890) by Taranek (1881) under the name *Nebela bursella* Vejdovský. To resolve this issue we decided to synonymize *Alabasta* (*Nebela*) *militaris* and *Nebela bursella* and to conserve the junior name valid because *N. bursella* has not been used after 1964 whereas *N. militaris* has been consistently used for more than one century. In order to operate this taxonomic issue properly, we invoked the article 23.9.3 of the International Code of Zoological Nomenclature and reviewed the context (Appendix B).

Besides Hyalospheniformes, there are many incertae sedis arcellinid taxa that remain to be investigated and redefined such as *Argynnia*, (Vucetich, 1974), *Heleopera*, (Leidy, 1879), *Pseudonebela*, (Gauthier-Liévre, 1953) and *Physochila*, (Jung, 1942; Kosakyan et al., 2016a). I expect that further taxonomical and systematical studies on “wastebin” genera, such as *Difflugia*, will permit to delineate precisely monophyletic and morphologically homogeneous clades and improved, *in fine*, survey on this group. Finally, developing a sound taxonomy and phylogeny is of primordial importance to obtain reproducible ecological and biogeographical surveys (Mitchell and Meisterfeld, 2005).

7.5 Ecology and biogeography achievements, limitations and perspectives

Hyalospheniformes were taken as a model group for studying the distribution patterns in amoebozoan protists at the continental scale (chapter 5 and appendix D) and at the local scale along an elevation gradient (chapter 6). Biogeography and macroecology are two disciplines that developed over >100 years based on plants and animals (Ebach, 2015; Humboldt, 1807; Wallace, 1876). Most importantly, all the theoretical background and the hypotheses that have been formulated are based on species, a taxonomical rank that is complex to evaluate for microorganisms (Martiny et al., 2006; Prosser et al., 2007). In order to be able to test biogeographical or macroecological hypotheses on protists, it is thus of capital importance that these organisms could (1) be handy to sort individually (impossible with smaller protists) and (2) be discriminated at the species level, either molecularly or morphologically. Hyalospheniformes fulfill these requirements with their relatively large size that allows individual cell picking (Kosakyan et al., 2012), and also because COI proved to be efficient in identifying species (Singer et al., 2018).

7.5.1 Hyalospheniformes biogeographical patterns have been underestimated

In Chapter 5, we observed that genera *Apodera* and *Alocodera*, which were previously believed to be monospecific, are composed of species that have probably limited geographical distributions and, possibly, different ecological optima. The genus *Padaungiella* was already expected to include several species, some of them assumed to be cosmopolite with broad ecological niches. However, we detected only a single clade of *Padaungiella* that occurred in the Northern Hemisphere, while in New Zealand and Southern South America they appeared genetically more diverse. Finally, we also bring arguments to suggest that the Southern Hemisphere may most probably be the origin of the *Alocodera, Apodera* and *Padaungiella* clade . All these preliminary observations still need to be confirmed by larger sampling efforts, especially toward more underexplored regions (e.g. Africa and Asia) and diverse ecosystems, but already illustrate the present lack of knowledge on Southern Hemisphere testate amoeba diversity. Further investigations on these clades will certainly reveal a richness of Hyalospheniformes species comparable to the one of Northern Hemisphere, as well as new genera previously overlooked. Moreover, I expect this diversity to present biogeographical and macroecological patterns comparable to macroorganisms once examined on similar taxonomical resolution (Soininen, 2012).

In addition, our phylogeographical survey on the *Hyalosphenia papilio* species complex confirmed the non-random lineages distribution observed by (Heger et al., 2013). However, we completed their previous conclusion showing that climatic conditions are the main variable explaining distribution patterns. Indeed, although our analysis demonstrated also limited dispersal capacity of these organisms, we showed that historical factors, i.e. the influence of Quaternary glaciations on *Sphagnum*-peatland expansions and contractions, explain better the present distribution than climatic conditions or current suitable habitats repartition (Appendix D). We based this assumption on the evidence that highest richness is observed in ice-free refugia that allowed *Sphagnum* growth during glaciations (Shaw et al., 2013) and that most "cosmopolitan" lineages have a broader climatic niche. Further studies based on other Holarctic testate amoebae inhabiting this highly specific environment, such as genus *Nebela*, may confirm and clarify these findings.

These two projects, investigating specimens at higher taxonomical resolution, i.e. with molecular barcoding of the COI marker, revealed the first evidences of fine distribution patterns on Hyalospheniformes that were previously undetected based on morphological observations alone (Fontaneto and Hortal, 2013; Hanson et al., 2012; Mitchell and Meisterfeld, 2005). However in a larger scale, besides the few taxa investigated during this thesis, most Hyalospheniformes are still believed to be cosmopolitan. I expect that molecular data and larger databases will reveal more limited restriction and clarify some dubious findings such as *Alabasta kivuense* geographical distribution. Indeed, this species was first described from Congo (Gauthier-Lièvre and Thomas, 1961) but recently reported in Canada (Nicholls, 2015), which lead Nicholls to propose that in this case the assumed endemism was the consequence of species rarity. However, this may also be another case of cryptic species with different geographical distribution and more work is required to clarify this.

7.5.2 Ecological patterns highlighted in genus *Nebela*

Our analysis of the diversity of genus *Nebela* along an elevation gradient showed that distribution patterns of some Hyalospheniformes at least are correlated with altitude (Chapter 6), just like macroscopic organisms (Soininen, 2012). Indeed, the abundance and diversity of species of this genus tended to decrease in high elevation and the richness seemed to correspond to a typical mid-domain effect (Colwell and Lees, 2000). In addition, we noticed that the species previously described are most abundant and dominant at low and mid elevation, whereas previously unknown species were rare and occurred mostly at high elevation. Finally, our co-occurrence analyzes indicated that some species present competitive exclusion, with low competitors restricted to the higher elevation.

Altogether, these results demonstrate that investigating microbial diversity at fine taxonomical resolution is required to investigate pattern of distribution at regional scale. Further investigations with this metabarcoding approach on other Hyalospheniformes and others environments will allow to determine if our findings are unique to highly restricted habitas such as peatlands or if this is characteristic of all elevation gradients. Finally, this methods could similarly been used to survey the impact of latitudinal gradients and to test accurate ecological hypotheses (Fernández et al., 2016; Prosser et al., 2007).

7.6 Conclusion

The Amoebozoa are one of the most numerous, diverse and characteristic group in soil, thus playing important roles in ecosystems functioning. However, their taxonomy has always been difficult to investigate due to the lack of stable morphological traits. The recent developments of molecular tools such as metabarcoding have revealing large new deep environmental clades and inflate their diversity estimates (Mahé et al., 2017; Santos et al., 2020; van Hannen et al., 1999). But this discovery of hidden diversity was usually not followed by taxonomic refinement, leading to many uncharacterized clades. Testate amoebae, such as Hyalospheniformes, are comparatively easier to identify because of their rigid test (Meisterfeld, 2002). In addition, testate amoebae have often been used as biogeographical and ecological protist models thanks to their large size and the possibility to observe them easily in the environment (Foissner, 2006; Krashevska et al., 2007; Smith et al., 2008). However, barcoding tools have recently unveiled important pseudocryptic or cryptic species diversity within most taxa (Kosakyan et al., 2013; Mulot et al., 2017; Singer et al., 2015), which could contradict previous findings.

In this thesis, we improved the characterization of several amoebozoans, especially within the Hyalospheniformes taxa, which principally resulted in the description of two genera and one species. In addition, we demonstrated that traditional isolation and cultivation methods (Ekelund and Rønn, 1994) were still efficient and required for this purpose. Moreover, we developed the new mitochondrial marker NAD that could improve, in addition to COI, the characterization of the Hyalospheniformes. Further investigations using barcoding might reveal morphological traits, which then allow discrimination between species presenting a continuum variation of shape, such as *Padaungiella lageniformis* (Luketa, 2015, 2017b). Our results from chapter 5 and 6 showed that Hyalospheniformes diversity suggest narrow biogeographical patterns, and demonstrated ecological variation in community structures, along an elevation gradient. These findings are similar to those obtained in recent studies on this infraorder (Heger et al., 2013; Singer et al., 2018, 2019).

More globally, these results confirmed that both endemic and cosmopolitan species exist in Amoebozoa (Dagamac et al., 2017), in accordance with the moderate endemicity hypothesis (Foissner, 2007). More importantly, these diversity patterns suggest that some protists have more in common with macroscopic organisms than previously assumed (Martiny et al., 2006; Hanson et al., 2012; Soininen, 2012). Indeed, similar spatial structures were observed between animals, plants and protists between Neotropical forests, both at fine taxonomic and phylogenetic levels (Lentendu et al., 2018). Microscopic animals such as rotifers, nematodes and tardigrades present also broad diversity in biogeographical patterns with some confirmed endemic and cosmopolite taxa (Fontaneto, 2019). More globally this implies that our ecological perception of microbial organisms' complexity was generally underestimated because of their important morphological crypticity (Fontaneto et al., 2009).

Penard used to refer to Amoebozoa as “animal” and believed that they display a capacity of reasoning, which is to some extent similar to “true” animals (Penard, 1938, 1941). Similarly, Leidy was very enthusiastic about its description of *Hyalosphenia papilio* and then wrote “From its delicacy and transparency, its bright colors and form, as it moves among the leaves of *sphagnum*, desmids, and diatoms, I have associated it with the idea of a butterfly hovering among flowers” (Leidy, 1879). With this thesis, I hope that I succeeded too in demonstrating that these organisms are indeed fascinating and deserve as much attention as animals or plants.

Bibliography

Adl, S. M. and Gupta, V. V. Protists in soil ecology and forest nutrient cycling. *Canadian Journal of Forest Research*, 36(7):1805–1817, 2006. ISSN 0045-5067. Page 2

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., Soo Park, J., Pawlowski, J. W., Powell, M. J., Richter, D. J., Rueckert, S., Shadwick, L., Shimano, S., Spiegel, F. W., Torruella, G., Youssef, N., Zlatogursky, V., and Zhang, Q. Revisions to the classification, nomenclature, and diversity of eukaryotes. *Journal of Eukaryotic Microbiology*, 66(1):4–119, 2019. ISSN 1066-5234. Page 4

Amacker, N., Mitchell, E. A. D., Ferrari, B. J. D., and Chèvre, N. Development of a new ecotoxicological assay using the testate amoeba *Euglypha rotunda* (Rhizaria; Euglyphida) and assessment of the impact of the herbicide S-metolachlor. *Chemosphere*, 201:351–360, 2018. ISSN 0045-6535. doi:https://doi.org/10.1016/j.chemosphere.2018.03.001. Page 41

Amesbury, M. J., Swindles, G. T., Bobrov, A., Charman, D. J., Holden, J., Lamentowicz, M., Mallon, G., Mazei, Y., Mitchell, E. A. D., Payne, R. J., Roland, T. P., Turner, T. E., and Warner, B. G. Development of a new pan-European testate amoeba transfer function for reconstructing peatland palaeohydrology. *Quaternary Science Reviews*, 152:132–151, 2016. ISSN 0277-3791. doi:https://doi.org/10.1016/j.quascirev.2016.09.024. Page 41

Amesbury, M. J., Charman, D. J., Fyfe, R. M., Langdon, P. G., and West, S. Bronze Age upland settlement decline in southwest England: testing the climate change hypothesis. *Journal of Archaeological Science*, 35 (1):87–98, 2008. ISSN 0305-4403. doi:10.1016/j.jas.2007.02.010. Page 4

Amesbury, M. J., Booth, R. K., Roland, T. P., Bunbury, J., Clifford, M. J., Charman, D. J., Elliot, S., Finkelstein, S., Garneau, M., Hughes, P. D. M., Lamarre, A., Loisel, J., Mackay, H., Magnan, G., Markel, E. R., Mitchell, E. A. D., Payne, R. J., Pelletier, N., Roe, H., Sullivan, M. E., Swindles, G. T., Talbot, J., van Bellen, S., and Warner, B. G. Towards a Holarctic synthesis of peatland testate amoeba ecology: Development of a new continental-scale palaeohydrological transfer function for North America and comparison to European data. *Quaternary Science Reviews*, 201:483–500, 2018. ISSN 0277-3791. doi:https://doi.org/10.1016/j.quascirev.2018.10.034. Page 5, Page 9

Anderson, O. R. Abundance of terrestrial gymnamoebae at a northeastern US site: A four-year study, including the El Nino winter of 1997-1998. *Journal of Eukaryotic Microbiology*, 47(2):148–155, 2000. Page 10

Anderson, O. R. Laboratory and field-based studies of abundances, small-scale patchiness, and diversity of gymnamoebae in soils of varying porosity and organic content: evidence of microbiocoenoses. *Journal of Eukaryotic Microbiology*, 49(1):17–23, 2002. ISSN 1066-5234. Page 10

Anderson, O. R. A half-century of research on free-living amoebae (1965-2017): Review of biogeographic, ecological and physiological studies. *Acta Protozoologica*, 57(1):1–28, 2018. ISSN 1689-0027. Page 11

Avery, S. V., Harwood, J. L., and Lloyd, D. Quantification and characterization of phagocytosis in the soil amoeba *Acanthamoeba castellanii* by flow cytometry. *Appl. Environ. Microbiol.*, 61(3):1124–1132, 1995. ISSN 0099-2240. Page 2

Awerintzew, S. Die Susswasserrhizopoden. *Imperatorskago Sankt-Peterburgskago Obschestva Estestvoisptatelei Trud*, 36:1–351, 1906. Page 44

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyerreil, L. A., and Thingstad, F. The ecological role of water-column microbes in the sea. *Marine Ecology Progress Series*, 10(3):257–263, 1983. ISSN 0171-8630. doi:Doi 10.3354/Meps010257. Page 2, Page 31

Baas Becking, L. Geobiologie of inleiding tot de milieukunde. *WP Van Stockum and Zoon, The Hague, the Netherlands*, 1934. Page 8, Page 51

Barrett, R. D. H. and Hebert, P. D. N. Identifying spiders through DNA barcodes. *Canadian Journal of Zoology*, 83(3):481–491, 2005. ISSN 0008-4301. Page 14

Barth, D., Krenek, S., Fokin, S. I., and Berendonk, T. U. Intraspecific genetic variation in *Paramecium* revealed by mitochondrial cytochrome c oxidase I sequences. *Journal of Eukaryotic Microbiology*, 53(1):20–25, 2006. ISSN 1066-5234. doi:https://doi.org/10.1111/j.1550-7408.2005.00068.x. Page 14

Bates, S. T., Clemente, J. C., Flores, G. E., Walters, W. A., Parfrey, L. W., Knight, R., and Fierer, N. Global biogeography of highly diverse protistan communities in soil. *The ISME journal*, 7(3):652, 2013. ISSN 1751-7370. doi:10.1038/ismej.2012.147. Page 13

Baumgartner, M., Yapi, A., Gröbner-Ferreira, R., and Stetter, K. O. Cultivation and properties of *Echinamoeba thermarum* n. sp., an extremely thermophilic amoeba thriving in hot springs. *Extremophiles*, 7:267–274, 2003. Page 3

Berney, C., Fahrni, J., and Pawlowski, J. How many novel eukaryotic'kingdoms'? Pitfalls and limitations of environmental DNA surveys. *BMC biology*, 2(1):1, 2004. ISSN 1741-7007. Page 19

Berney, C., Geisen, S., Van Wichelen, J., Nitsche, F., Vanormelingen, P., Bonkowski, M., and Bass, D. Expansion of the 'reticulosphere': diversity of novel branching and network-forming amoebae helps to define Variosea (Amoebozoa). *Protist*, 166(2):271–295, 2015. ISSN 1434-4610. Page 19

Bian, P., Strano, J., Zheng, P., Steinitz-Kannan, M., Clarson, S. J., Kannan, R., and McCarthy, T. J. Amoebae assemble synthetic spherical particles to form reproducible constructs. *Langmuir*, 35(14):5069–5074, 2019. ISSN 0743-7463. doi:DOI: 10.1021/acs.langmuir.9b00333. Page 11

Blandenier, Q., Lara, E., Mitchell, E. A. D., Alcântara, D. M. C., Siemensma, F. J., Todorov, M., and Lahr, D. J. G. NAD9/NAD7 (mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene)—A new "Holy Grail" phylogenetic and DNA-barcoding marker for Arcellinida (Amoebozoa)? *European journal of protistology*, 58:175–186, 2017a. ISSN 0932-4739. Page 14, Page 41, Page 52

Blandenier, Q., Seppey, C. V. W., Singer, D., Vlimant, M., Simon, A., Duckert, C., and Lara, E. *Mycamoeba gemmipara* nov. gen., nov. sp., the first cultured member of the environmental Dermamoebidae Clade LKM74 and its unusual life cycle. *Journal of Eukaryotic Microbiology*, 64(2):257–265, 2017b. ISSN 1066-5234. Page 4, Page 10, Page 12

Blandenier, Q., Duckert, C., Kupferschmid, F. A. L., Kosakyan, A., Mitchell, E. A. D., Lara, E., and Singer, D. En garde! Redefinition of *Nebela militaris* (Arcellinida, Hyalospheniidae) and erection of *Alabasta* gen. nov. *European Journal of Protistology*, 66:156–165, 2018. ISSN 0932-4739. doi:https://doi.org/10.1016/j.ejop.2018.08.005. Page 6, Page 11, Page 51, Page 52

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., and Abebe, E. Defining operational taxonomic units using DNA barcode data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462):1935–1943, 2005. ISSN 0962-8436. doi:https://doi.org/10.1098/rstb.2005.1725. Page 12

Blaxter, M. L. The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1444):669–679, 2004. ISSN 0962-8436. doi:10.1098/rstb.2003.1447. Page 11, Page 13

Bobrov, A. and Kosakyan, A. A new species from mountain forest soils in Japan: *Porosia paracarinata* sp. nov., and taxonomic concept of the genus *Porosia* Jung, 1942. *Acta Protozoologica*, 2015(4):289–294, 2015. ISSN 1689-0027. doi:DOI: 10.4467/16890027AP.15.024.3538. Page 5, Page 7, Page 74

Bonnet, L. L'émission pseudopodique chez les thécamoebiens endogés (I). *Bulletin de la Société Zoologique de France*, 86:17–28, 1961. Page 4

Bonnet, L. Le peuplement thécamobiens des sols. *Revue d'Écologie et de Biologie du Sol*, 1(2):123–408, 1964. Page 4, Page 5, Page 7

Bonnet, L. Le peuplement thécamoebien des mousses corticoles. *Protistologica*, 9(3):319–338, 1973. Page 7

Bonnet, L. Faunistique et biogeographie des thécamoebiens. II. Thécamoebiens des sols de la Guadeloupe et de la Martinique. *Bull. soc. Hist. Nat. Toulouse*, 113:45–49, 1977a. Page 9

Bonnet, L. Le peuplement thécamoebien des sols du Népal et son intérêt biogéographique. *Extrait du bulletin de la société d'hisstoire naturelle de Toulouse*, 1977b. Page 9

Bonnet, L. Thécamoebiens (Rhizopoda, Testacea). *Biologie de sols. CNFRA*, 48:23–32, 1981. Page 9

Bonnet, L. Données écologiques sur quelques Hyalospheniidae et Paraquadrulidae (Thécamoebiens) des sols (2ème partie : genre *Nebela*). *Bulletin de la Société d'Histoire Naturelle de Toulouse*, 126:9–17, 1990. Page 7, Page 8, Page 74

Booth, R. K. Ecology of testate amoebae (Protozoa) in two lake superior coastal wetlands: Implications for paleoecology and environmental monitoring. *Wetlands*, 21(4):564–576, 2001. Page 7

Borcard, D. and Matthey, W. La vie animale des tourbières Neuchâteloises. *Bulletin de la Société Neuchâteloise de Sciences Naturelles*, 119:3–18, 1996. Page 63

Borcard, D. Les Oribates des tourbières du Jura suisse (Acari, Oribatei): écologie. II: Les relations Oribates-environnement à la lumière du test de Mantel. *Revue d'écologie et de biologie du sol*, 28(3):323–339, 1991. ISSN 0035-1822. Page 63

Bovee, E. C. Morphological identification of free-living Amoebida. In *Proceedings of the Iowa Academy of Science*, volume 60, pages 599–615, 1953. ISBN 0085-2236. Page 2, Page 11

Bragazza, L., Buttler, A., Robroek, B. J. M., Albrecht, R., Zaccone, C., Jassey, V. E. J., and Signarbieux, C. Persistent high temperature and low precipitation reduce peat carbon accumulation. *Global Change Biology*, 22(12):4114–4123, 2016. ISSN 1354-1013. Page 68

Brown, M. W., Spiegel, F. W., and Silberman, J. D. Amoeba at attention: phylogenetic affinity of *Sappinia pedata*. *Journal of Eukaryotic Microbiology*, 54(6):511–519, 2007. ISSN 1066-5234. doi:10.1111/j.1550-7408.2007.00292.x. Page 36

Brown, M. W., Spiegel, F. W., and Silberman, J. D. Phylogeny of the "forgotten" cellular slime mold, *Fonticula alba*, reveals a key evolutionary branch within Opisthokonta. *Molecular Biology and Evolution*, 26(12): 2699–2709, 2009. ISSN 0737-4038. doi:10.1093/molbev/msp185. Page 36

Brown, M. W. and Silberman, J. D. *The non-dictyostelid sorocarpic amoebae*, pages 219–242. Springer, 2013. Page 10

Brown, W. M., George, M., and Wilson, A. C. Rapid evolution of animal mitochondrial-DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 76(4):1967–1971, 1979. ISSN 0027-8424. doi:DOI 10.1073/pnas.76.4.1967. Page 14, Page 19

Bullerwell, C. E., Burger, G., Gott, J. M., Kourennaia, O., Schnare, M. N., and Gray, M. W. Abundant 5S rRNA-like transcripts encoded by the mitochondrial genome in amoebozoa. *Eukaryotic cell*, 9(5):762–773, 2010. ISSN 1535-9778. Page 21

Burger, G., Plante, I., Lonergan, K. M., and Gray, M. W. The mitochondrial-DNA of the ameboid protozoan, *Acanthamoeba-castellanii* - complete sequence, gene content and genome organization. *Journal of Molecular Biology*, 245(5):522–537, 1995. ISSN 0022-2836. doi:DOI 10.1006/jmbi.1994.0043. Page 21, Page 25

Burki, F., Roger, A. J., Brown, M. W., and Simpson, A. G. B. The new tree of eukaryotes. *Trends in ecology & evolution*, 2019. ISSN 0169-5347. Page 2

Cavalier-Smith, T., Chao, E. E. Y., and Oates, B. Molecular phylogeny of Amoebozoa and the evolutionary significance of the unikont *Phalansterium*. *European Journal of Protistology*, 40(1):21–48, 2004. ISSN 0932-4739. doi:10.1016/j.ejop.2003.10.001. Page 36

Cavalier-Smith, T. and Chao, E. E. Y. Phylogeny and classification of phylum Cercozoa (Protozoa). *Protist*, 154 (3-4):341–358, 2003. ISSN 1434-4610. doi:https://doi.org/10.1078/143446103322454112. Page 4

Certes, A. Mission scientifique du Cap Horn. Zoologie. Protozoaires. Gauthier-Villars & fils, Imprimeurs-Libraires de l'école polytechnique, du bureau des longitudes, Paris. *Ministères de la marine et de l'instruction publique*, 1889. Page 4

Chardez, D. *Difflugia oblonga* Ehrenberg et ses variétés. *Bull. Inst. Agron. Stat. Rech. Gembloux.*, new series 2: 589–595, 1967. doi:DOI 10.1006/jmbi.1994.0043. Page 20

Chardez, D. Protozoaires prédateurs de thécamoebiens. *Protistologica*, 21(2):187–194, 1985. Page 4

Charman, D. J. and Warner, B. G. The ecology of testate amoebae (Protozoa : Rhizopoda) in oceanic peatlands in Newfoundland, Canada: Modelling hydrological relationships for palaeoenvironmental reconstruction. *Ecoscience*, 4(4):555–562, 1997. Page 4, Page 7

Chase, M. W., Salamin, N., Wilkinson, M., Dunwell, J. M., Kesanakurthi, R. P., Haidar, N., and Savolainen, V. Land plants and DNA barcodes: short-term and long-term goals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462):1889–1895, 2005. ISSN 0962-8436. doi:10.1098/rstb.2005.1720. Page 13

Chomczynski, P. and Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate phenol chloroform extraction. *Analytical Biochemistry*, 162(1):156–159, 1987. ISSN 0003-2697. doi:DOI 10.1006/abio.1987.9999. Page 11, Page 20, Page 32, Page 42, Page 52

Clarholm, M. Soil protozoa: an under-researched microbial group gaining momentum. *Soil biology & biochemistry*, 2005. ISSN 0038-0717. Page 2

Clarholm, M. Protozoan grazing of bacteria in soil—impact and importance. *Microbial Ecology*, 7(4):343–350, 1981. ISSN 0095-3628. Page 2

Clymo, R. S. The growth of *Sphagnum*: some effects of environment. *The Journal of Ecology*, pages 849–869, 1973. Page 57

Cole, R. A. and Williams, K. L. The *Dictyostelium-discoideum* mitochondrial genome - a primordial system using the universal code and encoding hydrophilic proteins atypical of metazoan mitochondrial-DNA. *Journal of Molecular Evolution*, 39(6):579–588, 1994. ISSN 0022-2844. doi:Doi 10.1007/Bf00160403. Page 25

Colwell, R. K. and Lees, D. C. The mid-domain effect: geometric constraints on the geography of species richness. *Trends in ecology & evolution*, 15(2):70–76, 2000. ISSN 0169-5347. Page 76

Corsaro, D. and Venditti, D. Molecular phylogenetics evidence for a novel lineage of amoebae within Discosea (Amoebozoa: Lobosa). *Acta Protozoologica*, 52(4):309–316, 2013. ISSN 0065-1583. doi:10.4467/16890027AP.13.0028.1320. Page 31, Page 34, Page 35, Page 36, Page 74

Coûteaux, M.-M., Munsch, A., and Ponge, J.-F. Le genre *Euglypha*: essai de taxinomie numérique. 1979. Page 72

Couteaux, M. M. Relationships between testate amebas and fungi in humus microcosms. *Soil Biology & Biochemistry*, 17(3):339–345, 1985. ISSN 0038-0717. Page 4

Cowling, A. J. Free-living heterotrophic flagellates: methods of isolation and maintenance, including sources of strains in culture. *The biology of free-living heterotrophic flagellates. Oxford University Press, Oxford*, pages 477–492, 1991. Page 10

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., Potter, C., and Bik, H. M. The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution*, 7 (9):1008–1018, 2016. ISSN 2041-210X. doi:https://doi.org/10.1111/2041-210X.12574. Page 12

Cristóbal, E., Ayuso, S. V., Justel, A., and Toro, M. Robust optima and tolerance ranges of biological indicators: a new method to identify sentinels of global warming. *Ecological research*, 29(1):55–68, 2014. ISSN 0912-3814. Page 65

Dagamac, N. H. A., Novozhilov, Y. K., Stephenson, S. L., Lado, C., Rojas, C., dela Cruz, T. E. E., Unterseher, M., and Schnittler, M. Biogeographical assessment of myxomycete assemblages from Neotropical and Asian Palaeotropical forests. *Journal of biogeography*, 44(7):1524–1536, 2017. Page 77

Dahl, M. B., Shchepin, O., Schunk, C., Menzel, A., Novozhilov, Y. K., and Schnittler, M. A four year survey reveals a coherent pattern between occurrence of fruit bodies and soil amoebae populations for nivicolous myxomycetes. *Scientific Reports*, 8(1):11662, 2018. ISSN 2045-2322. doi:10.1038/s41598-018-30131-3. Page 13, Page 63

Darbyshire, J. F., Wheatley, R. E., Greaves, M. P., and Inkson, R. H. E. A rapid micromethod for estimating bacterial and protozoan populations in soil. *Revue d'Ecologie et de Biologie du Sol (France)*, 1974. ISSN 0035-1822. Page 10

Darwin, C. An account of the fine dust which often falls on vessels in the Atlantic Ocean. *Quarterly Journal of the Geological Society*, 2(1-2):26–30, 1846. ISSN 0370-291X. Page 8, Page 51

de Groot, G. A., Laros, I., and Geisen, S. *Molecular identification of soil eukaryotes and focused approaches targeting protist and faunal groups using high-throughput metabarcoding*, pages 125–140. Springer, 2016. doi:DOI 10.1007/978-1-4939-3369-3$_7$. Page 13

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J. M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horak, A., Jaillon, O., Lima-Mendez, G., Lukes, J., Malviya, S., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., 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., and Coordinators, T. O. Eukaryotic plankton diversity in the sunlit ocean. *Science*, 348(6237), 2015. ISSN 0036-8075. doi:Artn 1261605 10.1126/Science.1261605. Page 2, Page 13, Page 32, Page 73

Decelle, J., Romac, S., Sasaki, E., Not, F., and Mahé, F. Intracellular diversity of the V4 and V9 regions of the 18S rRNA in marine protists (radiolarians) assessed by high-throughput sequencing. *PloS one*, 9(8): e104297–e104297, 2014. ISSN 1932-6203. doi:10.1371/journal.pone.0104297. Page 13

Decloitre, L. *Thécamoebiens de la XIIe Expédition antarctique française*, volume 259. Expéditions polaires françaises. Missions Paul-Emile Victor, 1964. Page 7, Page 56

Decloître, L. Contribution à la faune du Congo (Brazzaville?). Mission A. Dexarpentries et A. Villiers. III. Rhizopodes Thécamoebiens. *Bulletin de l'Institut Français d'Afrique Noire, sér. A,*, 27:165–184, 1965. Page 5

Dedysh, S. N., Pankratov, T. A., Belova, S. E., Kulichevskaya, I. S., and Liesack, W. Phylogenetic analysis and in situ identification of bacteria community composition in an acidic *Sphagnum* peat bog. *Appl. Environ. Microbiol.*, 72(3):2110–2117, 2006. ISSN 0099-2240. Page 63

Deflandre, G. Materiaux pour la faune rhizopodique de France. III. *Bull. Soc. Zool. France*, 32:496–519, 1927. Page 7

Deflandre, G. Le genre *Arcella* Ehrenberg: Morphologie-Biologie. Essai phylogénétique et systématique. *Archiv für Protistenkunde*, new series 2:589–595, 1928. doi:DOI 10.1006/jmbi.1994.0043. Page 8, Page 20

Deflandre, G. Etude monographique sur le genre *Nebela* Leidy. *Annales de Protistologie*, 1936. Page 4, Page 5, Page 7, Page 8, Page 9, Page 54, Page 56, Page 74

del Campo, J., Mallo, D., Massana, R., de Vargas, C., Richards, T. A., and Ruiz-Trillo, I. Diversity and distribution of unicellular opisthokonts along the European coast analysed using high-throughput sequencing. *Environmental Microbiology*, 17(9):3195–3207, 2015. ISSN 1462-2912. doi:10.1111/1462-2920.12759. Page 31

del Campo, J., Sieracki, M. E., Molestina, R., Keeling, P., Massana, R., and Ruiz-Trillo, I. The others: our biased perspective of eukaryotic genomes. *Trends in ecology & evolution*, 29(5):252–259, 2014. ISSN 0169-5347. doi:https://doi.org/10.1016/j.tree.2014.03.006. Page 2

Di Filippo, M. M., Santoro, M., Lovreglio, P., Monno, R., Capolongo, C., Calia, C., Fumarola, L., D'Alfonso, R., Berrilli, F., and Di Cave, D. Isolation and molecular characterization of free-living Amoebae from different water sources in Italy. *International Journal of Environmental Research and Public Health*, 12(4):3417–3427, 2015. ISSN 1660-4601. doi:10.3390/ijerph120403417. Page 31, Page 74

Diaconu, A., Toth, M., Lamentowicz, M., Heiri, O., Kuske, E., Tanţău, I., Panait, A., Braun, M., and Feurdean, A. How warm? How wet? Hydroclimate reconstruction of the past 7500 years in northern Carpathians, Romania. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 482:1–12, 2017. ISSN 0031-0182. Page 41

Dieffenbach, C. W., Lowe, T. M. J., and Dveksler, G. S. General concepts for PCR primer design. *PCR-Methods and Applications*, 3(3):S30–S37, 1993. ISSN 1054-9803. Page 23

Douglas, T. E., Kronforst, M. R., Queller, D. C., and Strassmann, J. E. Genetic diversity in the social amoeba *Dictyostelium discoideum*: Population differentiation and cryptic species. *Molecular phylogenetics and evolution*, 60(3):455–462, 2011. ISSN 1055-7903. doi:https://doi.org/10.1016/j.ympev.2011.05.007. Page 11

Dumack, K., Baumann, C., and Bonkowski, M. A bowl with marbles: revision of the thecate amoeba genus *Lecythium* (Chlamydophryidae, Tectofilosida, Cercozoa, Rhizaria) including a description of four new species and an identification key. *Protist*, 167(5):440–459, 2016. ISSN 1434-4610. Page 10, Page 41, Page 45

Dumack, K., Kahlich, C., Lahr, D. J. G., and Bonkowski, M. Reinvestigation of *Phryganella paradoxa* (Arcellinida, Amoebozoa) Penard 1902. *Journal of Eukaryotic Microbiology*, 66(2):232–243, 2019. ISSN 1066-5234. doi:https://doi.org/10.1111/jeu.12665. Page 4, Page 10

Dvorský, M., Altman, J., Kopecký, M., Chlumská, Z., Řeháková, K., Janatková, K., and Doležal, J. Vascular plants at extreme elevations in eastern Ladakh, northwest Himalayas. *Plant Ecology & Diversity*, 8(4):571–584, 2015. ISSN 1755-0874. Page 68

Dykstra, M. J. and Keller, H. W. *Class Mycetozoa De Bary, 1859*, pages 952–981. Allen Press, Lawrence, Kansas, U.S.A, 2000. Page 36

Ebach, M. C. Origins of biogeography. *History, Philosophy and Theory of the Life Sciences*, 13:1–20, 2015. Page 75

Ehrenberg, C. G. Fortgesetzte Beobachtungen uber jetzt herrschende atmospharische mikroskopische Verhaltnisse. *Bericht uber die zur Bekanntmachung geeigneten Verhandlungen der Koniglichen Preussischen Akademie der Wissenschaften zu Berlin*, 13:370–381, 1848. Page 4

Ekelund, F. and Rønn, R. Notes on protozoa in agricultural soil with special emphasis on heterotrophic flagellates and naked amoebae and their ecology. *FEMS Microbiology Reviews*, 15:321–353, 1994. Page 2, Page 10, Page 11, Page 76

Evans, K. M., Wortley, A. H., and Mann, D. G. An assessment of potential diatom “barcode” genes (cox1, rbcL, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist*, 158(3):349–364, 2007. ISSN 1434-4610. doi:https://doi.org/10.1016/j.protis.2007.04.001. Page 14

Fahrni, J. F., Bolivar, I., Berney, C., Nassonova, E., Smirnov, A., and Pawlowski, J. Phylogeny of lobose amoebae based on actin and small-subunit ribosomal RNA genes. *Molecular Biology and Evolution*, 20(11):1881–1886, 2003. ISSN 0737-4038. doi:10.1093/molbev/msg201. Page 13

Fenchel, T. and Finlay, B. J. Is microbial diversity fundamentally different from biodiversity of larger animals and plants? *European Journal of Protistology*, 39(4):486–490, 2003. ISSN 0932-4739. Page 51

Fenchel, T., Esteban, G. F., and Finlay, B. J. Local versus global diversity of microorganisms: cryptic diversity of ciliated protozoa. *Oikos*, 80(2):220–225, 1997. ISSN 0030-1299. doi:DOI: 10.2307/3546589. Page 8, Page 9

Fernández, L. D., Fournier, B., Rivera, R., Lara, E., Mitchell, E. A. D., and Hernández, C. E. Water-energy balance, past ecological perturbations and evolutionary constraints shape the latitudinal diversity gradient of soil testate amoebae in south-western South America. *Global Ecology and Biogeography*, 25(10):1216–1227, 2016. Page 76

Fernández, L. D., Lara, E., and Mitchell, E. A. D. Checklist, diversity and distribution of testate amoebae in Chile. *European Journal of Protistology*, 51(5):409–424, 2015. ISSN 0932-4739. doi:https://doi.org/10.1016/j.ejop.2015.07.001. Page 9, Page 45, Page 74

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, 281(5374):237–240, 1998. doi:10.1126/science.281.5374.237. Page 2

Finlay, B. J. Global dispersal of free-living microbial eukaryote species. *Science*, 296(5570):1061–1063, 2002. ISSN 0036-8075. doi:DOI: 10.1126/science.1070710. Page 8

Finlay, B. J. and Fenchel, T. Divergent perspectives on protist species richness. *Protist*, 150(3):229–233, 1999. ISSN 1434-4610. doi:https://doi.org/10.1016/S1434-4610(99)70025-8. Page 8, Page 51

Finlay, B. J., Esteban, G. F., and Fenchel, T. Protist diversity is different? *Protist*, 155(1):15, 2004. ISSN 1434-4610. Page 51

Fiore-Donno, A. M., Nikolaev, S. I., Nelson, M., Pawlowski, J., Cavalier-Smith, T., and Baldauf, S. L. Deep phylogeny and evolution of slime moulds (Mycetozoa). *Protist*, 161(1):55–70, 2010. ISSN 1434-4610. doi:https://doi.org/10.1016/j.protis.2009.05.002. Page 13

Fiz-Palacios, O., Leander, B. S., and Heger, T. J. Old lineages in a new ecosystem: diversification of Arcellinid Amoebae (Amoebozoa) and peatland mosses. *PloS one*, 9(4), 2014. Page 57

Foissner, W. Soil Protozoa: fundamental problems, ecological significance. adaptations in ciliates and testaceans, bioindicators. and guide to the literature. *Prog. Protistol.*, 2:69–212, 1987. Page 10

Foissner, W. *Some new ciliates (Protozoa, Ciliophora) from an Austrian floodplain soil, including a giant, red „Flagship “, Cyrtohymena (Cyrtohymenide) aspoecki nov. subgen., nov. spec.* 2004. Page 9, Page 51

Foissner, W. Biogeography and dispersal of micro-organisms: A review emphasizing protists. *Acta Protozoologica*, 45(2):111–136, 2006. Page 9, Page 51, Page 74, Page 76

Foissner, W. Protist diversity: estimates of the near-imponderable. *Protist*, 150(4):363–368, 1999. ISSN 1434-4610. doi:https://doi.org/10.1016/S1434-4610(99)70037-4. Page 9

Foissner, W. Two new “flagship” ciliates (Protozoa, Ciliophora) from Venezuela: *Sleighophrys pustulata* and *Luporinophrys micelae*. *European Journal of Protistology*, 41(2):99–117, 2005. ISSN 0932-4739. Page 51

Foissner, W. Protist diversity and distribution: some basic considerations. In *Protist diversity and geographical distribution*, pages 1–8. Springer, 2007. Page 58, Page 77

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.*, 3:294–299, 1994. Page 14, Page 42, Page 52, Page 64

Fontaneto, D. Long-distance passive dispersal in microscopic aquatic animals. *Movement ecology*, 7(1):10, 2019. ISSN 2051-3933. Page 77

Fontaneto, D. and Hortal, J. At least some protist species are not ubiquitous. *Molecular ecology*, 22(20):5053–5055, 2013. ISSN 0962-1083. doi:https://doi.org/10.1111/mec.12507. Page 76

Fontaneto, D., Herniou, E. A., Boschetti, C., Caprioli, M., Melone, G., Ricci, C., and Barraclough, T. G. Independently evolving species in asexual bdelloid rotifers. *PLoS biology*, 5(4), 2007. Page 67

Fontaneto, D., Kaya, M., Herniou, E. A., and Barraclough, T. G. Extreme levels of hidden diversity in microscopic animals (Rotifera) revealed by DNA taxonomy. *Molecular Phylogenetics and Evolution*, 53(1):182–189, 2009. ISSN 1055-7903. Page 77

Fournier, B., Lara, E., Jassey, V. E. J., and Mitchell, E. A. D. Functional traits as a new approach for interpreting testate amoeba palaeo-records in peatlands and assessing the causes and consequences of past changes in species composition. *The Holocene*, 25(9):1375–1383, 2015. ISSN 0959-6836. doi:http://dx.doi.org/10.1177/0959683615585842. Page 7

Fučíková, K. and Lahr, D. J. G. Uncovering cryptic diversity in two amoebozoan species using complete mitochondrial genome sequences. *Journal of Eukaryotic Microbiology*, 63(1):112–122, 2016. ISSN 1066-5234. doi:10.1111/jeu.12253. Page 14, Page 19, Page 20

Gao, Y., Zhang, X., He, G., Shchepin, O. N., Yan, S., and Chen, S. Influence of forest type on dark-spored myxomycete community in subtropical forest soil, China. *Soil Biology and Biochemistry*, 138:107606, 2019. ISSN 0038-0717. doi:https://doi.org/10.1016/j.soilbio.2019.107606. Page 13

García-Martín, J. M., Zamora, J. C., and Lado, C. Evidence of intra-individual SSU polymorphisms in dark-spored Myxomycetes (Amoebozoa). *Protist*, 170(5):125681, 2019. ISSN 1434-4610. doi:https://doi.org/10.1016/j.protis.2019.125681. Page 13, Page 72

Garcés-Pastor, S., Wangensteen, O. S., Pérez-Haase, A., Pèlachs, A., Pérez-Obiol, R., Cañellas-Boltà, N., Mariani, S., and Vegas-Vilarrúbia, T. DNA metabarcoding reveals modern and past eukaryotic communities in a high-mountain peat bog system. *Journal of Paleolimnology*, 62(4):425–441, 2019. ISSN 0921-2728. doi:https://doi.org/10.1007/s10933-019-00097-x. Page 73

Gaston, K. J. Global patterns in biodiversity. *Nature*, 405(6783):220–227, 2000. ISSN 1476-4687. Page 63

Gauthier-Lièvre, L. and Thomas, R. Troisième note sur les Nebelinae d'Afrique. *Bull. Soc. Hist. Natur. Afrique du Nord.*, 52:41–48, 1961. Page 45, Page 47, Page 48, Page 76

Gauthier-Liévre, L. Les genres *Nebela, Paraquadrulea* et *Pseudonebela* (Rhizopodes testaces) en Afrique. *Bull. Soc. d'Hist. Natur. d'Afrique du Nord.*, 43:324–366, 1953. Page 8, Page 75

Gauthier-Liévre, L. Additions aux *Nebela* d'Afrique. *Bulletin de la Société d'Histoire Naturelle d'Afrique du Nord*, 48:494–525, 1957. Page 5

Geisen, S., Kudryavtsev, A., Bonkowski, M., and Smirnov, A. Discrepancy between species borders at morphological and molecular levels in the genus *Cochliopodium* (Amoebozoa, Himatismenida), with the description of *Cochliopodium plurinucleolum* n. sp. *Protist*, 165(3):364–383, 2014. ISSN 1434-4610. doi:10.1016/j.protis.2014.04.002. Page 19

Geisen, S., Tveit, A. T., Clark, I. M., Richter, A., Svenning, M. M., Bonkowski, M., and Urich, T. Metatranscriptomic census of active protists in soils. *Isme Journal*, 9(10):2178–2190, 2015a. ISSN 1751-7362. doi:10.1038/ismej.2015.30. Page 13, Page 19, Page 24, Page 31

Geisen, S., Mitchell, E. A. D., Adl, S., Bonkowski, M., Dunthorn, M., Ekelund, F., Fernández, L. D., Jousset, A., Krashevska, V., Singer, D., Spiegel, F. W., Walochnik, J., and Lara, E. Soil protists: a fertile frontier in soil biology research. *FEMS microbiology reviews*, 42(3):293–323, 2018. ISSN 1574-6976. Page 3, Page 45

Geisen, S., Laros, I., Vizcaíno, A., Bonkowski, M., and De Groot, G. A. Not all are free-living: high-throughput DNA metabarcoding reveals a diverse community of protists parasitizing soil metazoa. *Molecular Ecology*, 24(17):4556–4569, 2015b. ISSN 0962-1083. doi:doi: 10.1111/mec.13238. Page 13

Geisen, S., Rosengarten, J., Koller, R., Mulder, C., Urich, T., and Bonkowski, M. Pack hunting by a common soil amoeba on nematodes. *Environmental microbiology*, 17(11):4538–4546, 2015c. ISSN 1462-2912. Page 2, Page 4

Geisen, S., Koller, R., Hünninghaus, M., Dumack, K., Urich, T., and Bonkowski, M. The soil food web revisited: diverse and widespread mycophagous soil protists. *Soil Biology and Biochemistry*, 94:10–18, 2016. ISSN 0038-0717. Page 2

Geisen, S., Mitchell, E. A. D., Wilkinson, D. M., Adl, S., Bonkowski, M., Brown, M. W., Fiore-Donno, A. M., Heger, T. J., Jassey, V. E. J., Krashevska, V., Lahr, D. J. G., Marcisz, K., Mulot, M., Payne, R., Singer, D., Anderson, O. R., Charman, D. J., Ekelund, F., Griffiths, B. S., Rønn, R., Smirnov, A., Bass, D., Belbahri, L., Berney, C., Blandenier, Q., Chatzinotas, A., Clarholm, M., Dunthorn, M., Feest, A., Fernández, L. D., Foissner, W., Fournier, B., Gentekaki, E., Hájek, M., Helder, J., Jousset, A., Koller, R., Kumar, S., La Terza, A., Lamentowicz, M., Mazei, Y., Santos, S. S., Seppey, C. V. W., Spiegel, F. W., Walochnik, J., Winding, A., and Lara, E. Soil protistology rebooted: 30 fundamental questions to start with. *Soil Biology and Biochemistry*, 111:94–103, 2017. ISSN 0038-0717. doi:https://doi.org/10.1016/j.soilbio.2017.04.001. Page 2, Page 63

Geisen, S., Vaulot, D., Mahé, F., Lara, E., de Vargas, C., and Bass, D. A user guide to environmental protistology: primers, metabarcoding, sequencing, and analyses. *bioRxiv*, page 850610, 2019. Page 13, Page 72, Page 73

Gilbert, D. and Mitchell, E. A. D. *Microbial diversity in Sphagnum peatlands*, pages 287–317. Developments in Earth Surface Processes series. Elsevier, Amsterdam, 2006. ISBN ISBN-13: 978-0-444-52883-4 ISBN-10: 0-444-52883-0. Page 5

Gilbert, D., Amblard, C., Bourdier, G., Francez, A.-J., and Mitchell, E. A. D. Le régime alimentaire des thécamoebiens. *L'Année Biologique*, 39:57–68, 2000. Page 4, Page 5

Gile, G. H., Stern, R. F., James, E. R., and Keeling, P. J. DNA barcoding of Chlorarachniophytes using nucleomorph ITS sequences . *Journal of Phycology*, 46(4):743–750, 2010. ISSN 0022-3646. doi:https://doi.org/10.1111/j.1529-8817.2010.00851.x. Page 13

Glockner, G. *Genome analysis of social amoebae*, pages 35–47. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. ISBN 978-3-642-38487-5. doi:10.1007/978-3-642-38487-5$_2$. Page 23

Golemansky, V. Faune muscicole de Guinee forestiere (Rhisopodes testaces). *Recherches Africaines, Nouvelle série*, 4:33–60, 1962. Page 45

Gomaa, F., Todorov, M., Heger, T. J., Mitchell, E. A. D., and Lara, E. SSU rRNA phylogeny of Arcellinida (Amoebozoa) reveals that the largest Arcellinid genus, *Difflugia* Leclerc 1815, is not monophyletic. *Protist*, 163(3):389–399, 2012. ISSN 1434-4610. doi:10.1016/j.protis.2011.12.001. Page 4, Page 20, Page 24, Page 26

Gomaa, F., Kosakyan, A., Heger, T. J., Corsaro, D., Mitchell, E. A. D., and Lara, E. One alga to rule them all: unrelated mixotrophic testate amoebae (Amoebozoa, Rhizaria and Stramenopiles) share the same symbiont (Trebouxiophyceae). *Protist*, 165(2):161–176, 2014. ISSN 1434-4610. doi:10.1016/j.protis.2014.01.002. Page 3, Page 4, Page 42

Gomaa, F., Yang, J., Mitchell, E. A. D., Zhang, W. J., Yu, Z., Todorov, M., and Lara, E. Morphological and molecular diversification of asian endemic *Difflugia tuberspinifera* (Amoebozoa, Arcellinida): a case of fast morphological evolution in protists? *Protist*, 166(1):122–130, 2015. ISSN 1434-4610. doi:10.1016/j.protis.2014.11.004. Page 24, Page 26

Gomaa, F., Lahr, D. J. G., Todorov, M., Li, J., and Lara, E. A contribution to the phylogeny of agglutinating Arcellinida (Amoebozoa) based on SSU rRNA gene sequences. *European journal of protistology*, 59:99–107, 2017. Page 11

González-Resendiz, L., Johansen, J. R., León-Tejera, H., Sanchéz, L., Segal-Kischinevzky, C., Escobar-Sánchez, V., and Morales, M. A bridge too far in naming species: A total evidence approach does not support recognition of four species in *Desertifilum* (Cyanobacteria). *Journal of phycology*, 2019. ISSN 0022-3646. doi:https://doi.org/10.1111/jpy.12867. Page 14

Gottsberger, G. and Nannenga-Bremekamp, N. E. A new species of *Didymium* from Brazil. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Series C-Biological and Medical Sciences*, 74(3): 264–268, 1971. ISSN 0023-3374. Page 36

Gray, M. W. Evolutionary origin of RNA editing. *Biochemistry*, 51(26):5235–5242, 2012. ISSN 0006-2960. doi:10.1021/bi300419r. Page 26

Gray, M. W., Lang, B. F., and Burger, G. Mitochondria of protists. *Annual Review of Genetics*, 38:477–524, 2004. ISSN 0066-4197. doi:DOI 10.1146/annurev.genet.37.110801.142526. Page 19

Grebecki, A. Membrane and cytoskeleton flow in motile cells with emphasis an the contribution of free-living amoebae. *International Review of Cytology - a Survey of Cell Biology, Vol 148*, 148:37–80, 1994. ISSN 0074-7696. Page 19

Gregorich, E. G. and Carter, M. R. *Soil sampling and methods of analysis.* CRC press, 2007. ISBN 1420005278. Page 63

Greninger, A. L., Messacar, K., Dunnebacke, T., Naccache, S. N., Federman, S., Bouquet, J., Mirsky, D., Nomura, Y., Yagi, S., and Glaser, C. Clinical metagenomic identification of *Balamuthia mandrillaris* encephalitis and assembly of the draft genome: the continuing case for reference genome sequencing. *Genome medicine*, 7 (1):1, 2015. ISSN 1756-994X. Page 21

Grospietsch, T. Monographische Studie der Gattung Hyalosphenia Stein. *Hydrobiologia*, 26:211–241, 1965. Page 5

Grospietsch, T. Rhizopoda. Beitrag zur ökologie der testaceen Rhizopoden von Marion Island. *Marion and Prince Edward Islands*, 37:411–427, 1971. Page 10, Page 56

Grossmann, L., Jensen, M., Pandey, R. V., Jost, S., Bass, D., Psenner, R., and Boenigk, J. Molecular investigation of protistan diversity along an elevation transect of alpine lakes. *Aquatic Microbial Ecology*, 78(1):25–37, 2016. ISSN 0948-3055. Page 63

Guillou, L., Viprey, M., Chambouvet, A., Welsh, R. M., Kirkham, A. R., Massana, R., Scanlan, D. J., and Worden, A. Z. Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environmental Microbiology*, 10(12):3349–3365, 2008. ISSN 1462-2912. doi:10.1111/j.1462-2920.2008.01731.x. Page 31

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., and Decelle, J. The protist ribosomal reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic acids research*, 41(D1):D597–D604, 2012. ISSN 0305-1048. Page 12

Gužvić, M. The history of DNA sequencing/ISTORIJAT SEKVENCIRANJA DNK. *Journal of medical biochemistry*, 32(4):301–312, 2013. ISSN 1452-8266. doi:DOI: 10.2478/jomb-2014-0004. Page 12, Page 13

Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*, 41:95–98, 1999. Page 21, Page 42, Page 52, Page 64

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., and Martiny, J. B. H. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology*, 10(7):497–506, 2012. ISSN 1740-1534. doi:doi:10.1038/nrmicro2795. Page 63, Page 76, Page 77

Harder, C. B., Rønn, R., Brejnrod, A., Bass, D., Al-Soud, W. A., and Ekelund, F. Local diversity of heathland Cercozoa explored by in-depth sequencing. *The ISME journal*, 10(10):2488–2497, 2016. ISSN 1751-7370. Page 63

Harnisch, O. Weitere Daten zur Rhizopodenfauna Lapplands: Ein Beitrag zur Ökologie sphagnophiler Rhizopoden. *Zoologischer Anzeiger*, 124(5)(6):138–150, 1938. Page 45

Heal, O. W. Morphological variation in certain Testacea (Protozoa: Rhizopoda). *Archiv für Protistenkunde*, 106:351–368, 1963. Page 7

Hebert, P. D. N., Ratnasingham, S., and deWaard, J. R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B-Biological Sciences*, 270: S96–S99, 2003. ISSN 0962-8452. doi:10.1098/rsbl.2003.0025. Page 11, Page 14, Page 19, Page 41, Page 51, Page 72

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., and Hallwachs, W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences*, 101(41):14812–14817, 2004a. ISSN 0027-8424. Page 14

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S., and Francis, C. M. Identification of birds through DNA barcodes. *PLoS biology*, 2(10):e312, 2004b. ISSN 1545-7885. Page 14

Heger, T. J., Pawlowski, J., Golemanski, V., Todorov, M., Lara, E., and Mitchell, E. A. D. Comparing potential COI and SSU rDNA barcodes for assessing the diversity and phylogenetic relationships of cyphoderiid testate amoebae (Rhizaria: Euglyphida). *Protist*, 162(1):131–141, 2011a. doi:10.1016/j.protis.2010.05.002. Page 41

Heger, T. J., Mitchell, E. A. D., and Leander, B. S. Holarctic phylogeography of the testate amoeba *Hyalosphenia papilio (Amoebozoa: Arcellinida)* reveals extensive genetic diversity explained more by environment than dispersal limitation. *Molecular ecology*, 22(20):5172–5184, 2013. ISSN 1365-294X. Page 9, Page 10, Page 14, Page 41, Page 43, Page 46, Page 51, Page 74, Page 75, Page 77

Heger, T. J., Mitchell, E. A. D., Ledeganck, P., Vincke, S., Van de Vijver, B., and Beyens, L. The curse of taxonomic uncertainty in biogeographical studies of free-living terrestrial protists: a case study of testate amoebae from Amsterdam Island. *Journal of biogeography*, 36(8):1551–1560, 2009. ISSN 0305-0270. doi:doi:10.1111/j.1365-2699.2009.02094.x. Page 8, Page 10, Page 51, Page 63

Heger, T. J., Booth, R. K., Sullivan, M. E., Wilkinson, D. M., Warner, B. G., Asada, T., Mazei, Y., Meisterfeld, R., and Mitchell, E. A. D. Rediscovery of *Nebela ansata* (Amoebozoa: Arcellinida) in eastern North America: biogeographical implications. *Journal of biogeography*, 38(10):1897–1906, 2011b. ISSN 0305-0270. Page 9, Page 58

Heger, T. J., Lara, E., and Mitchell, E. A. D. Arcellinida testate amoebae (Amoebozoa: Arcellinida): model of organisms for assessing microbial biogeography. *Biogeography of Microscopic Organisms: Everything Small Everywhere*, pages 111–129, 2011c. Page 9

Heger, T. J., Edgcomb, V. P., Kim, E., Lukeš, J., Leander, B. S., and Yubuki, N. A resurgence in field research is essential to better understand the diversity, ecology, and evolution of microbial eukaryotes. *Journal of Eukaryotic Microbiology*, 61(2):214–223, 2014. ISSN 1066-5234. doi:doi:10.1111/jeu.12095. Page 73

Heger, T. J., Derungs, N., Theurillat, J.-P., and Mitchell, E. A. D. Testate amoebae like it hot: species richness decreases along a subalpine-alpine altitudinal gradient in both natural *Calluna vulgaris* litter and transplanted *Minuartia sedoides* cushions. *Microbial ecology*, 71(3):725–734, 2016. ISSN 0095-3628. Page 63

Heidel, A. J. and Glockner, G. Mitochondrial genome evolution in the social amoebae. *Molecular Biology and Evolution*, 25(7):1440–1450, 2008. ISSN 0737-4038. doi:10.1093/molbev/msn088. Page 19, Page 21, Page 25, Page 26

Hofstatter, P. G., Brown, M. W., and Lahr, D. J. G. Comparative genomics supports sex and meiosis in diverse Amoebozoa. *Genome biology and evolution*, 10(11):3118–3128, 2018. ISSN 1759-6653. Page 3

Hoogenraad, H. and De Groot, A. Die Geographische Verbreitung der Süsswasser-Rhizopoden. *Aquatic Ecology*, 13(2):152–171, 1979. Page 8

Humboldt, A. v. Le voyage aux régions equinoxiales du Nouveau Continent, fait en 1799–1804, par Alexandre de Humboldt et Aimé Bonpland, 1807. Page 75

Hurst, L. D., Hamilton, W. D., and Ladle, R. J. Covert sex. *Trends in ecology & evolution*, 7(5):144, 1992. ISSN 0169-5347. Page 3

Jamy, M., Foster, R., Barbera, P., Czech, L., Kozlov, A., Stamatakis, A., Bending, G., Hilton, S., Bass, D., and Burki, F. Long-read metabarcoding of the eukaryotic rDNA operon to phylogenetically and taxonomically resolve environmental diversity. *Molecular ecology resources*, 2019. ISSN 1755-098X. doi:https://doi.org/10.1111/1755-0998.13117. Page 73

Jassey, V. E. J., Shimano, S., Dupuy, C., Toussaint, M., and Gilbert, D. Characterizing the feeding habits of the testate amoebae *Hyalosphenia papilio* and *nebela* tincta along a narrow "fen-bog" gradient using digestive vacuole content and 13C and 15N isotopic analyses. *Protist*, 163(3):451–464, 2012. ISSN 1434-4610. Page 41

Jassey, V. E. J., Gilbert, D., Binet, P., Toussaint, M.-L., and Chiapusio, G. Effect of a temperature gradient on *Sphagnum fallax* and its associated living microbial communities: a study under controlled conditions. *Canadian journal of microbiology*, 57(3):226–235, 2011. ISSN 0008-4166. Page 57

Jassey, V. E. J., Meyer, C., Dupuy, C., Bernard, N., Mitchell, E. A. D., Toussaint, M.-L., Metian, M., Chatelain, A. P., and Gilbert, D. To what extent do food preferences explain the trophic position of heterotrophic and mixotrophic microbial consumers in a *Sphagnum* peatland? *Microbial ecology*, 66(3):571–580, 2013. ISSN 0095-3628. Page 4

Jung, W. Thekamöben ursprünglischer lebender deutscher Hochmoore. *Abhandlung Landesmuseum der Provinz Westfalen*, 7:1–87, 1936. Page 7

Jung, W. Illustrierte Thekamöben-Bestimmungstabellen. I. Die Systematik der Nebelinen. *Archiv für Protistenkunde*, 95(3):357–390, 1942. Page 5, Page 7, Page 55, Page 56, Page 74, Page 75

Kalla, S. E., Queller, D. C., Lasagni, A., and Strassmann, J. E. Kin discrimination and possible cryptic species in the social amoeba *Polysphondylium violaceum*. *BMC evolutionary biology*, 11(1):31, 2011. ISSN 1471-2148. doi:https://doi.org/10.1186/1471-2148-11-31. Page 11

Kang, S., Tice, A. K., Spiegel, F. W., Silberman, J. D., Pánek, T., Čepička, I., Kostka, M., Kosakyan, A., Alcântara, D. M. C., Roger, A. J., Shadwick, L. L., Smirnov, A., Kudryavtsev, A., Lahr, D. J. G., and Brown, M. W. Between a pod and a hard test: the deep evolution of amoebae. *Molecular biology and evolution*, 34(9):2258–2270, 2017. ISSN 0737-4038. Page 2, Page 46

Kappel, T. and Anken, R. H. An aquarium myxomycete: *Didymium nigripes*. *Mycologist*, 6:106–107, 1992. Page 36

Karpov, S. A., Mikhailov, K. V., Mirzaeva, G. S., Mirabdullaev, I. M., Mamkaeva, K. A., Titova, N. N., and Aleoshin, V. V. Obligately phagotrophic aphelids turned out to branch with the earliest-diverging fungi. *Protist*, 164(2):195–205, 2013. ISSN 1434-4610. doi:10.1016/j.protis.2012.08.001. Page 31

Keeling, P. J., Burger, G., Durnford, D. G., Lang, B. F., Lee, R. W., Pearlman, R. E., Roger, A. J., and Gray, M. W. The tree of eukaryotes. *Trends in ecology & evolution*, 20(12):670–676, 2005. ISSN 0169-5347. Page 2

Kent, W. *A manual of the infusoria, including a description of all known flagellate, ciliate, and tentaculiferous protozoa, British and foreign and an account of the organization and affinities of the sponges*. 1880. doi:https://doi.org/10.5962/bhl.title.1243. Page 4, Page 46

Koenig, I., Mulot, M., and Mitchell, E. A. D. Taxonomic and functional traits responses of *Sphagnum* peatland testate amoebae to experimentally manipulated water table. *Ecological Indicators*, 85:342–351, 2018. ISSN 1470-160X. Page 5, Page 7

Komac, B., Pladevall, C., Peñuelas, J., Conesa, J. V., and Domènech, M. Variations in functional diversity in snowbed plant communities determining snowbed continuity. *Plant Ecology*, 216(9):1257–1274, 2015. ISSN 1385-0237. Page 68

Kosakyan, A., Heger, T. J., Leander, B. S., Todorov, M., Mitchell, E. A. D., and Lara, E. COI barcoding of nebelid testate amoebae (Amoebozoa: Arcellinida): extensive cryptic diversity and redefinition of the Hyalospheniidae Schultze. *Protist*, 163(3):415–434, 2012. ISSN 1434-4610. doi:10.1016/j.protis.2011.10.003. Page 4, Page 6, Page 11, Page 19, Page 24, Page 41, Page 42, Page 43, Page 44, Page 45, Page 46, Page 51, Page 52, Page 53, Page 54, Page 55, Page 59, Page 72, Page 74, Page 75

Kosakyan, A., Gomaa, F., Mitchell, E. A. D., Heger, T. J., and Lara, E. Using DNA-barcoding for sorting out protist species complexes: a case study of the *Nebela tincta-collaris-bohemica* group (Amoebozoa; Arcellinida, Hyalospheniidae). *European Journal of Protistology*, 49(2):222–237, 2013. ISSN 0932-4739. doi:10.1016/j.ejop.2012.08.006. Page 6, Page 14, Page 19, Page 24, Page 26, Page 41, Page 43, Page 45, Page 51, Page 63, Page 64, Page 65, Page 66, Page 72, Page 74, Page 76

Kosakyan, A., Mulot, M., Mitchell, E. A. D., and Lara, E. Environmental DNA COI barcoding for quantitative analysis of protists communities: a test using the *Nebela collaris* complex (Amoebozoa; Arcellinida; Hyalospheniidae). *European Journal of Protistology*, 51(4):311–320, 2015. ISSN 0932-4739. doi:10.1016/j.ejop.2015.06.005. Page 20, Page 65

Kosakyan, A., Gomaa, F., Lara, E., and Lahr, D. J. G. Current and future perspectives on the systematics, taxonomy and nomenclature of testate amoebae. *European journal of protistology*, 2016a. ISSN 0932-4739. Page 8, Page 14, Page 19, Page 24, Page 26, Page 63, Page 75

Kosakyan, A., Lahr, D. J. G., Mulot, M., Meisterfeld, R., Mitchell, E. A. D., and Lara, E. Phylogenetic reconstruction based on COI reshuffles the taxonomy of hyalosphenid shelled (testate) amoebae and reveals the convoluted evolution of shell plate shapes. *Cladistics*, 32(6):606–623, 2016b. ISSN 1096-0031. Page 4, Page 6, Page 11, Page 14, Page 41, Page 43, Page 45, Page 56

Krashevska, V., Sandmann, D., Marian, F., Maraun, M., and Scheu, S. Leaf litter chemistry drives the structure and composition of soil testate amoeba communities in a tropical montane rainforest of the ecuadorian andes. *Microbial ecology*, 74(3):681–690, 2017. ISSN 0095-3628. Page 45

Krashevska, V., Malysheva, E., Klarner, B., Mazei, Y., Maraun, M., Widyastuti, R., and Scheu, S. Micro-decomposer communities and decomposition processes in tropical lowlands as affected by land use and litter type. *Oecologia*, pages 1–12, 2018. ISSN 0029-8549. Page 7, Page 41, Page 45

Krashevska, V., Bonkowski, M., Maraun, M., and Scheu, S. Testate amoebae (protista) of an elevational gradient in the tropical mountain rain forest of Ecuador. *Pedobiologia*, 51(4):319–331, 2007. ISSN 0031-4056. doi:10.1016/j.pedobi.2007.05.005. Page 9, Page 63, Page 76

Kudryavtsev, A. *Paravannella minima* n. g. n. sp (Discosea, Vannellidae) and distinction of the genera in the vannellid amoebae. *European Journal of Protistology*, 50(3):258–269, 2014. ISSN 0932-4739. doi:10.1016/j.ejop.2013.12.004. Page 14, Page 19, Page 35, Page 72

Kudryavtsev, A. and Pawlowski, J. *Cunea* n. g. (Amoebozoa, Dactylopodida) with two cryptic species isolated from different areas of the ocean. *European Journal of Protistology*, 51(3):197–209, 2015. ISSN 0932-4739. doi:10.1016/j.ejop.2015.04.002. Page 13, Page 26, Page 31, Page 35

Kudryavtsev, A., Bernhard, D., Schlegel, M., Chao, E. E. Y., and Cavalier-Smith, T. 18S ribosomal RNA gene sequences of *Cochliopodium* (Himatismenida) and the phylogeny of Amoebozoa. *Protist*, 156(2):215–224, 2005. ISSN 1434-4610. doi:10.1016/j.protis.2005.03.003. Page 36

Kudryavtsev, A., Pawlowski, J., and Hausmann, K. Description and phylogenetic relationships of *Spumochlamys perforata* n. sp and *Spumochlamys bryora* n. sp (Amoebozoa, Arcellinida). *Journal of Eukaryotic Microbiology*, 56(6):495–503, 2009. ISSN 1066-5234. doi:10.1111/j.1550-7408.2009.00430.x. Page 13, Page 19, Page 26

Kudryavtsev, A., Wylezich, C., and Pawlowski, J. *Ovalopodium desertum* n. sp and the phylogenetic relationships of Cochliopodiidae (Amoebozoa). *Protist*, 162(4):571–589, 2011. ISSN 1434-4610. doi:10.1016/j.protis.2011.04.002. Page 26, Page 36

Kudryavtsev, A. and Gladkikh, A. Two new species of *Ripella* (Amoebozoa, Vannellida) and unusual intragenomic variability in the SSU rRNA gene of this genus. *European journal of protistology*, 61:92–106, 2017. ISSN 0932-4739. doi:https://doi.org/10.1016/j.ejop.2017.09.003. Page 14

Kudryavtsev, A., Volkova, E., and Plotnikov, A. *Vannella samoroda* n. sp. (Amoebozoa)—First member of the genus from a continental saline habitat placed in a molecular tree. *European journal of protistology*, 71: 125634, 2019. ISSN 0932-4739. Page 4, Page 10

Lahr, D. J. G. and Katz, L. A. Reducing the impact of PCR-mediated recombination in molecular evolution and environmental studies using a new-generation high-fidelity DNA polymerase. *Biotechniques*, 47(4): 857–863, 2009. ISSN 0736-6205. doi:10.2144/000113219. Page 20

Lahr, D. J. G., Grant, J., Nguyen, T., Lin, J. H., and Katz, L. A. Comprehensive phylogenetic reconstruction of amoebozoa based on concatenated analyses of SSU-rDNA and actin genes. *PLoS One*, 6(7):e22780, 2011a. ISSN 1932-6203. Page 19, Page 25

Lahr, D. J. G., Nguyen, T. B., Barbero, E., and Katz, L. A. Evolution of the actin gene family in testate lobose amoebae (Arcellinida) is characterized by two distinct clades of paralogs and recent independent expansions. *Molecular Biology and Evolution*, 28(1):223–236, 2011b. ISSN 0737-4038. doi:10.1093/molbev/msq200. Page 14, Page 19, Page 25, Page 26, Page 72

Lahr, D. J. G., Bosak, T., Lara, E., and Mitchell, E. A. D. The Phanerozoic diversification of silica-cycling testate amoebae and its possible links to changes in terrestrial ecosystems. *PeerJ*, 3:e1234, 2015a. ISSN 2167-8359. Page 4, Page 5, Page 41, Page 51, Page 52

Lahr, D. J. G., Grant, J., Molestina, R., Katz, L. A., and Anderson, O. R. *Sapocribrum chincoteaguense* n. gen. n. sp.: a small, scale-bearing Amoebozoan with Flabellinid affinities. *Journal of Eukaryotic Microbiology*, 62(4): 444–453, 2015b. ISSN 1066-5234. doi:10.1111/jeu.12199. Cl9wo Times Cited:0 Cited References Count:45. Page 35

Lahr, D. J. G. and Lopes, S. G. B. C. Morphology, biometry, ecology and biogeography of five species of *Difflugia* Leclerc, 1815 (Arcellinida: Difflugiidae), from Tiete River, Brazil. *Acta Protozool*, 45:77–90, 2006. Page 11

Lahr, D. J. G., Parfrey, L. W., Mitchell, E. A. D., Katz, L. A., and Lara, E. The chastity of amoebae: re-evaluating evidence for sex in amoeboid organisms. *Proceedings of the Royal Society B: Biological Sciences*, 278(1715): 2081–2090, 2011c. ISSN 0962-8452. Page 2, Page 3, Page 11

Lahr, D. J. G., Grant, J. R., and Katz, L. A. Multigene phylogenetic reconstruction of the Tubulinea (Amoebozoa) corroborates four of the six major lineages, while additionally revealing that shell composition does not predict phylogeny in the Arcellinida. *Protist*, 164(3):323–339, 2013. ISSN 1434-4610. doi:http://dx.doi.org/10.1016/j.protis.2013.02.003. Page 25

Lahr, D. J. G., Kosakyan, A., Lara, E., Mitchell, E. A. D., Morais, L., Porfirio-Sousa, A. L., Ribeiro, G. M., Tice, A. K., Pánek, T., Kang, S., and Bornw, M. W. Phylogenomics and morphological reconstruction of Arcellinida testate amoebae highlight diversity of microbial eukaryotes in the Neoproterozoic. *Current Biology*, 29(6): 991–1001. e3, 2019. ISSN 0960-9822. Page 4, Page 6

Lamentowicz, M., Bragazza, L., Buttler, A., Jassey, V. E. J., and Mitchell, E. A. D. Seasonal patterns of testate amoeba diversity, community structure and species–environment relationships in four *Sphagnum*-dominated peatlands along a 1300 m altitudinal gradient in Switzerland. *Soil Biology and Biochemistry*, 67:1–11, 2013. ISSN 0038-0717. Page 63

Lamentowicz, M. and Mitchell, E. A. D. The ecology of testate amoebae (Protists) in *Sphagnum* in north-western Poland in relation to peatland ecology. *Microbial Ecology*, 50(1):48–63, 2005. Page 7

Lamentowicz, M., Van Der Knaap, W., Lamentowicz, u., Van Leeuwen, J. F. N., Mitchell, E. A. D., Goslar, T., and Kamenik, C. A near-annual palaeohydrological study based on testate amoebae from a sub-alpine mire: surface wetness and the role of climate during the instrumental period. *Journal of Quaternary Science*, 25 (2):190–202, 2010. ISSN 0267-8179. doi:https://doi.org/10.1002/jqs.1295. Page 4

Lamentowicz, M., Gałka, M., Obremska, M., Kühl, N., Lücke, A., and Jassey, V. E. J. Reconstructing climate change and ombrotrophic bog development during the last 4000 years in northern Poland using biotic proxies, stable isotopes and trait-based approach. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 418: 261–277, 2015. ISSN 0031-0182. doi:https://doi.org/10.1016/j.palaeo.2014.11.015. Page 7

Laminger, H. Die Testaceen (Protozoa, Rhizopoda) einiger Hochgebirgsgewässer von Mexiko, Costa Rica und Guatemala. *Int Rev Ges Hydrobiol*, 58(2):273–305, 1973. Page 9

Lara, E., Chatzinotas, A., and Simpson, A. G. B. *Andalucia* (n. Gen.) - the deepest branch within jakobids (Jakobida; Excavata), based on morphological and molecular study of a new flagellate from soil. *Journal of Eukaryotic Microbiology*, 53(2):112–120, 2006. ISSN 1066-5234. doi:10.1111/j.1550-7408.2005.00081.x. Page 32

Lara, E., Berney, C., Ekelund, F., Harms, H., and Chatzinotas, A. Molecular comparison of cultivable protozoa from a pristine and a polycyclic aromatic hydrocarbon polluted site. *Soil Biology & Biochemistry*, 39(1): 139–148, 2007. ISSN 0038-0717. doi:10.1016/j.soilbio.2006.06.017. Page 10, Page 31, Page 32

Lara, E., Heger, T. J., Ekelund, F., Lamentowicz, M., and Mitchell, E. A. D. Ribosomal RNA genes challenge the monophyly of the Hyalospheniidae (Amoebozoa : Arcellinida). *Protist*, 159(2):165–176, 2008. ISSN 1434-4610. doi:10.1016/j.protis.2007.09.003. Page 6, Page 11, Page 13, Page 26, Page 45, Page 51, Page 52, Page 72

Lara, E., Moreira, D., Vereshchaka, A., and Lopez-Garcia, P. Pan-oceanic distribution of new highly diverse clades of deep-sea diplonemids. *Environmental Microbiology*, 11(1):47–55, 2009. ISSN 1462-2912. doi:10.1111/j.1462-2920.2008.01737.x. Page 12, Page 31

Lara, E., Moreira, D., and Lopez-Garcia, P. The environmental clade LKM11 and Rozella form the deepest branching clade of fungi. *Protist*, 161(1):116–121, 2010. ISSN 1434-4610. doi:10.1016/j.protis.2009.06.005. Page 13, Page 31

Lara, E., Heger, T. J., Scheihing, R., and Mitchell, E. A. D. COI gene and ecological data suggest size-dependent high dispersal and low intra-specific diversity in free-living terrestrial protists (Euglyphida; Assulina). *Journal of Biogeography*, 38:640–650, 2011a. doi:10.1111/j.1365-2699.2010.02426.x. Page 45

Lara, E., Mitchell, E. A. D., Moreira, D., and Garcia, P. L. Highly diverse and seasonally dynamic protist community in a pristine peat bog. *Protist*, 162(1):14–32, 2011b. ISSN 1434-4610. doi:10.1016/j.protis.2010.05.003. Page 31, Page 74

Lauterborn, R. Protozoenstudien. III. Über eine Süßwasserart der Gattung *Multicilia* Cienkowsky (*M. lacustris* nov. spec.) und deren systematische Stellung. *Zeitschrift für wissenschaftliche Zoologie*, Bd.60:236–248, 1895. Page 2

Leadbeater, B. S. C. Irène Manton FRS 1904–1988. *Protist*, 155(1):127–140, 2004. ISSN 1434-4610. Page 72

Leidy, J. *Freshwater Rhizopods of North America*, volume Vol. 12. United States Geological Survey of the Territories, Washington, 1879. Page 4, Page 5, Page 45, Page 75, Page 77

Lejzerowicz, F., Pawlowski, J., Fraissinet-Tachet, L., and Marmeisse, R. Molecular evidence for widespread occurrence of Foraminifera in soils. *Environmental microbiology*, 12(9):2518–2526, 2010. ISSN 1462-2912. doi:10.1111/j.1462-2920.2010.02225.x. Page 13

Lentendu, G., Wubet, T., Chatzinotas, A., Wilhelm, C., Buscot, F., and Schlegel, M. Effects of long-term differential fertilization on eukaryotic microbial communities in an arable soil: a multiple barcoding approach. *Molecular Ecology*, 23(13):3341–3355, 2014. ISSN 0962-1083. doi:10.1111/mec.12819. Page 31

Lentendu, G., Mahé, F., Bass, D., Rueckert, S., Stoeck, T., and Dunthorn, M. Consistent patterns of high alpha and low beta diversity in tropical parasitic and free-living protists. *Molecular ecology*, 27(13):2846–2857, 2018. ISSN 0962-1083. Page 77

Lesaulnier, C., Papamichail, D., McCorkle, S., Ollivier, B., Skiena, S., Taghavi, S., Zak, D., and van der Lelie, D. Elevated atmospheric CO_2 affects soil microbial diversity associated with trembling aspen. *Environmental Microbiology*, 10(4):926–941, 2008. ISSN 1462-2912. doi:10.1111/j.1462-2920.2007.01512.x. Page 13, Page 31, Page 74

Leys, C., Klein, O., Dominicy, Y., and Ley, C. Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance. *Journal of Experimental Social Psychology*, 74:150–156, 2018. ISSN 0022-1031. Page 64

Lindley, L. A., Stephenson, S. L., and Spiegel, F. W. Protostelids and myxomycetes isolated from aquatic habitats. *Mycologia*, 99(4):504–509, 2007. ISSN 0027-5514. doi:DOI 10.3852/mycologia.99.4.504. Page 35

Loftus, B., Anderson, I., Davies, R., Alsmark, U. C. M., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R. P., Mann, B. J., Nozaki, T., Suh, B., Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M. A., Rabbinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guillén, N., Gilchrist, C., Stroup, S. E., Bhattacharya, S., Lohia, A., Foster, P. G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayed, N. M., Petri Jr, W. A., Clark, C. G., Embley, T. M., Barrell, B., Fraser, C. M., and Hall, N. The genome of the protist parasite *Entamoeba histolytica*. *Nature*, 433(7028):865–868, 2005. ISSN 0028-0836. Page 27

Lopez-Garcia, P., Rodriguez-Valera, F., Pedros-Alio, C., and Moreira, D. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, 409(6820):603–607, 2001. ISSN 0028-0836. doi:Doi 10.1038/35054537. Page 12, Page 31

Lukes, J., Flegontova, O., and Horak, A. Quick guide Diplonemids. *Current Biology*, 25(16):R702–R704, 2015. ISSN 0960-9822. Page 31

Luketa, S. Description of the family Padaungiellidae and morphological variability of *Padaungiella lageniformis* (Amoebozoides: Arcellinida) from the Vlasina Lake area, Serbia. *Arch. Biol. Sci*, 67:1331–1337, 2015. doi:DOI:10.2298/abs150312110l. Page 6, Page 76

Luketa, S. Morphological variability of *Nebela collaris* ss (Arcellinida: Hyalospheniidae) from Krečko Brdo Hill, East Herzegovina. *Biologia Serbica*, 39(2), 2017a. Page 11

Luketa, S. Morphological variability of *Padaungiella lageniformis* (Arcellinida: Padaungiellidae) from the central part of the Balkan Peninsula. *Protistology*, 11(1), 2017b. ISSN 1680-0826. doi:doi:10.21685/1680-0826-2017-11-1-2. Page 11, Page 76

Maciver, S. K. Ancestral eukaryotes reproduced asexually, facilitated by polyploidy: a hypothesis. *Bioessays*, 41(12):1900152, 2019. Page 4

Macumber, A. L., Blandenier, Q., Todorov, M., Duckert, C., Lara, E., Lahr, D. J. G., Mitchell, E. A. D., and Roe, H. M. Phylogenetic divergence within the Arcellinida (Amoebozoa) is congruent with test size and metabolism type. *European Journal of Protistology*, page 125645, 2019. ISSN 0932-4739. doi:https://doi.org/10.1016/j.ejop.2019.125645. Page 4, Page 14

Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. Swarm: robust and fast clustering method for amplicon-based studies. *Peerj*, 2, 2014. ISSN 2167-8359. doi:ARTN e593 10.7717/peerj.593. Page 33

Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ*, 3:e1420, 2015. ISSN 2167-8359. doi:doi: 10.7717/peerj.1420. Page 12, Page 64

Mahé, F., de Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., Singer, D., Mayor, J., Bunge, J., and Sernaker, S. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nature Ecology & Evolution*, 1(4):0091, 2017. ISSN 2397-334X. Page 2, Page 13, Page 76

Manton, I. and Leedale, G. F. Observations on the microanatomy of *Coccolithus pelagicus* and *Cricosphaera carterae*, with special reference to the origin and nature of coccoliths and scales. *Journal of the Marine Biological Association of the United Kingdom*, 49(1):1–16, 1969. ISSN 1469-7769. Page 72

Manton, I., Sutherland, J., and Leadbeater, B. S. C. Further observations on the fine structure of marine collared flagellates (Choanoflagellata) from arctic Canada and west Greenland: species of *Parvicorbicula* and *Pleurasiga*. *Canadian Journal of Botany*, 54(16):1932–1955, 1976. ISSN 0008-4026. Page 72

Manton, I. and Parke, M. Further observations on small green flagellates with special reference to possible relatives of *Chromulina pusilla* Butcher. *Journal of the Marine Biological Association of the United Kingdom*, 39(2):275–298, 1960. ISSN 1469-7769. Page 72

Margulis, L., Schwartz, K. V., and Gould, S. J. *Five kingdoms: an illustrated guide to the phyla of life on earth.* WH freeman New York, 1998. ISBN 0716730278. Page 3

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Øvreås, L., Reysenbach, A.-L., Smith, V. H., and Staley, J. T. Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, 4(2):102–112, 2006. ISSN 1740-1534. Page 75, Page 77

Massana, R., Castresana, J., Balague, V., Guillou, L., Romari, K., Groisillier, A., Valentin, K., and Pedros-Alio, C. Phylogenetic and ecological analysis of novel marine stramenopiles. *Applied and Environmental Microbiology*, 70(6):3528–3534, 2004. ISSN 0099-2240. doi:10.1128/Aem.70.6.3528-3534.2004. Page 31

Massana, R., del Campo, J., Sieracki, M. E., Audic, S., and Logares, R. Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles. *ISME Journal*, 8(4):854–866, 2014. ISSN 1751-7362. doi:10.1038/ismej.2013.204. Page 31

Massana, R., Terrado, R., Forn, I., Lovejoy, C., and Pedrós-Alió, C. Distribution and abundance of uncultured heterotrophic flagellates in the world oceans. *Environmental microbiology*, 8(9):1515–1522, 2006. ISSN 1462-2912. doi:https://doi.org/10.1111/j.1462-2920.2006.01042.x. Page 12

Mazei, Y. and Warren, A. A survey of the testate amoeba genus *Difflugia* Leclerc, 1815 based on specimens in the E. Penard and C. G. Ogden collections of the Natural History Museum, London. Part 2: species with shells that are pyriform or elongate. *Protistology*, 8(4):39, 2014. Page 25

Mazei, Y. A. and Bubnova, O. A. Species composition and structure of testate amoebae community in a *Sphagnum* bog at the initial stage of its formation. *Biology Bulletin*, 34(6):619–628, 2007. ISSN 1062-3590. Page 45

Medioli, F. S., Scott, D. B., and Abbott, B. H. A case study of protozoan intraclonal variability; taxonomic implications. *The Journal of Foraminiferal Research*, 17(1):28–47, 1987. ISSN 0096-1191. Page 6

Meisterfeld, R. Clusteranalytische Differenzierung der Testaceenzönosen (Rhizopoda, Testacea) in *Sphagnum*. *Archiv für Protistenkunde*, 121:270–307, 1979. Page 7

Meisterfeld, R. *Order Arcellinida Kent, 1880*, volume 2, pages 827–860. Society of protozoologists, Lawrence, Kansas, USA, second edition edition, 2002. Page 4, Page 5, Page 11, Page 41, Page 76

Meisterfeld, R. and Tan, L.-w. First records of testate amoebae (Protozoa: Rhizopoda) from Mount Buffalo National Park, Victoria: preliminary notes. *The Victorian Naturalist*, 1998. Page 9

Melton III, J. T., Wood, F. C., Branch, J., Singla, M., and Tekle, Y. I. Phylogenomics of Thecamoebida (Discosea, Amoebozoa) with the description of *Stratorugosa tubuloviscum* gen. nov. sp. nov., a freshwater amoeba with a perinuclear MTOC. *Protist*, 170(1):8–20, 2019. ISSN 1434-4610. Page 11

Miller, M. A., Pfeiffer, W., and Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Gateway Computing Environments Workshop (GCE), 2010*, pages 1–8. Ieee, 2010. ISBN 1424497523. Page 43, Page 53, Page 64

Mitchell, E. A. D., Buttler, A. J., Warner, B. G., and Gobat, J. M. Ecology of testate amoebae (Protozoa : Rhizopoda) in *Sphagnum* peatlands in the Jura mountains, Switzerland and France. *Ecoscience*, 6(4):565–576, 1999. Page 7

Mitchell, E. A. D., Charman, D. J., and Warner, B. G. Testate amoebae analysis in ecological and paleoecological studies of wetlands: past, present and future. *Biodiversity and Conservation*, 17:2115–2137, 2008. Page 4, Page 41

Mitchell, E. A. D. and Meisterfeld, R. Taxonomic confusion blurs the debate on cosmopolitanism versus local endemism of free-living protists. *Protist*, 156(3):263–267, 2005. doi:https://doi.org/10.1016/j.protis.2005.07.001. Page 8, Page 9, Page 10, Page 51, Page 52, Page 75, Page 76

Mitchell, E. A. D., Bragazza, L., and Gerdol, R. Testate amoebae (Protista) communities in *Hylocomium splendens* (Hedw.) BSG (Bryophyta): relationships with altitude, and moss elemental chemistry. *Protist*, 155(4):423–436, 2004. Page 63

Morais, L., Fairchild, T. R., Lahr, D. J. G., Rudnitzki, I. D., Schopf, J. W., Garcia, A. K., Kudryavtsev, A. B., and Romero, G. R. Carbonaceous and siliceous Neoproterozoic vase-shaped microfossils (Urucum Formation, Brazil) and the question of early protistan biomineralization. *Journal of Paleontology*, 91(3):393–406, 2017. ISSN 0022-3360. Page 4

Morard, R., Escarguel, G., Weiner, A. K. M., André, A., Douady, C. J., Wade, C. M., Darling, K. F., Ujiié, Y., Seears, H. A., Quillévéré, F., De Garidel-Thoron, T., De Vargas, C., and Kucera, M. Nomenclature for the nameless: a proposal for an integrative molecular taxonomy of cryptic diversity exemplified by planktonic foraminifera. *Systematic biology*, page syw031, 2016. ISSN 1063-5157. Page 13, Page 19

Mulot, M., Marcisz, K., Grandgirard, L., Lara, E., Kosakyan, A., Robroek, B. J. M., Lamentowicz, M., Payne, R. J., and Mitchell, E. A. D. Genetic determinism vs. phenotypic plasticity in protist morphology. *Journal of eukaryotic microbiology*, 64(6):729–739, 2017. ISSN 1550-7408. Page 6, Page 10, Page 41, Page 57, Page 76

Nassonova, E., Smirnov, A., Fahrni, J., and Pawlowski, J. Barcoding amoebae: comparison of SSU, ITS and COI genes as tools for molecular identification of naked lobose amoebae. *Protist*, 161(1):102–115, 2010. ISSN 1434-4610. doi:10.1016/j.protis.2009.07.003. Page 13, Page 14, Page 19, Page 26, Page 45, Page 72

Neff, R. J. Purification, axenic cultivation, and description of a soil amoeba, *Acanthamoeba* sp. *The Journal of Protozoology*, 4(3):176–182, 1957. ISSN 0022-3921. doi:https://doi.org/10.1111/j.1550-7408.1957.tb02505.x. Page 2

Nicholls, K. H. *Nebela kivuense* Gauthier-Lièvre et Thomas, 1961 (Amoebozoa, Arcellinida), missing for a half-century; found 11,500 km from "home". *Acta Protozoologica*, 54(4):283–288, 2015. ISSN 1689-0027. Page 45, Page 76

Nikolaev, S. I., Mitchell, E. A. D., Petrov, N. B., Berney, C., Fahrni, J., and Pawlowski, J. The testate lobose amoebae (order Arcellinida Kent, 1880) finally find their home within Amoebozoa. *Protist*, 156(2):191–202, 2005. ISSN 1434-4610. doi:10.1016/j.protis.2005.03.002. Page 11, Page 13, Page 19, Page 26, Page 41

Oda, E. *One Piece*, volume 113. 1999. Page 47

Odum, E. P. and Barrett, G. W. *Fundamentals of ecology*, volume 3. Saunders Philadelphia, 1971. Page 63

Ogawa, S., Yoshino, R., Angata, K., Iwamoto, M., Pi, M., Kuroe, K., Matsuo, K., Morio, T., Urushihara, H., Yanagisawa, K., and Tanaka, Y. The mitochondrial DNA of *Dictyostelium discoideum*: complete sequence, gene content and genome organization. *Molecular and General Genetics*, 263(3):514–519, 2000. ISSN 0026-8925. doi:Doi 10.1007/Pl00008685. Page 21, Page 25

Ogden, C. G. Comparative morphology of some pyriform species of *Difflugia* (Rhizopoda). *Archiv für Protistenkunde*, 122(1-2):143–153, 1979. ISSN 0003-9365. Page 11, Page 20

Ogden, C. G. Notes on some Difflugiidae from Norfolk (Rhizopodea, Protozoa). *Bull Brit Mus (Nat Hist) Zool*, 39:125–138, 1980. Page 11, Page 20

Ogden, C. G. Observations on the systematics of the genus *Difflugia* in Britain (Rhizopoda, Protozoa). *Bulletin of the Natural History Museum. Zoology series*, 44(1):1–73, 1983. Page 11

Ogden, C. G. The biology and ultrastructure of an agglutinate testate ameba *Difflugia geosphaira* sp. nov. (Protozoa, Rhizopoda). *Archiv Fur Protistenkunde*, 140(2-3):141, 1991. ISSN 0003-9365. Page 11

Ogden, C. G. and Meisterfeld, R. The taxonomy and systematics of some species of *Cucurbitella, Difflugia* and *Netzelia* (Protozoa: Rhizopoda); with an evaluation of diagnostic characters. *European journal of protistology*, 25(2):109–128, 1989. ISSN 0932-4739. Page 72

Ogden, C. G. and Meisterfeld, R. The biology and ultrastructure of the testate amoeba, *Difflugia lucida* Penard (Protozoa, Rhizopoda). *European journal of protistology*, 26(3-4):256–269, 1991. Page 10

Okamoto, N., Gawryluk, R. M. R., del Campo, J., Strassert, J. F. H., Lukeš, J., Richards, T. A., Worden, A. Z., Santoro, A. E., and Keeling, P. J. A revised taxonomy of diplonemids including the Eupelagonemidae n. fam. and a type species, *Eupelagonema oceanica* n. gen. & sp. *Journal of Eukaryotic Microbiology*, 66(3):519–524, 2019. ISSN 1066-5234. Page 73

Oksanen, J., Simpson, G., Solymos, P., Hannigan, G., Weedon, J., Szöcs, E., McGlinn, D., Legendre, P., Friendly, M., and Bolker, B. The Vegan Package, version v2. 3-5. *Community Ecology Package*, 2016. Page 65

Olive, L. S. The Protostelida: a new order of the Mycetozoa. *Mycologia*, 59(1):1–29, 1967. ISSN 00275514, 15572536. doi:10.2307/3756938. Page 36

Oliverio, A. M., Lahr, D. J. G., Nguyen, T., and Katz, L. A. Cryptic diversity within morphospecies of testate amoebae (Amoebozoa: Arcellinida) in New England bogs and fens. *Protist*, 165(2):196–207, 2014. ISSN 1434-4610. doi:10.1016/j.protis.2014.02.001. Page 19

Oliverio, A. M., Lahr, D. J. G., Grant, J., and Katz, L. A. Are microbes fundamentally different than macroorganisms? Convergence and a possible case for neutral phenotypic evolution in testate amoeba (Amoebozoa: Arcellinida). *Royal Society open science*, 2(12):150414, 2015. ISSN 2054-5703. Page 26

O'Malley, M. A. What did Darwin say about microbes, and how did microbiology respond? *Trends in microbiology*, 17(8):341–347, 2009. ISSN 0966-842X. Page 8, Page 51

Page, F. C. *An illustrated key to freshwater and soil amoebae.* Freshwater Biological Association, Ambleside; UK, 1976. Page 10

Patterson, R. T., Roe, H. M., and Swindles, G. T. Development of an Arcellacea (testate lobose amoebae) based transfer function for sedimentary phosphorus in lakes. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 348:32–44, 2012. ISSN 0031-0182. Page 4

Pavlović, L. M., Stojanovic, D. B., Mladenovic, E., Lakicevic, M., and Orlović, S. Potential elevation shift of the European beech stands (*Fagus sylvatica* L.) in Serbia. *Frontiers in plant science*, 10:849, 2019. ISSN 1664-462X. Page 68

Pawlowski, J. and Burki, F. Untangling the phylogeny of amoeboid protists. *Journal of Eukaryotic Microbiology*, 56(1):16–25, 2009. ISSN 1066-5234. doi:10.1111/j.1550-7408.2008.00379.x. Page 2, Page 36

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S. S., Čepička, I., Decelle, J., Dunthorn, M., Fiore-Donno, A. M., Gile, G. H., Holzmann, M., Jahn, R., Jirku, M., Keeling, P. J., Kostka, M., Kudryavtsev, A., Lara, E., Lukes, J., Mann, D. G., Mitchell, E. A. D., Nitsche, F., Romeralo, M., Saunders, G. W., Simpson, A. G. B., Smirnov, A. V., Spouge, J. L., Stern, R. F., Stoeck, T., Zimmermann, J., Schindel, D., and de Vargas, C. CBOL Protist Working Group: Barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *Plos Biology*, 10(11):e1001419, 2012. ISSN 1545-7885. doi:ARTN e1001419 10.1371/journal.pbio.1001419. Page 2, Page 13, Page 19, Page 24, Page 41, Page 72

Pawlowski, J., Lejzerowicz, F., Apotheloz-Perret-Gentil, L., Visco, J., and Esling, P. Protist metabarcoding and environmental biomonitoring: time for change. *European Journal of Protistology*, 55:12–25, 2016. ISSN 0932-4739. doi:doi: 10.1016/j.ejop.2016.02.003. Page 13

Pawlowski, J. The twilight of Sarcodina: a molecular perspective on the polyphyletic origin of amoeboid protists. *Protistology*, 5(4), 2008. ISSN 1680-0826. Page 2

Pawlowski, J. and Holzmann, M. A plea for DNA barcoding of foraminifera. *The Journal of Foraminiferal Research*, 44(1):62–67, 2014. ISSN 0096-1191. Page 19

Payne, R. J., Mitchell, E. A. D., Nguyen-Viet, H., and Gilbert, D. Can pollution bias peatland paleoclimate reconstruction? *Quaternary research*, 78(2):170–173, 2012. ISSN 0033-5894. doi:https://doi.org/10.1016/j.yqres.2012.05.004. Page 4

Pearson, W. R. *Flexible sequence similarity searching with the FASTA3 program package*, pages 185–219. Springer, 2000. Page 12

Penard, E. *Études sur les Rhizopodes d'eau douce.* Imprimerie Aubert-Schuchardt, 1890. Page 5, Page 24, Page 42, Page 43, Page 44, Page 45, Page 47, Page 48, Page 54, Page 74, Page 75

Penard, E. *Faune rhizopodique du bassin du Léman.* Kündig, Genève, 1902. Page 4, Page 5, Page 7, Page 8, Page 11, Page 44

Penard, E. *Contributions à l'étude des Rhizopodes du Léman.* éditeur non identifié, 1891. Page 5

Penard, E. *Les Rhizopodes de faune profonde dans le lac Léman.* W. Kündig, 1899. Page 5

Penard, E. Part VI - Sarcodina: Rhizopodes d'eau douce. Report, W. Heinemann, 1911. Page 7, Page 9, Page 56, Page 57

Penard, E. *Les infiniment petits dans leurs manifestations vitales.* Georg, 1938. Page 8, Page 77

Penard, E. *Protozoaires et psychologie.* éditeur non identifié, 1941. Page 77

Pérez-Juárez, H., Serrano-Vázquez, A., Kosakyan, A., Mitchell, E. A. D., Aguilar, V. M. R., Lahr, D. J. G., Moreno, M. M. H., Cuellar, H. M., Eguiarte, L. E., and Lara, E. *Quadrulella texcalense* sp. nov. from a Mexican desert: An unexpected new environment for hyalospheniid testate amoebae. *European journal of protistology*, 61: 253–264, 2017. ISSN 0932-4739. Page 5, Page 6, Page 41, Page 43

Perty, M. Mikroskopische Organismen der Alpen und der italienischen Schweiz. *Mittheilungen der Naturforschenden Gesellschaft in Bern*, 164(165):153–176, 1849. Page 25

Poche, F. The system of the protozoa. *Arch. Protistenkunde*, 30(3):125–321, 1913. Page 4

Pombert, J.-F., Janouškovec, J., Smirnov, A., Gray, M. W., James, E. R., and Keeling, P. J. The complete mitochondrial genome from an unidentified *Phalansterium* species. *Protist Genomics*, 1:25–32, 2013. Page 21

Porfírio-Sousa, A. L., Ribeiro, G. M., and Lahr, D. J. G. Morphometric and genetic analysis of *Arcella intermedia* and *Arcella intermedia laevis* (Amoebozoa, Arcellinida) illuminate phenotypic plasticity in microbial eukaryotes. *European journal of protistology*, 58:187–194, 2017. Page 10, Page 11, Page 24

Porter, S. M. and Knoll, A. H. Testate amoebae in the Neoproterozoic Era: evidence from vase-shaped microfossils in the Chuar Group, Grand Canyon. *Paleobiology*, 26(3):360–385, 2000. ISSN 0094-8373. Page 4

Prescott, D. M. and James, T. W. Culturing of *Amoeba proteus* on *Tetrahymena*. *Experimental cell research*, 8 (1):256–258, 1955. ISSN 0014-4827. Page 10

Prosser, J. I., Bohannan, B. J. M., Curtis, T. P., Ellis, R. J., Firestone, M. K., Freckleton, R. P., Green, J. L., Green, L. E., Killham, K., Lennon, J. J., Osborn, M. A., Solan, M., van der Gast, C. J., and Young, J. P. W. The role of ecological theory in microbial ecology. *Nature Reviews Microbiology*, 5(5):384–392, 2007. ISSN 1740-1534. doi:https://doi.org/10.1038/nrmicro1643. Page 75, Page 76

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., and Glöckner, F. O. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic acids research*, 35(21):7188–7196, 2007. ISSN 1362-4962. Page 12

Puillandre, N., Lambert, A., Brouillet, S., and ACHAZ. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular ecology*, 21(8):1864–1877, 2012. ISSN 0962-1083. Page 67

Qin, Y., Man, B., Kosakyan, A., Lara, E., Gu, Y., Wang, H., and Mitchell, E. A. D. *Nebela jiuhuensis* nov. sp. (Amoebozoa; Arcellinida; Hyalospheniidae): a new member of the *Nebela saccifera-equicalceus-ansata* group described from *Sphagnum* Peatlands in South-Central China. . *Journal of eukaryotic microbiology*, 63:558–566, 2016. Page 6, Page 43

Qvarnstrom, Y., Da Silva, A. J., Schuster, F. L., Gelman, B. B., and Visvesvara, G. S. Molecular confirmation of *Sappinia pedata* as a causative agent of amoebic encephalitis. *The Journal of infectious diseases*, 199(8): 1139–1142, 2009. ISSN 1537-6613. Page 2

Raikov, I. B. *The protozoan nucleus - morphology and evolution.* Springer-Verlag, 1982. ISBN 321181678X. Page 4

Raper, K. B. *The Dictyostelids.* Princeton University Press, Princeton, 1984. Page 3

Richards, T. A., Vepritskiy, A. A., Gouliamova, D. E., and Nierzwicki-Bauer, S. A. The molecular diversity of freshwater picoeukaryotes from an oligotrophic lake reveals diverse, distinctive and globally dispersed lineages. *Environmental Microbiology*, 7(9):1413–1425, 2005. ISSN 1462-2912. doi:10.1111/j.1462-2920.2005.00828.x. Page 31, Page 74

Riedman, L. A., Porter, S. M., and Calver, C. R. Vase-shaped microfossil biostratigraphy with new data from Tasmania, Svalbard, Greenland, Sweden and the Yukon. *Precambrian Research*, 319:19–36, 2018. ISSN 0301-9268. Page 4

Rippon, J. W. *Medical mycology : the pathogenic fungi and the pathogenic actinomycetes.* Saunders, Philadelphia, 2nd edition, 1982. ISBN 0721675867. Page 36

Rønn, R., Ekelund, F., and Christensen, S. Optimizing soil extract and broth media for MPN enumeration of naked amoebae and heterotrophic flagellates in soil. *Pedobiologia*, 39:10–19, 1995. ISSN 1873-1511. Page 10

Robba, L., Russell, S. J., Barker, G. L., and Brodie, J. Assessing the use of the mitochondrial cox1 marker for use in DNA barcoding of red algae (Rhodophyta). *American journal of botany*, 93(8):1101–1108, 2006. ISSN 0002-9122. doi:https://doi.org/10.3732/ajb.93.8.1101. Page 14

Robideau, G. P., de Cock, A. W. A. M., Coffey, M. D., Voglmayr, H., Brouwer, H., Bala, K., Chitty, D. W., Désaulniers, N., Eggertson, Q. A., Gachon, C. M. M., Hu, C. H., Küpper, F. C., Rintoul, T. L., Sarhan, E., Verstappen, E. C., Zhang, Y., Bonants, P. J., Ristaino, J. B., and Lévesque, C. A. DNA barcoding of oomycetes with cytochrome c oxidase subunit I and internal transcribed spacer. *Molecular ecology resources*, 11(6): 1002–1011, 2011. ISSN 1755-098X. doi:doi: 10.1111/j.1755-0998.2011.03041.x. Page 13

Rogerson, A. *Parvamoeba Rugata* n.g., n.sp., (Gymnamoebia, thecamoebidae): an exceptionally small marine naked amoeba. *European Journal of Protistology*, 29(4):446–452, 1993. ISSN 0932-4739. Page 35

Ronquist, F. and Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12):1572–1574, 2003. ISSN 1460-2059. Page 43

Rydin, H., Jeglum, J. K., and Bennett, K. D. *The biology of peatlands, 2e*. Oxford university press, 2013. ISBN 0199602999. Page 7

Santos, S. S., Schöler, A., Nielsen, T. K., Hansen, L., Schloter, M., and Winding, A. Land use as a driver for protist community structure in soils under agricultural use across Europe. *Science of The Total Environment*, page 137228, 2020. Page 2, Page 76

Schiaffino, M. R., Lara, E., Fernández, L. D., Balagué, V., Singer, D., Seppey, C. C. W., Massana, R., and Izaguirre, I. Microbial eukaryote communities exhibit robust biogeographical patterns along a gradient of Patagonian and Antarctic lakes. *Environmental microbiology*, 18(12):5249–5264, 2016. ISSN 1462-2912. Page 63

Schönborn, W. Zur Ökologie der sphagnicolen, bryokolen un terrikolen Testaceen. *Limnologica*, 1:231–254, 1962. Page 7

Schönborn, W. Comparative-studies on the production biology of protozoan communities in fresh-water and soil ecosystems. *Archiv Fur Protistenkunde*, 141(3):187–214, 1992. ISSN 0003-9365. Page 4

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W., and Consortium, F. B. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for fungi. *Proceedings of the National Academy of Sciences*, 109(16):6241–6246, 2012. ISSN 0027-8424. doi:10.1073/pnas.1117018109. Page 13

Schuler, G. A. and Brown, M. W. Description of *Armaparvus languidus* n. gen. n. sp. confirms ultrastructural unity of Cutosea (Amoebozoa, Evosea). *Journal of Eukaryotic Microbiology*, 66(1):158–166, 2019. ISSN 1066-5234. Page 11, Page 12

Schultze, F. E. Rhizopodenstudien VI. *Arch. Mikrosk. Anat.*, 13:9–30, 1877. Page 4, Page 5, Page 6, Page 46

Seppey, C. V. W., Fournier, B., Szelecz, I., Singer, D., Mitchell, E. A. D., and Lara, E. Response of forest soil euglyphid testate amoebae (Rhizaria: Cercozoa) to pig cadavers assessed by high-throughput sequencing. *International Journal of Legal Medicine*, 130(2):551–562, 2016. ISSN 0937-9827. doi:10.1007/s00414-015-1149-7. Page 32, Page 41

Seppey, C. V. W., Singer, D., Dumack, K., Fournier, B., Belbahri, L., Mitchell, E. A. D., and Lara, E. Distribution patterns of soil microbial eukaryotes suggests widespread algivory by phagotrophic protists as an alternative pathway for nutrient cycling. *Soil Biology and Biochemistry*, 112:68–76, 2017. ISSN 0038-0717. doi:https://doi.org/10.1016/j.soilbio.2017.05.002. Page 2, Page 13

Seppey, C. V. W., Broennimann, O., Buri, A., Yashiro, E., Pinto-Figueroa, E., Singer, D., Blandenier, Q., Mitchell, E. A. D., Hirzel, H. N., Guisan, A., and Lara, E. Soil protist diversity in the Swiss western Alps is better predicted by topo-climatic than by edaphic variables. *Journal of Biogeography*, page 1–13, 2019. doi:https://doi.org/10.1111/jbi.13755. Page 8, Page 64, Page 73

Shadwick, L. L., Spiegel, F. W., Shadwick, J. D. L., Brown, M. W., and Silberman, J. D. Eumycetozoa = Amoebozoa?: SSUrDNA phylogeny of protosteloid slime molds and its significance for the amoebozoan supergroup. *Plos One*, 4(8), 2009. ISSN 1932-6203. doi:ARTN e6754 10.1371/journal.pone.0006754. Page 3, Page 31, Page 36

Shaw, A. J., Golinski, G. K., Clark, E. G., Shaw, B., Stenøien, H. K., and Flatberg, K. I. Intercontinental genetic structure in the amphi-Pacific peatmoss *Sphagnum miyabeanum* (Bryophyta: Sphagnaceae). *Biological journal of the Linnean society*, 111(1):17–37, 2013. ISSN 0024-4066. Page 76

Shchepin, O. N., Schnittler, M., Erastova, D. A., Prikhodko, I. S., Dahl, M. B., Azarov, D. V., Chernyaeva, E. N., and Novozhilov, Y. K. Community of dark-spored myxomycetes in ground litter and soil of taiga forest (Nizhne-Svirskiy Reserve, Russia) revealed by DNA metabarcoding. *Fungal ecology*, 39:80–93, 2019. ISSN 1754-5048. doi:https://doi.org/10.1016/j.funeco.2018.11.006. Page 13

Shen, C., Liang, W., Shi, Y., Lin, X., Zhang, H., Wu, X., Xie, G., Chain, P., Grogan, P., and Chu, H. Contrasting elevational diversity patterns between eukaryotic soil microbes and plants. *Ecology*, 95(11):3190–3202, 2014. ISSN 1939-9170. Page 63

Simpson, A. G. B. and Roger, A. J. The real 'kingdoms' of eukaryotes. *Current biology*, 14(17):R693–R696, 2004. ISSN 0960-9822. Page 2

Singer, D., Kosakyan, A., Pillonel, A., Mitchell, E. A. D., and Lara, E. Eight species in the *Nebela collaris* complex: *Nebela gimlii* (Arcellinida, Hyalospheniidae), a new species described from a Swiss raised bog. *European Journal of Protistology*, 51(1):79–85, 2015. ISSN 0932-4739. doi:10.1016/j.ejop.2014.11.004. Page 6, Page 14, Page 19, Page 41, Page 43, Page 45, Page 51, Page 63, Page 64, Page 65, Page 66, Page 72, Page 74, Page 76

Singer, D., Lara, E., Steciow, M. M., Seppey, C. V. W., Paredes, N., Pillonel, A., Oszako, T., and Belbahri, L. High-throughput sequencing reveals diverse oomycete communities in oligotrophic peat bog micro-habitat. *Fungal Ecology*, 23:42–47, 2016. ISSN 1754-5048. Page 20

Singer, D., Kosakyan, A., Seppey, C. V. W., Pillonel, A., Fernández, L. D., Fontaneto, D., Mitchell, E. A. D., and Lara, E. Environmental filtering and phylogenetic clustering correlate with the distribution patterns of cryptic protist species. *Ecology*, 99(4):904–914, 2018. ISSN 0012-9658. Page 8, Page 14, Page 41, Page 45, Page 51, Page 63, Page 64, Page 65, Page 66, Page 67, Page 68, Page 73, Page 75, Page 77

Singer, D., Mitchell, E. A. D., Payne, R. J., Blandenier, Q., Duckert, C., Fernández, L. D., Fournier, B., Hernández, C. E., Granath, G., Rydin, H., Bragazza, L., Koronatova, N. G., Goia, I., Harris, L. I., Kajukało, K., Kosakyan, A., Lamentowicz, M., Kosykh, N. P., Vellak, K., and Lara, E. Dispersal limitations and historical factors determine the biogeography of specialized terrestrial protists. *Molecular ecology*, 28:3089– 3100, 2019. ISSN 0962-1083. doi:https://doi.org/10.1111/mec.15117. Page 10, Page 51, Page 57, Page 58, Page 77

Smirnov, A. Amoebas, lobose. *Encyclopedia of Microbiology. Elsevier, Oxford*, 558577, 2008. Page 2, Page 11

Smirnov, A. and Goodkov, A. Ultrastructure and geographic distribution of the genus *Paradermamoeba* (Gymnamoebia, Thecamoebidae). *European Journal of Protistology*, 40(2):113–118, 2004. ISSN 0932-4739. doi:10.1016/j.ejop.2003.12.001. Page 11, Page 36

Smirnov, A., Nassonova, E., Berney, C., Fahrni, J., Bolivar, I., and Pawlowski, J. Molecular phylogeny and classification of the lobose amoebae. *Protist*, 156(2):129–142, 2005. ISSN 1434-4610. doi:10.1016/j.protis.2005.06.002. Page 19

Smirnov, A. V., Nassonova, E. S., Chao, E., and Cavalier-Smith, T. Phylogeny, evolution, and taxonomy of vannellid amoebae. *Protist*, 158(3):295–324, 2007. ISSN 1434-4610. doi:10.1016/j.protis.2007.04.004. Page 36

Smirnov, A. V., Bedjagina, O. M., and Goodkov, A. V. *Dermamoeba algensis* n. sp (Amoebozoa, Dermamoebidae) - an algivorous lobose amoeba with complex cell coat and unusual feeding mode. *European Journal of Protistology*, 47(2):67–78, 2011a. ISSN 0932-4739. doi:10.1016/j.ejop.2010.12.002. Page 2, Page 11, Page 36

Smirnov, A. V., Chao, E., Nassonova, E. S., and Cavalier-Smith, T. A revised classification of naked lobose amoebae (Amoebozoa: Lobosa). *Protist*, 162(4):545–570, 2011b. ISSN 1434-4610. doi:10.1016/j.protis.2011.04.004. Page 32, Page 36

Smirnov, A., Nassonova, E., Geisen, S., Bonkowski, M., Kudryavtsev, A., Berney, C., Glotova, A., Bondarenko, N., Dyková, I., Mrva, M., Fahrni, J., and Pawlowski, J. Phylogeny and systematics of leptomyxid amoebae (Amoebozoa, Tubulinea, Leptomyxida). *Protist*, 168(2):220–252, 2017. ISSN 1434-4610. Page 12

Smirnov, A. V. and Goodkov, A. V. An illustrated list of basic morphotypes of Gymnamoebia (Rhizopoda, Lobosea). *Protistology*, 1(1):20–29, 1999. ISSN 1680-0826. Page 2

Smith, H., Bobrov, A., and Lara, E. Diversity and biogeography of testate amoebae. *Biodiversity and Conservation*, 17(2):1822–1831, 2008. Page 6, Page 9, Page 41, Page 51, Page 76

Smith, H. G. and Wilkinson, D. M. Not all free-living microorganisms have cosmopolitan distributions – the case of *Nebela* (*Apodera*) *vas* Certes (Protozoa: Amoebozoa: Arcellinida). *Journal of Biogeography*, 34: 1822–1831, 2007. Page 6, Page 9, Page 41, Page 51, Page 57

Soininen, J. Macroecology of unicellular organisms–patterns and processes. *Environmental microbiology reports*, 4(1):10–22, 2012. ISSN 1758-2229. doi:https://doi.org/10.1111/j.1758-2229.2011.00308.x. Page 63, Page 75, Page 76, Page 77

Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313, 2014. ISSN 1367-4803. Page 43, Page 53, Page 64

Stamatakis, A., Hoover, P., and Rougemont, J. A rapid bootstrap algorithm for the RAxML web servers. *Syst. Biol.*, 57:758–771, 2008. Page 32, Page 35

Stephenson, S. L., Fiore-Donno, A. M., and Schnittler, M. Myxomycetes in soil. *Soil Biology and Biochemistry*, 43(11):2237–2242, 2011. ISSN 0038-0717. doi:https://doi.org/10.1016/j.soilbio.2011.07.007. Page 13, Page 72

Stoneking, M. *Mitochondrial Genome: Evolution.* John Wiley & Sons, Ltd, 2013. ISBN 9780470015902. doi:10.1002/9780470015902.a0005074.pub3. Page 14, Page 19

Sudzuki, M. Zur biologischen Analyse der mikroskopischen Süsswassertierwelt geringster. *Zool. Mag.*, 73: 165–174, 1964. Page 44

Swindles, G. T., Baird, A. J., Kilbride, E., Low, R., and Lopez, O. Testing the relationship between testate amoeba community composition and environmental variables in a coastal tropical peatland. *Ecological Indicators*, 91:636–644, 2018. ISSN 1470-160X. doi:https://doi.org/10.1016/j.ecolind.2018.03.021. Page 9

Szelecz, I., Fournier, B., Seppey, C., Amendt, J., and Mitchell, E. Can soil testate amoebae be used for estimating the time since death? A field experiment in a deciduous forest. *Forensic Science International*, 236:90–98, 2014. ISSN 0379-0738. doi:10.1016/j.forsciint.2013.12.030. Page 32, Page 35

Takano, H., Abe, T., Sakurai, R., Moriyama, Y., Miyazawa, Y., Nozaki, H., Kawano, S., Sasaki, N., and Kuroiwa, T. The complete DNA sequence of the mitochondrial genome of *Physarum polycephalum*. *Molecular and General Genetics MGG*, 264(5):539–545, 2001. ISSN 1432-1874. doi:10.1007/s004380000357. Page 21

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12):2725–2729, 2013. ISSN 0737-4038. doi:10.1093/molbev/mst197. Page 21

Taranek, K. J. Beiträge zur Kenntniss der Süsswasser-Rhizopoden Böhmens. *Sitzungsberichte der Königlichen Böhmischen Gesellschaft der Wissenschaften in Prag*, 1881:220–235, 1881. Page 44, Page 47, Page 75

Taranek, K. J. Monographie der Böhmen's: Ein Beitrag zut Kenntniss der Süsswasser-Monothalamien. *Abh. Königl. Böhm. Ges. Wiss.*, 2:1–56, 1882. Page 5, Page 44, Page 47, Page 48

Tarnawski, S. E. and Lara, E. From environmental sequences to morphology: observation and characterisation of a paulinellid testate amoeba (*Micropyxidiella edaphonis* gen. nov. sp. nov. Euglyphida, Paulinellidae) from soil using fluorescent in situ hybridization. *Protist*, 166(2):264–270, 2015. ISSN 1434-4610. doi:10.1016/j.protis.2015.04.002. Page 31, Page 74

Team, R. D. C. *R: A Language and Environment for Statistical Computing.* Vienna, Austria, 2011. Page 33

Tekle, Y. I., Grant, J., Anderson, O. R., Nerad, T. A., Cole, J. C., Patterson, D. J., and Katz, L. A. Phylogenetic placement of diverse amoebae inferred from multigene analyses and assessment of clade stability within 'Amoebozoa' upon removal of varying rate classes of SSU-rDNA. *Molecular Phylogenetics and Evolution*, 47(1):339–352, 2008. ISSN 1055-7903. doi:10.1016/j.ympev.2007.11.015. Page 19

Tekle, Y. I., Anderson, O. R., Katz, L. A., Maurer-Alcala, X. X., Romero, M. A. C., and Molestina, R. Phylogenomics of 'Discosea': a new molecular phylogenetic perspective on Amoebozoa with flat body forms. *Molecular Phylogenetics and Evolution*, 99:144–154, 2016. ISSN 1055-7903. doi:10.1016/j.ympev.2016.03.029. Page 19, Page 36

Tekle, Y. I. DNA barcoding in Amoebozoa and challenges: the example of *Cochliopodium*. *Protist*, 165(4): 473–484, 2014. ISSN 1434-4610. doi:https://doi.org/10.1016/j.protis.2014.05.002. Page 14

Tice, A. K., Shadwick, L. L., Fiore-Donno, A. M., Geisen, S., Kang, S., Schuler, G. A., Spiegel, F. W., Wilkinson, K. A., Bonkowski, M., Dumack, K., Lahr, D. J. G., Voelcker, E., Clauß, S., Zhang, J., and Brown, M. W. Expansion of the molecular and morphological diversity of Acanthamoebidae (Centramoebida, Amoebozoa) and identification of a novel life cycle type within the group. *Biology direct*, 11(1):69, 2016. ISSN 1745-6150. Page 4

Todorov, M. Observation on the soil and moss testate amoebae (Protozoa: Rhizopoda) from Pirin Mountain (Bulgaria). *Acta zool. bulg*, 50(2/3):19–29, 1998. Page 63

Todorov, M. Morphology, biometry and ecology of *Nebela bigibbosa* Penard, 1890 (Protozoa: Rhizopoda). *Acta Protozoologica*, 41:239–244, 2002. Page 5, Page 7, Page 74

Todorov, M. and Bankov, N. An atlas of *Sphagnum*-dwelling testate amoebae in Bulgaria. *Advanced Books*, 1: e38685, 2019. Page 5, Page 54, Page 55

Todorov, M. and Golemansky, V. Morphological variability of *Difflugia urceolata* Carter, 1864 (Testacealobosia: Difflugiidae) and taxonomical status of its varieties *D. urceolata* var. *olla* Leidy, 1879, and *D. urceolata* var. *sphaerica* Playfair, 1917. *Acta Zool. Bulg.*, 59:3–10, 2007. Page 51

Todorov, M. and Golemansky, V. Soil testate amoebae (Amoebozoa, Rhizaria) of a tropical rainforest in Madagascar. *Acta Zoologica Bulgarica*, 66(4):469–476, 2014. ISSN 0324-0770. Page 9

Todorov, M., Golemansky, V., and Meisterfeld, R. Is *Difflugia nebeloides* (Amoebozoa: Arcellinida) really a *Difflugia*? Re-description and new combination. *Acta Zool. Bulg*, 62:13–20, 2010. Page 54

Tolonen, K., Warner, B. G., and Vasander, H. Ecology of testaceans (Protozoa, Rhizopoda) in mires in Southern Finland .1. Autecology. *Archiv für Protistenkunde*, 142(3-4):119–138, 1992. Page 7

Tolonen, K., Warner, B. G., and Vasander, H. Ecology of testaceans (Protozoa, Rhizopoda) in mires in Southern Finland .2. Multivariate-Analysis. *Archiv für Protistenkunde*, 144(1):97–112, 1994. Page 7

Traphagen, S. J., Dimarco, M. J., and Silliker, M. E. RNA editing of 10 *Didymium iridis* mitochondrial genes and comparison with the homologous genes in *Physarum polycephalum*. *RNA*, 16(4):828–838, 2010. ISSN 1355-8382. Page 22

Tsyganov, A. N., Babeshko, K. V., and Mazei, Y. A. A guide to testate amoebae with the keys to genera: monograph. *. ISBN 978-5-906913-19-7.*, 2016. Page 4, Page 5, Page 41

Tsyganov, A. N., Milbau, A., and Beyens, L. Environmental factors influencing soil testate amoebae in herbaceous and shrubby vegetation along an altitudinal gradient in subarctic tundra (Abisko, Sweden). *European journal of protistology*, 49(2):238–248, 2013. ISSN 0932-4739. Page 63

Tyml, T. and Dyková, I. *Sappinia* sp. (Amoebozoa: Thecamoebida) and *Rosculus* sp. (SAR: Cercozoa) isolated from King Penguin guano collected in the Subantarctic (South Georgia, Salisbury Plain) and their coexistence in culture. *Journal of Eukaryotic Microbiology*, 65(4):544–555, 2018. ISSN 1066-5234. Page 11

Udalov, I. A. *Pseudoparamoeba microlepis* n. sp., *Korotnevella fousta* n. sp. (Amoebozoa, Dactylopodida), with notes on the evolution of scales among dactylopodid amoebae. *European journal of protistology*, 54:33–46, 2016. ISSN 0932-4739. Page 11

Valentini, A., Pompanon, F., and Taberlet, P. DNA barcoding for ecologists. *Trends in ecology & evolution*, 24 (2):110–117, 2009. ISSN 0169-5347. Page 12

Väliranta, M., Blundell, A., Charman, D. J., Karofeld, E., Korhola, A., Sillasoo, U., and Tuittila, E. Reconstructing peatland water tables using transfer functions for plant macrofossils and testate amoebae: a methodological comparison. *Quaternary International*, 268:34–43, 2012. ISSN 1040-6182. Page 41

van Hannen, E. J., Zwart, G., van Agterveld, M. P., Gons, H. J., Ebert, J., and Laanbroek, H. J. Changes in bacterial and eukaryotic community structure after mass lysis of filamentous cyanobacteria associated with viruses. *Applied and Environmental Microbiology*, 65(2):795–801, 1999. ISSN 0099-2240. Page 12, Page 31, Page 74, Page 76

Van Leewenhoeck, A. Observations, communicated to the publisher by Mr. Antony van Leewenhoeck, in a Dutch Letter of the 9th of Octob. 1676. Here English'd: concerning little animals by him observed in rain-well-sea and snow water; as also in water wherein pepper had lain infused. *Philosophical Transactions (1665-1678)*, 12:821–831, 1676. ISSN 0370-2316. Page 2

Van Oye, P. Rhizopodes du district sub-alpin de la Belgique. *Arch. Naturgesc., Abt. B*, 2:538–573, 1933. Page 44, Page 48

Van Oye, P. On the Thecamoeban fauna of New Zealand with description of four new species and biogeografical discussion. *Hydrobiologia*, 8:16–37, 1956. Page 5, Page 8, Page 9, Page 45, Page 56

Van Oye, P. Rhizopodes de Java. *Bijdragen tot de Dierkunde*, 28(1):327–352, 1949. ISSN 0067-8546. Page 9

Veech, J. A. A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography*, 22 (2):252–260, 2013. ISSN 1466-822X. Page 65, Page 67

Vejdovský, F. *Thierische Organismen der Brunnenwasser von Prag*. Selbstverlag, Prag, 1882. Page 44, Page 48

Visvesvara, G. S., Moura, H., and Schuster, F. L. Pathogenic and opportunistic free-living amoebae: *Acanthamoeba* spp., *Balamuthia mandrillaris*, *Naegleria fowleri*, and *Sappinia diploidea*. *FEMS Immunology & Medical Microbiology*, 50(1):1–26, 2007. ISSN 0928-8244. Page 2

Volkova, E. and Kudryavtsev, A. Description of *Neoparamoeba longipodia* n. sp. and a new strain of *Neoparamoeba aestuarina* (Page, 1970)(Amoebozoa, Dactylopodida) from deep-sea habitats. *European journal of protistology*, 61:107–121, 2017. ISSN 0932-4739. Page 2

Volkova, E. and Smirnov, A. Regeneration of test in testate amoebae of the genus *Arcella* (Tubulinea, Arcellinida). *European journal of protistology*, 55:128–140, 2016. ISSN 0932-4739. Page 11

Von Wintzingerode, F., Göbel, U. B., and Stackebrandt, E. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *Fems Microbiol. Rev*, 21:213–229, 1997. Page 13

Voss, C., Fiore-Donno, A. M., Guerreiro, M. A., Peršoh, D., and Bonkowski, M. Metatranscriptomics reveals unsuspected protistan diversity in leaf litter across temperate beech forests, with Amoebozoa the dominating lineage. *FEMS microbiology ecology*, 95(10):fiz142, 2019. ISSN 0168-6496. doi:https://doi.org/10.1093/femsec/fiz142. Page 2, Page 13, Page 73

Vucetich, M. C. Comentarios criticos sobre *Argynnia* Jung, 1942 (Rhizopoda, Testacea). *Neotropica*, 20:126–128, 1974. Page 75

Wailes, G. H. Freshwater Rhizopoda from North and South America. *Zoological Journal of the Linnean Society*, 32(216):201–218, 1913. ISSN 0024-4082. Page 9, Page 56

Wallace, A. R. *The geographical distribution of animals: with a study of the relations of living and extinct faunas as elucidating the past changes of the earth's surface*, volume 1. Cambridge University Press, 1876. Page 75

Wanner, M. A review on the variability of testate amoebae: Methodological approaches, environmental influences and taxonomical implications. *Acta Protozoologica*, 38(1):15–29, 1999. Page 6, Page 41

Wanner, M. and Meisterfeld, R. Effects of some environmental factors on the shell morphology of testate amoebae (Rhizopoda, Protozoa). *European Journal of Protistology*, 30(2):191–195, 1994. Page 6, Page 10, Page 41

Warner, B. G. *Testate amoebae (Protozoa)*, volume Reprint Series 5, pages 65–74. Geoscience Canada, St. John's, Newfoundland, 1990. Page 4

Warner, B. G. and Chmielewski, J. G. Testate amoebae (Protozoa) as indicators of drainage in a forested mire, northern Ontario, Canada. *Archiv für Protistenkunde*, 141(3):179–183, 1992. Page 7

Wood, F. C., Heidari, A., and Tekle, Y. I. Genetic evidence for sexuality in *Cochliopodium* (Amoebozoa). *Journal of Heredity*, 108(7):769–779, 2017. ISSN 0022-1503. doi:doi:10.1093/jhered/esx078. Page 3

Yeates, G. W. and Foissner, W. Testate amoebae as predators of nematodes. *Biology and Fertility of Soils*, 20: 1–7, 1995. Page 5, Page 51, Page 52, Page 56

Zancan, S., Trevisan, R., and Paoletti, M. G. Soil algae composition under different agro-ecosystems in North-Eastern Italy. *Agriculture, ecosystems & environment*, 112(1):1–12, 2006. ISSN 0167-8809. doi:https://doi.org/10.1016/j.agee.2005.06.018. Page 2

Zapata, J. and Fernandez, L. Morphology and morphometry of *Apodera vas* (Certes, 1889)(Protozoa: Testacea) from two peatlands in Southern Chile. *Acta Protozoologica*, 47(4):389, 2008. ISSN 0065-1583. Page 7, Page 51, Page 52, Page 56, Page 57

Zhao, F., Filker, S., Xu, K., Li, J., Zhou, T., and Huang, P. Effects of intragenomic polymorphism in the SSU rRNA gene on estimating marine microeukaryotic diversity: A test for ciliates using single-cell high-throughput DNA sequencing. *Limnology and Oceanography: Methods*, 2019. ISSN 1541-5856. doi:https://doi.org/10.1002/lom3.10330. Page 13

Appendix A

Molecular and morphological tools to study genus *Difflugia*

Phylogenetic divergence within the Arcellinida (Amoebozoa) is congruent with test size and metabolism type

Andrew L. Macumber[1], Quentin Blandenier[2, 3], Milcho Todorov[4], Clément Duckert[2], Enrique Lara[3], Daniel J.G. Lahr[5], Edward A.D. Mitchell[2, 6] and Helen M. Roe[1]

[1]School of Natural and Built Environment, Queen's University Belfast, Belfast, BT7 1NN, United Kingdom
[2]Laboratory of Soil Biodiversity, University of Neuchâtel, Neuchâtel, 2000, Switzerland
[3]Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, Sofia, 1113, Bulgaria
[4]Real Jardin Botánico de Madrid, Spanish National Research Council, Madrid, E-28014, Spain
[5]Department of Zoology, Institute of Biosciences, University of São Paulo, São Paulo, 05508-090, Brazil
[6]Botanical Garden of Neuchâtel, Neuchâtel, 2000, Switzerland

European Journal of Protistology, 2020, Vol 72

DOI: 10.1016/j.ejop.2019.125645

Abstract. Arcellinida (lobose testate amoebae) are abundant and diverse in many ecosystems, especially in moist to aquatic environments. Molecular phylogeny has shown that overall test morphology (e.g., spherical or elongate) is generally conserved in Arcellinida lineages, but the taxonomic value of other traits (e.g., size, ornamentation, mixotrophy/heterotrophy metabolism type) has not been systematically evaluated. Morphological and physiological traits that correspond to genetic differences likely represent adaptive traits of ecological significance. We combined high-resolution phylogenetics (NAD9-NAD7 genes) and advanced morphometrics to assess the phylogenetic signal of morphological traits of a group of elongate *Difflugia* species (Arcellinida). The phylogenetic analyses revealed two clades which could be reliably separated by test size and the presence/absence of mixotrophy. Differences in test size may reflect trophic level, with smaller organisms occupying lower trophic levels. In addition to having larger tests, elongate mixotrophic *Difflugia* are characterised by wide, flat bases and an inflation of the lower two thirds of their test. These morphological traits may provide additional volume for endosymbionts and/or increased surface area to aid light transmission. Our results showcase greater diversity within the elongate *Difflugia* and highlight morphological traits of ecological and evolutionary significance.

Keywords. *Difflugia*; Centroid size; Geometric morphometrics; Molecular barcoding; Morphometry; NADH gene complex

Introduction

Arcellinida Kent, 1880 are an order of testate (shelled) Amoebozoa, characterized by broad and thick (lobose) pseudopods used for feeding and locomotion. Like other testate amoebae they enclose their cell body within a test, the morphology and composition of which can be used to identify species (Meisterfeld, 2002). Several features make Arcellinida valuable ecological indicators: (1) they occur in great abundances in moist habitats, from forest soils to peatlands to lakes and even brackish waters (Mitchell et al., 2008; Ogden and Hedley, 1980); (2) a relatively fast reproduction rate and a large scope of morphological plasticity within species could result in relatively rapid evolutionary change (Charman, 2001); (3) the tolerance of some taxa to harsh environmental conditions (Pérez-Juárez et al., 2017); (4) the presence of a simple test facilitates identification and has resulted in several key taxonomic publications over two centuries of study (Leclerc, 1815; Mitchell et al., 2008); and (5) they occupy several trophic levels according to their feeding behaviour (i.e., bacterivory, detritivory, predatory and mixotrophy; Jassey et al. 2013). In addition, the excellent preservation potential of their test (Lahr et al., 2019; Mitchell et al., 2008; Swindles and Roe, 2007), especially in Holocene and acidic sediments, means they are an excellent proxy for paleoenvironmental reconstructions.

Poor taxonomy remains a major obstacle for the systematic use of Arcellinida as bioindicators for present and past environmental conditions (Kosakyan et al., 2016). Uncertainty as to the range of phenotypic plasticity – i.e., the natural variation of test morphologies not mirrored by variation in the genomes (Mulot et al., 2017) – lowers taxonomic precision in the Arcellinida. Mulot et al. (2017) demonstrated that test shape and pore numbers varied with environmental conditions in wild and cultured populations of *Hyalosphenia papilio* Leidy, 1874. Jennings (1916), in an early study of a clonal culture of *Difflugia corona* (now *Netzelia corona* (Wallich, 1864) Gomaa et al., 2017), observed a variation in both number and placement of spines that was independent of environmental conditions. Porfírio-Sousa et al. (2017) observed very little variability in the test or aperture diameter in clonal cultures of *Arcella intermedia*, but found that wild populations occupied a distinct morphometric space compared to clonal populations, corroborating previous studies (Wanner, 1999). Tremendous phenotypic plasticity present in wild Arcellinida populations was illustrated by Bobrov and Mazei (2004). For each species, results of test length and breadth measurements formed diffuse clouds of transitional morphotypes plotting around weak central peaks rather than dense distributions. Thus, differences in test Arcellinida morphology may not be the result of genomic differences, but merely reflect a species' range of phenotypic variability. Notwithstanding this, Jennings (1916) showed that specific traits (e.g., number of spines and aperture crenulations) could be artificially selected across successive generations, possibly highlighting a pathway for speciation (West-Eberhard, 1989).

Speciation events could result from specific morphotypes that have greater success at exploiting a novel niche (parapatric speciation) (Weisse, 2008). One such pathway could be the development of mixotrophy, whereby testate amoeba host symbiotic green algae (Treboyxiophyceae) in their cytoplasmic vacuoles (Gomaa et al., 2014; Lara and Gomaa, 2017). Mixotrophic taxa represent valuable ecological indicators as they signal nutrient poor conditions, especially in peatlands (Jassey et al., 2013). Jassey et al. (2013) demonstrated that species with large body sizes occupied higher trophic levels and consumed many food sources as compared to smaller species that shared low trophic levels with more specific food strategies. Thus, morphological changes that correspond to genetic differences could be of ecological relevance. Greater taxonomic precision would lead to more robust ecological inferences.

Over a decade of molecular research (Nikolaev et al., 2005) has demonstrated several morphological features to be reliable predictors of genetic relatedness among the Arcellinida: (1) aperture outline (e.g., circular vs. slit); (2) presence of a neck; (3) shell outline (in both apertural and lateral view); and (4) shell composition (Gomaa et al., 2012; Kosakyan et al., 2016; Lahr et al., 2019). Lahr et al. (2019) presented a phylogenomic reconstruction based on 250 genes, obtained using single-cell transcriptomic techniques, which divided Arcellinida into five infraorders: (1) Sphaerothecina Kosakyan et al., 2016; Longithecina Lahr et al., 2019; Excentrostoma Lahr et al., 2019; Hyalospheniformes Lahr et al., 2019; and Vonustoma Lahr et al., 2019. For the Hyalospheniformes, ongoing molecular research has demonstrated that other test morphological features (e.g., compression, size, girth) can be used to reliably distinguish taxa at the genus and species level and how these features might be related to an organism's autecology. This remains to be done for the other infraorders of Arcellinida, such as Longithecina, containing the family Difflugiidae Wallich, 1864, one of the most diverse and abundant in wild assemblages.

To better define the boundary between baseline phenotypic plasticity and evolutionarily significant morphological differences, we require faster evolving genetic markers that can distinguish at the genus or species level and advanced morphometrics to characterise all features of the Arcellinida test. Blandenier et al. (2017) developed the application of the mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene marker (NADH) to construct Arcellinida molecular phylogenies that could both characterise deep phylogenetic relationships as well as finer (i.e., genus and species level) relationships. They were able to replicate the topology of the small subunit ribosomal gene (18S rRNA) trees but could also discriminate between two isolates of a single species (i.e., *Difflugia nodosa* (Leidy, 1879) and *Difflugia pyriformis* Perty, 1849). Previous morphometric analyses used the average measurements (n = 20-30) of several Arcellinida test features (e.g., length, breadth, width) to construct "ideal-individuals" (Wanner, 1999). This approach does not separate size variation from shape variation and these measurements are often correlated since they overlap in what they are measuring (Zelditch et al., 2012). Geometric morphometric analysis provides a detailed summary of morphological variability across the entire specimen with the added benefit of separating size as an independent variable (Adams and Otárola-Castillo, 2013). Shapes are defined by the geometric configuration of landmarks (fixed anatomically definable locations found on all specimens) and by boundary curves drawn between the landmarks. Geometric morphometrics has been applied to paleontological studies (Adams and Otárola-Castillo, 2013), including microfossils (e.g. diatoms; Beszteri et al. 2005) but has yet to be applied to Arcellinida.

By combining high-resolution molecular markers with advanced morphometrics analysis and applying them to Arcellinida with similar test morphologies we aim to further explore the boundary between phenotypic variability and evolutionarily significant morphological differences. We focus on the elongate type difflugiids as they are abundant in lacustrine habitats, and, like most difflugiids, are poorly characterised by current phylogenetic reconstructions. The results will improve ecological inferences based on morphotype assemblages, as a clearer distinction can be made on which morphological features are evolutionarily significant and likely represent an important ecological function. This can further be used to better understand the functional ecological role of Arcellinida and give a better estimate of the true diversity of this group.

Material and Methods

Sample collection and documentation

Difflugia specimens were isolated from *Sphagnum* mosses on Pirin Mountain (*Difflugia acuminata* Ehrenberg, 1830) and lacustrine submerged vegetation from Sofia Southern Park Artificial Lake (*Difflugia lanceolata* Penard, 1890 and *Difflugia oblonga* Ehrenberg, 1838), both in Bulgaria (Table 1). Cells were identified based on several reference articles (e.g., Chardez 1967; Ogden 1979, 1980, 1983; see Appendix A Table A.1) with an Axio Imager M2-Carl Zeiss compound microscope with a digital camera (ProgRes C7) and specialised software (CapturePro Software 2.8). Tests were documented using scanning electron microscopy as described previously (Todorov et al., 2010). We analyzed a single cell extraction for both *D. lanceolata* and *D. oblonga*, and a sample consisting of five morphologically identical cells from the same site for *D. acuminata* (Table 1).

DNA extraction, PCR amplification and sequencing

DNA extraction was performed with a thiocyanate guanidinium solution following the protocol of Duckert et al. (2018) adapted from Chomczynski and Sacchi (1987). New sequences of NADH dehydrogenases were obtained with a novel nested PCR protocol using forward and reverse primers developed by Blandenier et al. (2017) (Table 2). The *D. acuminata* sequence was obtained using the nested PCR product, whereas we cloned the nested PCR product of *D. lanceolata* and *D. oblonga* using a PCR2.1 Topo TA cloning vector and transformed into *Escherichia coli* TOP10′ One Shots cells (Invitrogen kit) following the manufacturer's instructions. For all three taxa, the final PCR product was purified using Millipore kit and sent for sequencing with an ABI3730XL DNA sequencer (Applied Biosystems) at Macrogen, Amsterdam, The Netherlands. Sequences were deposited in GenBank with the following accession numbers: MN524965-MN524967.

Alignment and phylogenetic analyses

Using BioEdit (Hall, 1999), we aligned the three obtained sequences with NAD sequences of Blandenier et al. (2017) and sequences from other Amoebozoan mitochondrial genomes retrieved on GenBank. In order to obtain a better resolution, the nucleotide sequences were converted into amino acid sequences and the intergenic region was removed (see Appendix B). The root was placed on the Discosea based on the phylogenetic reconstruction of Kang et al. (2017). Phylogenetic reconstructions were conducted using the CIPRES Portal (Miller et al., 2010). We used three distinct phylogenetic reconstruction methods: (1) we built a Maximum Likelihood phylogenetic tree using the RaxML v.8.2.10 algorithm (Stamatakis, 2014) with the JTT + model. Bootstrapping was stopped after 150 replicates based on MRE bootstrapping criterion; (2) we built a Bayesian phylogenetic tree using MrBayes (Ronquist and Huelsenbeck, 2003) with a GTR+GAMMA+INVARIANT model ran on two independent chains for 200,000 generations sampled every 100 generations resulting in 4,000 trees, of which 25% were discarded as the burn-in; and (3) parsimony analysis was performed using TNT version 1.1. (Goloboff et al., 2008), running 1,000 bootstrap replicates using the PAM 250 step matrix in a New Technology search allowing sectorial search, ratchet, drift and tree fusing parameters.

Geometric morphometric analyses

To quantify test size and shape variation we followed a modified geometric morphometric protocol based on MacLeod (2008) and Zelditch et al. (2012). Appendix C Table C.1 provides detailed information on the source of the images used for geometric morphometric analysis. Images were digitised using the “tps” series of software (Rohlf, 2015). To digitise an image, a landmark was placed on the left and right edge of the test aperture and a set of 28 equally spaced semilandmarks were used to outline the tests between landmarks 1 and 2 (Fig. 1). Landmark and semilandmark coordinates underwent a Procrustes superimposition using sliders to align all specimens and isolate size variation as an independent variable (Rohlf, 2015; Zelditch et al., 2012). Size variation is quantified by centroid size, defined as the square root of the sum of squared distances of all the landmarks of an object to their centroid (Fig. 1; MacLeod 2008). The centroid is the center of gravity of an object obtained by averaging the x and y coordinates of all landmarks (Fig. 1). We modeled the Arcellinida test shape space using a principal components analysis (PCA) of the Procrustes superimposed landmarks using the “geomorph” R-package (Adams and Otárola-Castillo, 2013). This package was also used to create five shape models along the first principal component (PC1). We used the FIJI distribution of “ImageJ” with the “ObjectJ” plugin to measure test length (mm), width (mm) and aperture width (mm) from scaled Arcellinida test images. Surface area was estimated using the equation for the surface area of an ellipse.

Results

NADH phylogenetic tree

We obtained sequences (254-257 bp) of the NAD9-NAD7 genes of *D. oblonga*, *D. acuminata* and *D. lanceolata* (Fig. 1 and Table 1). The sequence of *D. oblonga* possesses a short intergenic region (indel of 6 bp), whereas the sequences of *D. acuminata* and *D. lanceolata* possess an overlap of 1 bp as described for *Arcella* sequences in Blandenier et al. (2017). The phylogenetic reconstruction (Fig. 2) places the Arcellinida in a single clade (98 bootstrap support (bs)/1.00 posterior probability (pp)). Within Arcellinida, infraorder Sphaerothecina represents a robust clade (93 bs/0.99 pp) and our new sequences plot on two distinct clades. *Difflugia lanceolata* and *D. acuminata* form the “Lanceolate” clade (95 bs/0.99 pp) and *D. oblonga* along with *D. nodosa* and *D. pyriformis* comprise the “Pyriform” clade (95 bs/0.99 pp). The node supports for these three clades within Arcellinida are too low to make any inferences on their placement and relationship to one another (Fig. 2).

Geometric morphometric analysis

The first two Procrustes PCA components explained 82.9% of test morphological variance (Table 3), with the first principal component (PC1) explaining 72.5%. We display geometric morphometric analysis results as a biplot of PC1 scores and centroid size (Fig. 3; see Appendix C Table C.1 for all results). We grouped data points by phylogenetic clade (symbol colour) and by taxa (symbol shape; Fig. 3). Specimens that have been sequenced are represented by larger symbols.

Centroid size has a positive linear to superlinear relationship with test length (mm), test width (mm) and test surface area (mm^2) (Appendix D Fig. D.1). Thus, centroid size serves as a proxy for test surface area and test volume for radially symmetrical tests such as *D. lanceolata*, *D. acuminata*, *D. pyriformis* and *D. oblonga*. Centroid size can be used to distinguish the "Pyriform" clade (centroid size > 486.5; Appendix C Table C.1) from the "Lanceolate" clade with minimal (n = 1) overlap (Fig. 3). Centroid size can also be used to characterise members of the "Lanceolate" clade as *D. lanceolata* has a smaller average centroid size (centroid size < 300; Appendix C Table C.1; Fig. 3). We note that the sequenced specimen of *D. oblonga* (small morphotype) from Gomaa et al. (2017) and Lahr et al. (2019) falls within the centroid size and PC1 score ranges of the "Lanceolate" clade rather than the "Pyriform" clade (Fig. 3).

In Fig. 3, we show PC1 shape models ranging from negative to positive PC1 scores. Negative PC1 scores reflect columnar morphologies and acuminate bases, while positive PC1 scores reflect pyriform morphologies and wide flat bases. *Difflugia oblonga* has similar PC1 scores as members of the "Lanceolate" clade, making PC1 scores an unreliable metric to distinguish between the "Pyriform" and "Lanceolate" clades. However, test shape can be used, with minimal overlap, to distinguish between members of the "Pyriform" clade. Taxa known to house endosymbiotic algae (i.e., mixotrophic taxa) have positive PC1 scores, reflective of wide flat bases and inflation of the lower two thirds of the test.

Discussion

Previous phylogenetic reconstructions of the Arcellinida have either lacked "Pyriform" taxa (Gomaa et al., 2017; Lahr et al., 2019) or "Lanceolate" taxa (Blandenier et al., 2017). The addition of two "Lanceolate" taxa and a third "Pyriform" taxon to the NADH phylogenetic tree allows the molecular comparison of these two morphotypes and reveals significant diverging between them (Fig. 2). Importantly, the test morphologies of both the "Lanceolate" and "Pyriform" taxa are relatively similar, having elongate tests with terminal apertures (Fig. 1). It is thus appropriate to investigate the evolutionary significance of morphological traits other than an elongate test with a terminal rounded aperture.

NADH phylogenetic tree

Our NADH phylogenetic tree replicates the global topology of the 18S rRNA phylogenetic reconstruction of Gomaa et al. (2017) and Lahr et al. (2019) and the NADH phylogenetic reconstruction of Blandenier et al. (2017). They all show Arcellinida as monophyletic containing three well resolved infraorders: Sphaerothecina, Hyalospheniformes and Longithecina. Our tree suggests distinct diverging within the Longithecina between a "Lanceolate" clade consisting of *D. lanceolata* and *D. acuminata*, and a "Pyriform" clade that includes *D. oblonga*, *D. nodosa*, and *D. pyriformis*.

Our phylogenetic placement of *D. oblonga* contradicts that of the 18S rRNA reconstructions of Gomaa et al. (2017) and Lahr et al. (2019) that placed it amongst the "Lanceolate" group. It should be noted that the same sequence for *D. oblonga* was used by both Gomaa et al. (2017) and Lahr et al. (2019). Fig. 3 demonstrates that their specimen of *D. oblonga* is morphologically similar, in both shape and centroid size, to other members of the "Lanceolate" group. Gomaa et al. (2017) first described this specimen as *Difflugia parva* (Thomas, 1954) Ogden, 1983 but reclassified it as *D. oblonga* (small morphotype) following the recommendations of Mazei and Warren (2014), who indicated *D. parva* as a junior synonym of *D. oblonga.* The similarity in shape and centroid size of *D. lanceolata*, *D. acuminata* and *D. parva* could explain why they group together in the

18s rRNA tree of Gomaa et al. (2017). Thus, we recommend that because of its distinct size, *D. parva* should not be synonymized with *D. oblonga* sensu Mazei and Warren (2014).

If the *D. oblonga* (small morphotype) sequenced in the 18S rRNA phylogenetic reconstructions is in fact *D. parva*, a member of the "Lanceolate" group and a distinct species from *D. oblonga*, this leaves no barcoded taxa from the "Pyriform" clade with the 18S rRNA marker. Mazei and Warren (2014) define *D. oblonga* as being of similar shape to *D. pyriformis* but being relatively narrower in breadth. Our NADH phylogenetic reconstruction further demonstrates that within the lanceolate to oblong shape range, centroid size may be a more reliable morphological trait than PC1 scores (i.e., shape) in predicting genetic relatedness. Our specimens of *D. oblonga* are distinct in centroid size from members of the "Lanceolate" group and represent an intermediate shape between the Lanceolate" group and *D. pyriformis* (Fig. 3). Centroid size is also an important characteristic in distinguishing other members of the "Pyriform" clade from the "Lanceolate" clade (Fig. 3). We propose that centroid size is a defining feature in the classification of *D. oblonga* sensu Mazei and Warren (2014) and differentiating it from other members of the "Lanceolate" group.

There are four remaining difflugiids with 18S rRNA sequences, listed in decreasing PC1 score order (Fig. 3): *Difflugia compressa* Carter, 1864*, Difflugia hiraethogii* Ogden, 1983, *Difflugia bryophila* (Penard, 1902) Jung 1942 and *Difflugia bacillariarum* Perty, 1849. *Difflugia bacillariarum* plots within the "Lanceolate" group on Fig. 3 and with the "Lanceolate" group in the 18S rRNA reconstructions (Gomaa et al., 2017; Lahr et al., 2019). *Difflugia bryophila* has the smallest centroid size and PC1 scores that fall within the oblong shape range, but it is not found within the "Lanceolate" group in the 18S rRNA reconstruction (Lahr et al., 2019). *Difflugia hiraethogii* has a compressed and pyriform shape but a centroid size that falls within "Lanceolate" range and groups along with *D. bacillariarum* with the "Lanceolate" group in the 18S rRNA reconstructions (Gomaa et al., 2017; Lahr et al., 2019). *Difflugia compressa* has a centroid size at the lower end of the "Pyriform" range and a shape like that of *D. nodosa*. *Difflugia compressa* groups distinctly, along with *D. bryophila*, from the "Lanceolate" group in the 18S rRNA reconstruction (Lahr et al., 2019). Although *D. compressa* seems like a candidate for the "Pyriform" clade and thus could place this clade within the 18S rRNA reconstructions; grouping of *D. compressa* with *D. bryophila* in the 18S rRNA reconstructions, who do not share similar morphological characteristics, complicates the situation. It might be that other features unite these two taxa. The position of the "Pyriform" clade on 18S rRNA phylogenetic reconstructions remains uncertain and thus members of this clade should be the target of future molecular phylogenies.

Test centroid size and its ecological significance

Based on the data presented, test centroid size is a more reliable metric in differentiating between the "Lanceolate" and "Pyriform" clades than test shape (Fig. 3). Unlike estimates of biovolume that reduce test shape to an idealized shape (i.e., ellipse; Fournier et al., 2012), centroid size is based on the actual dimensions of the test (i.e., distance of each (semi)landmark from the centroid). Thus, centroid size is likely to be the most precise and accurate method for estimating test surface area and estimating test volume for radially symmetrical tests (i.e., *D. lanceolata*, *D. acuminata*, *D. pyriformis* and *D. oblonga*; see Appendix D Fig. D.1).

The different test centroid size ranges of the "Lanceolate" and "Pyriform" clades might be related to ecological differences between these clades. Jassey et al. (2013) in their study of the microbial food web from a *Sphagnum*-peatland demonstrated that taxa with large body sizes, as estimated by test length and biovolume, occupy higher trophic levels with more varied food sources as compared to smaller taxa that occupy lower trophic levels, and have more specific feeding strategies. Coeval changes in test size of multiple taxa have been shown to occur in response to environmental stresses: (1) smaller biovolumes associated with higher temperatures could reflect cellular respiration increasing at a greater rate than oxygen diffusion (Wanner, 1999); (2) reduction in test size in response to warmer climates could be indirectly beneficial, by increasing CO_2 access for the *Chlorella* symbionts in mixotrophic taxa, whose metabolism increases in warmer temperatures (Mulot et al., 2017); (3) Marcisz et al. (2016) observed a decrease in average Arcellinida test size in response to multiple disturbances (e.g., fire and peat farming) in their study of *Sphagnum* peatlands in Poland. Arcellinida test size is thus a trait of major ecological importance and it is imperative to consider it in future studies of Arcellinida communities. We propose that test centroid size to be the most precise and accurate method (to date) of estimating test size and recommend its utilization.

Mixotrophy

Mixotrophic taxa represent valuable ecological indicators as they signal nutrient poor conditions in peatlands (Jassey et al., 2013) and perhaps in other environments such as lakes but this has yet to be tested. Penard (1902) observed that mixotrophic specimens could survive dysoxic conditions and lack of nutrition longer than non-mixotrophic specimens and hypothesized that it was due to the inclusions of zoochlorellae. Leidy (1879) characterised *D. pyriformis* and *D. nodosa* as being bright green and full of "chlorophyll grains" and stated that the cells of *D. acuminata* and *D. lanceolata* are colourless. The absence of mixotrophy in the "Lanceolate" clade might partly explain the genetic divergence from the "Pyriform" clade.

All members of the "Pyriform" clade share a similar centroid size range (Fig. 3) but, unlike the other members, *D. oblonga* does not house endosymbionts (unpublished observation) making centroid size alone an unreliable metric to define mixotrophic taxa. Within the "Pyriform" clade, the combination of centroid size and test shape (i.e., PC1 scores) can characterise mixotrophic taxa as *D. pyriformis* and *D. nodosa* tend to have positive PC1 scores reflecting broader shapes with flatter bases and are greater centroid sizes on average as compared to *D. oblonga* (Fig. 3). In addition, *D. nodosa* is compressed lateral further helping to distinguish it from both *D. oblonga* and *D. nodosa*. Other mixotrophic taxa such as *Difflugia capreolata* Penard, 1902 and *Difflugia bidens* Penard, 1902 are also characterised by larger average sizes and broad morphologies, with *D. bidens* being compressed as well (Penard, 1902). There are several small ovoid arcellinid taxa that have been observed to house symbiotic algae: *Netzelia gramen* (Penard, 1902) Gomaa et al., 2017, *Difflugia lobostoma* Leidy, 1879, *Difflugia labiosa* Wailes, 1919, *Difflugia urceolata* Carter, 1864 and *Cucurbitella mespiliformis* Penard, 1902. However, the 18S rRNA phylogenetic reconstruction shows that these taxa, apart from *D. urceolata* and *C. mespiliformis* that have yet to be phylogenetically placed, belong to a separate infraorder, namely Sphaerothecina, characterised by ovoid to spherical and hemispherical tests (Gomaa et al., 2017). Based on our results, mixotrophic taxa with elongate tests and terminal apertures (i.e., Longithecina) are characterised by both larger average centroid sizes and broader morphologies, especially in the lower portion of the test. This hypothesis requires further testing through the sequencing of missing members of Longithecina. The larger average size of the tests could be the result of the greater amounts of energy provided by the endosymbionts, facilitating increases in test size and growth rate. The broader shapes could provide additional space to house endosymbionts in the cell body. In addition, broadening of the test increases surface area and could also allow for greater transmission of light, with the extreme being compression of the test as seen in *D. nodosa* and *D. bidens*.

Conclusions

Our approach of applying fast evolving genetic markers and advanced morphometric analysis to Arcellinida with comparable test morphologies was successful in highlighting ecologically relevant morphological traits, namely centroid size and characterising test morphology for mixotrophic taxa. Moreover, it represents a valuable tool for future taxonomic studies in creating detailed morphological descriptions of the phenotypic range of taxa. We did not erect new taxa names as the inability to anchor the clades within the Arcellinida makes inferences on their taxonomic rank (e.g., genus) uncertain. Our results highlight the importance of test centroid size when studying elongate difflugiids and the need for continued sequencing. A large test size range could be indicative of an organism's ability to exploit several trophic levels or perhaps a range of different species. Mixotrophic elongate difflugiids are characterised by the inflation of the lower portion of the test, larger centroid size and in some cases lateral compression. Continued research into the ecological relevance of morphological traits will serve to strengthen ecological inferences, increasing the importance of Arcellinida as environmental proxies.

Acknowledgements

Thanks, are expressed to Molly Baur, Amandine Pillonel and David Singer from the Laboratory of Soil Biodiversity at the University of Neuchâtel, Switzerland, and Sarah Hellyar, School of Biological Sciences, Queen's University Belfast for support with molecular barcoding NADH gene of the Arcellinida. We are grateful to Ferry J. Siemensma for sharing his observations of mixotrophic taxa and insights into Arcellinida taxonomy. This work was supported by a Horizon 2020 Marie Skłodowska-Curie Actions Fellowship awarded by the European Commission [Project Number 703381, 2016-2018]; the Swiss NSF [31003A_163254]; Consejera de Educación, Juventud y Deporte, Comunidad de Madrid "Atracción de Talento Investigador" [2017-T1/AMB-5210]; and a FAPESP/QUB SPRINT grant [2015/50327-1].

References

Adams, D.C., Otárola-Castillo, E., 2013. geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods Ecol. Evol. 4, 393–399.

Beszteri, B., Ács, É., Medlin, L., 2005. Conventional and geometric morphometric studies of valve ultrastructural variation in two closely related *Cyclotella* species (Bacillariophyta). Eur. J. Phycol. 40, 89–103.

Blandenier, Q., Lara, E., Mitchell, E.A.D., Alcantara, D.M.C., Siemensma, F.J., Todorov, M., Lahr, D.J.G., 2017. NAD9/NAD7 (mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene)—A new "Holy Grail" phylogenetic and DNA-barcoding marker for Arcellinida (Amoebozoa)? Eur. J. Protistol. 58, 175–186. https://doi.org/10.1016/j.ejop.2016.12.002

Bobrov, A., Mazei, Y., 2004. Morphological Variability of Testate Amoebae (Rhizopoda: Testacealobosea: Testaceafilosea) in Natural Populations. Acta Protozool. 43, 133–146.

Chardez, D., 1967. *Difflugia oblonga* Ehrenberg et ses variétés. Bull Inst Agron Stat Rech Gembloux New Ser. 2, 589–595.

Charman, D.J., 2001. Biostratigraphic and palaeoenvironmental applications of testate amoebae. Quat. Sci. Rev. 20, 1753–1764. https://doi.org/10.1016/S0277-3791(01)00036-1

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156–159.

Duckert, C., Blandenier, Q., Kupferschmid, F.A., Kosakyan, A., Mitchell, E.A., Lara, E., Singer, D., 2018. En garde! Redefinition of *Nebela militaris* (Arcellinida, Hyalospheniidae) and erection of *Alabasta* gen. nov. Eur. J. Protistol. 66, 156–165.

Fournier, B., Malysheva, E., Mazei, Y., Moretti, M., Mitchell, E.A.D., 2012. Toward the use of testate amoeba functional traits as indicator of floodplain restoration success. Eur. J. Soil Biol. 49, 85–91. https://doi.org/10.1016/j.ejsobi.2011.05.008

Goloboff, P.A., Farris, J.S., Nixon, K.C., 2008. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.

Gomaa, F., Kosakyan, A., Heger, T.J., Corsaro, D., Mitchell, E.A.D., Lara, E., 2014. One Alga to Rule them All: Unrelated Mixotrophic Testate Amoebae (Amoebozoa, Rhizaria and Stramenopiles) Share the Same Symbiont (Trebouxiophyceae). Protist 165, 161–176. https://doi.org/10.1016/j.protis.2014.01.002

Gomaa, F., Lahr, D.J.G., Todorov, M., Li, J., Lara, E., 2017. A contribution to the phylogeny of agglutinating Arcellinida (Amoebozoa) based on SSU rRNA gene sequences. Eur. J. Protistol. 59, 99–107. https://doi.org/10.1016/j.ejop.2017.03.005

Gomaa, F., Todorov, M., Heger, T.J., Mitchell, E.A.D., Lara, E., 2012. SSU rRNA Phylogeny of Arcellinida (Amoebozoa) Reveals that the Largest Arcellinid Genus, *Difflugia* Leclerc 1815, is not Monophyletic. Protist 163, 389–399. https://doi.org/10.1016/j.protis.2011.12.001

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, in: Nucleic Acids Symposium Series. [London]: Information Retrieval Ltd., c1979-c2000., pp. 95–98.

Jassey, V.E.J., Meyer, C., Dupuy, C., Bernard, N., Mitchell, E.A.D., Toussaint, M.-L., Metian, M., Chatelain, A.P., Gilbert, D., 2013. To What Extent Do Food Preferences Explain the Trophic Position of Heterotrophic and Mixotrophic Microbial Consumers in a *Sphagnum* Peatland? Microb. Ecol. 66, 571–580. https://doi.org/10.1007/s00248-013-0262-8

Jennings, H.S., 1916. Heredity, variation and the results of selection in the uniparental reproduction of *Difflugia corona*. Genetics 1, 407.

Kang, S., Tice, A.K., Spiegel, F.W., Silberman, J.D., Pánek, T., epi ka, I., Kostka, M., Kosakyan, A., Alcântara, D.M.C., Roger, A.J., Shadwick, L.L., Smirnov, A., Kudryavtsev, A., Lahr, D.J.G., Brown, M.W., 2017.

Between a Pod and a Hard Test: The Deep Evolution of Amoebae. Mol. Biol. Evol. 34, 2258–2270. https://doi.org/10.1093/molbev/msx162

Kosakyan, A., Gomaa, F., Lara, E., Lahr, D.J.G., 2016. Current and future perspectives on the systematics, taxonomy and nomenclature of testate amoebae. Eur. J. Protistol. 55, 105–117. https://doi.org/10.1016/j.ejop.2016.02.001

Lahr, D.J.G., Kosakyan, A., Lara, E., Mitchell, E.A.D., Morais, L., Porfírio-Sousa, A.L., Ribeiro, G.M., Tice, A.K., Panek, T., Kang, S., 2019. Phylogenomics and Morphological Reconstruction of Arcellinida Testate Amoebae Highlight Diversity of Microbial Eukaryotes in the Neoproterozoic. Curr. Biol. 29, 1–11. https://doi.org/10.1016/j.cub.2019.01.078

Lara, E., Gomaa, F., 2017. Chapter 13: Symbiosis between Testate Amoebae and Photosynthetic Organisms, in: Algal and Cyanobacteria Symbioses. World Scientific, pp. 399–419.

Leclerc, L., 1815. Note sur la Difflugie, nouveau genre de polype amorphe. Memoires Paris Mus. Hist. Nat. 2, 474–479.

Leidy, J., 1879. Fresh-water rhizopods of North America. Government Printing Office, Washington.

MacLeod, N., 2008. PalaeoMaths 101 | The Palaeontological Association [WWW Document]. URL https://www.palass.org/publications/newsletter/palaeomath-101/ (accessed 11.12.18).

Marcisz, K., Colombaroli, D., Jassey, V.E.J., Tinner, W., Kołaczek, P., Gałka, M., Karpińska-Kołaczek, M., Słowiński, M., Lamentowicz, M., 2016. A novel testate amoebae trait-based approach to infer environmental disturbance in *Sphagnum* peatlands. Sci. Rep. 6. https://doi.org/10.1038/srep33907

Mazei, Y., Warren, A., 2014. A survey of the testate amoeba genus *Difflugia* Leclerc, 1815 based on specimens in the E. Penard and C.G. Ogden collections of the Natural History Museum, London. Part 2: Species with shells that are pyriform or elongate. Protistology 8, 133–171.

Meisterfeld, R., 2002. Order Arcellinida Kent, 1880, in: Lee, J.J., Leedale, G.F., Bradbury, F. (Eds.), The Illustrated Guide to the Protozoa. Society of Protozoologists, Lawrence, Kansas, USA, pp. 827–860.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, in: Gateway Computing Environments Workshop (GCE), 2010. Ieee, pp. 1–8.

Mitchell, E.A.D., Charman, D.J., Warner, B.G., 2008. Testate amoebae analysis in ecological and paleoecological studies of wetlands: past, present and future. Biodivers. Conserv. 17, 2115–2137. https://doi.org/10.1007/s10531-007-9221-3

Mulot, M., Marcisz, K., Grandgirard, L., Lara, E., Kosakyan, A., Robroek, B.J., Lamentowicz, M., Payne, R.J., Mitchell, E.A., 2017. Genetic determinism vs. phenotypic plasticity in protist morphology. J. Eukaryot. Microbiol. 64, 729–739.

Nikolaev, S.I., Mitchell, E.A.D., Petrov, N.B., Berney, C., Fahrni, J., Pawlowski, J., 2005. The Testate Lobose Amoebae (Order Arcellinida Kent, 1880) Finally Find their Home within Amoebozoa. Protist 156, 191–202. https://doi.org/10.1016/j.protis.2005.03.002

Ogden, C.G., 1983. Observations on the systematics of the genus *Difflugia* in Britain (Rhizopoda, Protozoa). British Museum (Natural History).

Ogden, C.G., 1980. Notes on some Difflugiidae from Norfolk (Rhizopodea, Protozoa). Bull Brit Mus Nat Hist Zool 39, 125–138.

Ogden, C.G., 1979. Comparative morphology of some pyriform species of *Difflugia* (Rhizopoda). Arch. Für Protistenkd. 122, 143–153.

Ogden, G.G., Hedley, R.H., 1980. An Atlas of Freshwater Testate Amoebae: Soil Sci. 130, 176. https://doi.org/10.1097/00010694-198009000-00013

Penard, E., 1902. Faune rhizopodique du bassin du Léman. Geneva.

Pérez-Juárez, H., Serrano-Vázquez, A., Kosakyan, A., Mitchell, E.A.D., Rivera Aguilar, V.M., Lahr, D.J.G., Hernández Moreno, M.M., Cuellar, H.M., Eguiarte, L.E., Lara, E., 2017. *Quadrulella texcalense* sp. nov. from a Mexican desert: An unexpected new environment for hyalospheniid testate amoebae. Eur. J. Protistol. 61, 253–264. https://doi.org/10.1016/j.ejop.2017.06.008

Porfírio-Sousa, A.L., Ribeiro, G.M., Lahr, D.J.G., 2017. Morphometric and genetic analysis of *Arcella intermedia* and *Arcella intermedia* laevis (Amoebozoa, Arcellinida) illuminate phenotypic plasticity in microbial eukaryotes. Eur. J. Protistol. 58, 187–194. https://doi.org/10.1016/j.ejop.2016.11.003

Rohlf, F.J., 2015. The tps series of software. Hystrix 26.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

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

Swindles, G.T., Roe, H.M., 2007. Examining the dissolution characteristics of testate amoebae (Protozoa: Rhizopoda) in low pH conditions: Implications for peatland palaeoclimate studies. Palaeogeogr. Palaeoclimatol. Palaeoecol. 252, 486–496. https://doi.org/10.1016/j.palaeo.2007.05.004

Todorov, M., Golemansky, V., Meisterfeld, R., 2010. Is *Difflugia nebeloides* (Amoebozoa: Arcellinida) really a *Difflugia*? Re-description and new combination. Acta Zool Bulg 62, 13–20.

Wanner, M., 1999. A review on the variability of testate amoebae: methodological approaches, environmental influences and taxonomical implications. Acta Protozool. 38, 15–29.

Weisse, T., 2008. Distribution and diversity of aquatic protists: an evolutionary and ecological perspective. Biodivers. Conserv. 17, 243–259. https://doi.org/10.1007/s10531-007-9249-4

West-Eberhard, M.J., 1989. Phenotypic Plasticity and the Origins of Diversity. Annu. Rev. Ecol. Syst. 2, 249–278.

Zelditch, M.L., Swiderski, D.L., Sheets, H.D., 2012. Geometric morphometrics for biologists: a primer. Academic Press.

Figures

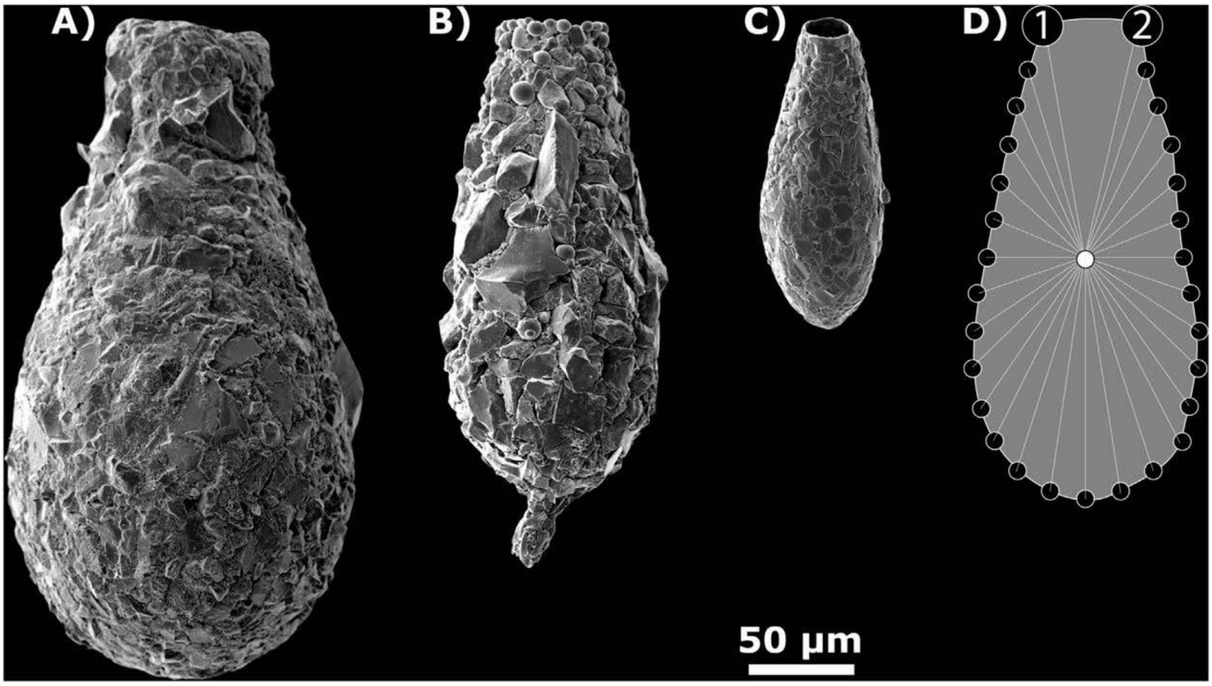

Fig. 1. A-D. Scanning electron microscope images of taxa studied: (**A**) *Difflugia oblonga*; (**B**) *D. acuminata*; (**C**) *D. lanceolata.* All images are to scale. (**D**) Schematic diagram showing geometric morphometric digitization: numbers represent landmarks; small black circles semilandmarks; the white circle is the centroid from which centroid size is derived.

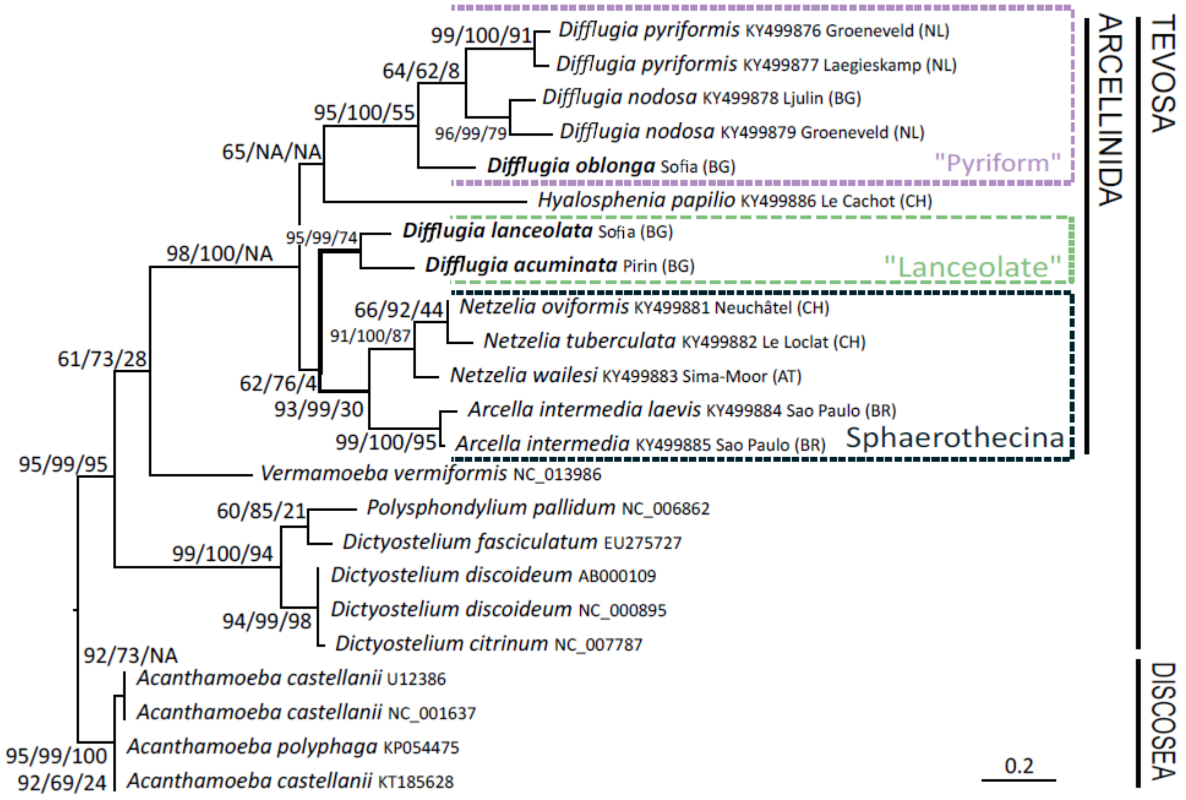

Fig. 2. Maximum likelihood phylogenetic tree of Arcellinida based on NAD9-NAD7 gene sequences. Numbers at the nodes correspond respectively to bootstrap values (bs), posterior probabilities (p.p.), as calculated with Bayesian inference, and parsimony scores (ps) obtained by maximum parsimony analysis. NA values correspond to cases where the topology was diverging between analysis. The tree was rooted with Discosea, as shown by Kang et al. (2017), and new sequences are highlighted in bold.

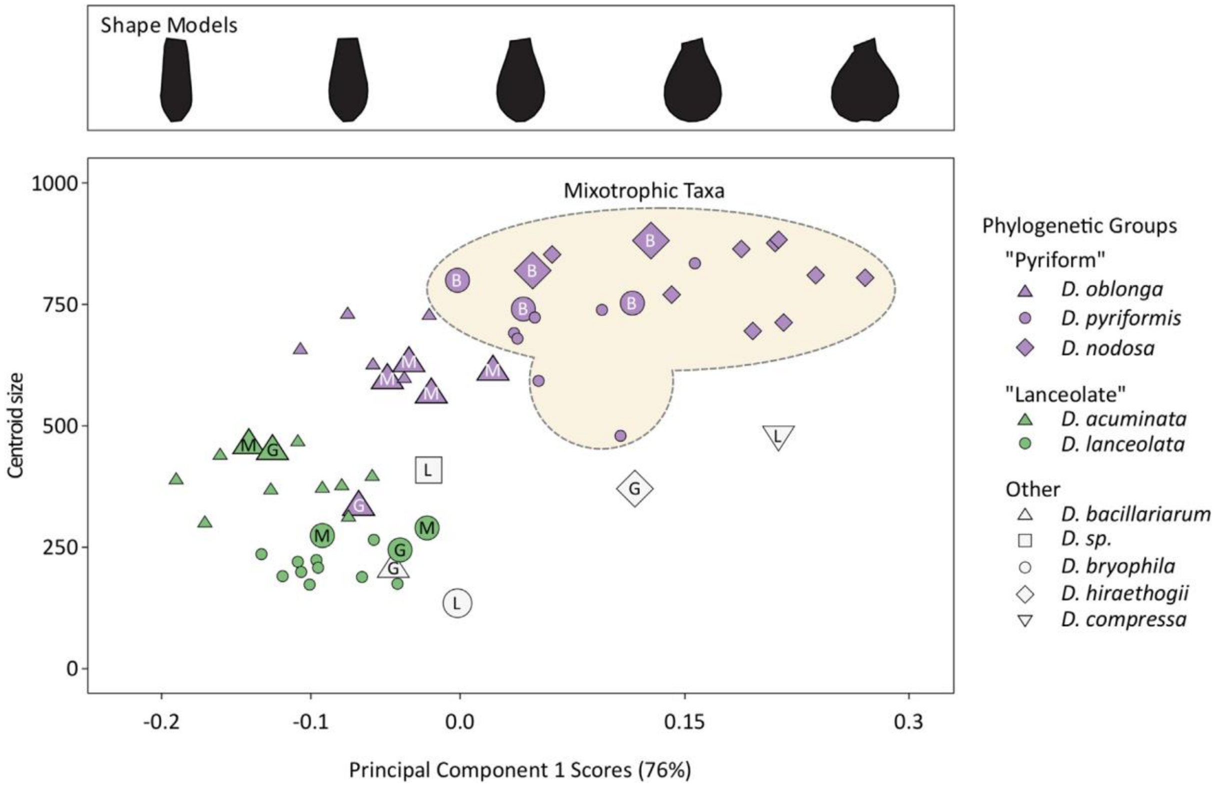

Fig. 3. A biplot of principal component 1 (PC-1) scores (x-axis) and centroid sizes (y-axis) from geometric morphometric analysis. Specimens are illustrated as coloured (phylogenetic clade) symbols (taxa). PC-1 scores are translated into shape models illustrated along the top of the figure. Larger symbols represent specimens that have been sequenced and the letter reflects where the results were published: B - (Blandenier et al. 2017); G - (Gomaa et al. 2012, 2017); L - (Lahr et al. 2019); M - present study.

Tables

Table 1. List of sequenced species, sampling locations and sequence details.

Species	Environments	Location	Latitude	Longitude	Extraction (# of cells)	Sequence (# of clones)	Sequence length (bp)	Genbank Number
Difflugia acuminata	*Sphagnum* mosses	Pirin Mountain, Bulgaria	41.70	23.4	5		254	MN524967
Difflugia lanceolata	Submerged vegetation	Southern Park Artificial Lake, Bulgaria	42.66	23.31	1	5	254	MN524966
Difflugia oblonga	Submerged vegetation	Southern Park Artificial Lake, Bulgaria	42.66	23.31	1	2	257	MN524965

Table 2. Nested PCR protocol.

		PCR1	PCR2
Primer	Forward	NAD9 386F	NAD9 386F
	Reverse	NAD7 1036R	NAD7 67R
Denature	Time (s)	180	180
	Temp. (°C)	94	94
	Cycles (n)	45	45
Denature	Time (s)	30	30
	Temp. (°C)	94	94
Anneal	Time (s)	30	30
	Temp. (°C)	50	61
Elongate	Time (s)	60	60
	Temp. (°C)	72	72
Elongate	Time (s)	600	600
	Temp. (°C)	72	72

Primer sequences	
NAD9 386 F	5′-TGG TTA GAA CGA GAA GTT TGG GAT ATG T-3′
NAD7 1036R	5′-CCA AAT TCT CCT TTT GGH GCY TC-3′
NAD7 67R	5′-GTG CGC AGC AGG RTG TTG WGG WCC-3′

Table 3. Summary of procrustes principal component analysis (first 7 of 60 components).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Standard deviation	0.1184	0.0372	0.0325	0.0246	0.0213	0.0166	0.1291
Proportion of variance	0.7591	0.075	0.057	0.0327	0.0246	0.0149	0.009
Cumulative proportion	0.7591	0.8342	0.8912	0.9239	0.9485	0.9634	0.9724

Appendices (available on https://www.sciencedirect.com/science/article/pii/S0932473919300823)

Appendix A. Taxonomic details for Arcellinida in this study (**Table A.1**).

Appendix B. Alignment after converting nucleotide sequences to amino acid sequences and the removal of the intergenic region (**Table B.1**).

Appendix C. Details of the images used for morphometric analysis, including results from geometric morphometric analysis and ImageJ analysis (**Table C.1**).

Appendix D. Biplots demonstrating the relationship between test centroid size (x-axis) and test length (mm), test width (mm) and test aperture size (mm) (**Fig. D.1**).

Appendix B
Nebela bursella a complicated story

Case 3782 – *Nebela militaris* Penard, 1890 (Arcellinida, HYALOSPHENIIDAE): proposed conservation of the specific name by giving it precedence over *Nebela bursella* Taranek, 1881

Quentin Blandenier[1,2*], Clément Duckert[1,*], Fanny A.L. Kupferschmid[1], Anush Kosakyan[3], Edward A.D. Mitchell[1,4], Enrique Lara[1], David Singer[1,5]

[1]Laboratory of Soil Biodiversity, University of Neuchâtel, Émile-Argand 11, 2000 Neuchâtel, Switzerland
[2]Real Jardín Botánico de Madrid, CSIC, Department of Mycology, Plaza de Murillo 2, 28014 Madrid
[3]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branišovská 31, České Budějovice 37005 Czech Republic
[4]Botanical Garden of Neuchâtel, Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland
[5]Department of Zoology, Institute of Biosciences, University of São Paulo, Brazil
* Co-first authors

Bulletin of Zoological Nomenclature, 2020

Abstract. The purpose of this application, under Article 23.9.3 of the Code, is to conserve the specific name *Nebela militaris* Penard, 1890, a junior subjective synonym of *Nebela bursella* Taranek, 1881 – referred to as *Nebela bursella* Vejdovský in the literature. Due to the absence of any type or reference specimen and due to the confusing original description, doubts about the taxonomic status of *N. bursella* persist. A review of the literature revealed that the names *N. militaris* and *N. bursella* originally referred to the same species, with the name *N. bursella* later being applied erroneously to another species. According to the Principle of Priority, *N. bursella* is the valid name of the species generally known as *N. militaris*, but there has been no mention of the former taxon since 1964 and its name is unknown to most active testate amoeba researchers. To avoid confusion, we propose to conserve the widely used species name *Nebela militaris* Penard, 1890 by granting it conditional precedence over *Nebela bursella* Taranek, 1881, and to designate a neotype.

Keywords. Nomenclature; taxonomy; testate amoebae; Arcellinida; HYALOSPHENIIDAE; *Nebela militaris*; *Nebela bursella*.

1. In 1880, František Vejdovský found specimens of an unknown species of testate amoeba, which he later described as *Nebela bursella* in Vejdovský (1882: 32, pl. II, fig. 2a-c). Before publishing, he mentioned his finding to a fellow naturalist, Karl J. Taranek, who reported this species in 1881 as *Nebela bursella* Vejdovský (Taranek, 1881: 230-231, fig. 2). In this publication, the description and illustration of *Nebela bursella* are congruent with those of Vejdovský (1882). Neither author designated a type specimen or provided any permanent reference slides, and the original material is considered to be lost (Duckert et al. 2018). Because Taranek (1881) attributed the name *Nebela bursella* to Vejdovský, various authors have subsequently referred to it as *N. bursella* Vejdovský, 1881 or as *N. bursella* Vejdovský, 1882. Despite this, Taranek did not explicitly attribute the text of his description to Vejdovský, and the specific name must be attributed solely to Taranek (1881). In consequence, Vejdovský's specific name is simultaneously a junior homonym and junior subjective synonym of Taranek's name.

2. Taranek (1882, pp. 36-38) reported on *Nebela bursella* again, and his illustrations clearly show that he included in this taxon specimens that differed markedly in morphology. While some of the illustrations (Taranek (1882): pl III, fig. 8, 12) are congruent with those of Taranek (1881) and Vejdovský (1882), others clearly do not match the original description. They depict larger and broader specimens which, based on the current taxonomy of the Arcellinida, should rather be assigned to another species of the genus *Nebela* (Taranek, 1882: pl. III, fig. 7; pl. IV, fig. 16), or potentially even to the genus *Heleopera* (Taranek, 1882: pl. III, fig. 9–11) (Duckert et al. 2018, see also Tsyganov et al. 2016 for a comparison between *Nebela* and *Heleopera*). This error was not corrected and led to the misinterpretation that *Nebela bursella* corresponded to the broader morphotypes, which ultimately resulted in *Nebela bursella* being synonymized with *Nebela tincta* (Leidy, 1879) (see Awerintzev, 1906) – now *Nebela tincta* sensu Kosakyan & Lara (2013).

3. Penard (1890, p.164, pl. VII, fig. 16-22) described *Nebela militaris* Penard, 1890 but he neither designated a type nor mentioned any isolated specimens in his description. However, he did mention in his description the location (Tröllhattan, Rosersberg, Marstrand in Sweden and the Vallée de Joux in Switzerland) of the specimens on which he established his species and several permanent slides that he made, of which the exact year of creation is not known but thought to postdate 1900, are still preserved in two distinct collections in the Natural History Museum of London and the Natural History Museum of Geneva. We found five slides containing specimens of *Nebela militaris* after inspection of Penard's permanent slides in both museums, but it is apparent that none of those specimens can be considered as type material, as the registers of the collections show that none have been isolated in the regions mentioned by Penard in his description of *Nebela militaris*. The registers of Penard's permanent slide collections in both the Natural History Museum of London and the Natural History Museum of Geneva have been transmitted to the Commission Secretariat. Given that Penard did not designate any name-bearing type, that no specimen on which he could have based his description was found in his collections of permanent slides and that the absence of type was (and unfortunately still is) more the rule than the exception in the field of taxonomy of testate amoebae, we are certain that there is not any type material for *Nebela militaris*. The specimens that Penard isolated as well as his original description and illustrations of *Nebela militaris* (Penard, 1890) are very similar to the first description of *Nebela bursella* by Taranek (1881). Nonetheless, Penard stated that his was truly a distinct species while acknowledging that some may consider it to be merely an aberrant form of *Nebela bursella*. He did not explain what features of *N. militaris* could be considered as abnormal in comparison to *N. bursella*, but in a subsequent major monograph Penard (1902) considered *N. bursella* a synonym of *N. tincta*, cited by him as *Hyalosphenia tincta* Leidy, 1879. This allows us to infer that Penard (1890) based his concept of *N. bursella* on Taranek's (1882) misidentified specimens and thought that his new species *N. militaris* might be regarded as an abnormally slender form of *Nebela tincta*.

4. The name *Nebela militaris* was rapidly accepted by the scientific community. A list of 26 works by 53 authors using this name as valid for the taxonomic species in question is held by the Commission Secretariat and available upon request. Among these works, we may mention Amesbury et al. (2016), Payne et al. (2011), Mitchell et al. (2008) and Charman et al. (2007). In contrast, the name *Nebela bursella*, having long been considered a junior synonym of *Nebela tincta* (Leidy, 1879) (Awerintzev, 1906) fell into disuse, being used as valid name for the last time by Sudzuki (1964). It is no longer found in its original meaning in major compendia such as Grospietsch (1958) and Tsyganov et al. (2016), and to revive it now as the valid senior subjective synonym of *N. militaris* would lead to confusion.

5. Due to the absence of any type and the lacunar state of the taxonomy of testate amoebae, it is also required that a neotype be designated for *Nebela militaris* Penard, 1890. It has been shown that species of testate amoeba can hide complexes of closely related species that differ only by slight variations in their morphology, even in the case of well-known species (see Kosakyan et al., 2013 as an example). As *Nebela militaris* has been recorded worldwide from a variety of different habitats it is likely that it corresponds to

such a species complex, with each species potentially having a restricted distribution and ecology. However, the illustrations and the descriptions originally made by Penard are not representative of the morphotype found in Europe, and establishing a species on specimens found in such a large area (Switzerland and Sweden) increases the chance of lumping distinct species.

6. To avoid future confusion over the identity of *Nebela militaris* and any potential closely related species yet to be described, we propose to designate a neotype representative of the specimens found in the Swiss Jura mountains and propose this region as the type locality, as it is from this region that the specimens studied by Duckert et al. (2018) and the specimens isolated by Penard originate. Accordingly, we selected the specimen 516-2 from the Penard Collection now deposited at the Natural History Museum of Geneva in Switzerland (https://commons.wikimedia.org/wiki/Category:Nebela_militaris#/media/File:Collection_Penard_MHNG_Specimen_516-2-1_Nebela_militaris.tif) to be set as the neotype.

7. This application is submitted to coincide with the publication of Duckert et al. (2018), who reappraised the taxonomic status of *Nebela militaris* and *Nebela bursella* and thereby became aware of this nomenclatural problem.

8. The International Commission on Zoological Nomenclature is accordingly asked:

(1) to use its plenary power to give the specific name *militaris* Penard, 1890, as published in the binomen *Nebela militaris*, precedence over the specific name *bursella* Taranek, 1881, as published in the binomen *Nebela bursella*, whenever the two are considered synonyms; and

(2) to place on the Official List of Specific Names in Zoology the following names:

(a) *militaris* Penard, 1890, as published in the binomen *Nebela militaris* and as defined by the neotype (specimen 516-2 from the Penard collection in the MHNG) designated in para. 7, with the endorsement that it is to be given precedence over the name *bursella* Taranek, 1881, as published in the binomen *Nebela bursella*, whenever the two are considered synonyms; and

(b) *bursella* Taranek, 1881, as published in the binomen *Nebela bursella*, with the endorsement that it is not to be given precedence over the name *militaris* Penard, 1890, as published in the binomen *Nebela militaris*, whenever the two are considered synonyms.

Acknowledgements

We thank the members of the Commission for their comments, which greatly improved the manuscript. This work was funded by the University of Neuchâtel, Switzerland and the Swiss National Science Foundation (projects No 31003A_141188 to E.M. and 31003A_163254 to E.L.).

References

Amesbury MJ, Swindles GT, Bobrov A, Charman DJ, Holden J, Lamentowicz M, Mallon G, Mazei Y, Mitchell EAD, Payne RJ, Roland TP, Turner TE, Warner BG (2016) Development of a new pan-European testate amoeba transfer function for reconstructing peatland palaeohydrology. Quaternary Science Reviews 152: 132-151.

Awerintzev S (1906) Die Süsswasser-Rhizopoden. Trudy Imperatorskago Sankt-Peterburgskago Obschestva Estestvoisptatelei 36: 249-250.

Charman D J, Blundell A, ACCROTELM members (2007) A new European testate amoebae transfer function for palaeohydrological reconstruction on ombrotrophic peatlands. Journal of Quaternary Science 22(3): 209-221.

Duckert C, Blandenier Q, Kupferschmid FAL, Kosakyan A, Mitchell EAD, Lara E, Singer D (2018) En garde! Redefinition of *Nebela militaris* (Arcellinida, Hyalospheniidae) and erection of *Alabasta* gen. nov. European journal of Protistology 66, 156-165.

Grospietsch T (1958) Wechseltierchen (Rhizopoden). Kosmos-Gesellschaft der Naturfreunde/Franckh’sche Verlagshandlung, Stuttgart, 87 pp.

Kosakyan A, Gomaa F, Mitchell EAD, Heger TJ, Lara E (2013) Using DNA-barcoding for sorting out protist species complexes: A case study of the *Nebela tincta-collaris-bohemica* group (Amoebozoa; Arcellinida, Hyalospheniidae). European Journal of Protistology 49: 222–237.

Leidy J (1879) Freshwater Rhizopods of North America. Report of the United States Geological Survey of the Territories. Volume XII. Government Printing Office, Washington, 324 pp., 48 pls.

Mitchell EAD, Charman DJ, Warner BG, (2008) Testate amoebae analysis in ecological and paleoecological studies of wetlands: past present and future. Biodiversity and Conservation 17: 2115–2137.

Payne RJ, Lamentowicz M, Mitchell EAD (2011) The perils of taxonomic inconsistency in quantitative palaeoecology: experiments with testate amoeba data. Boreas 40 (1): 15-27.

Penard E (1890) Études sur les Rhizopodes d'eau douce. Aubert-Schuchardt, Genève, 320 pp.

Penard E (1902) Faune rhizopodique du bassin du Léman. Henry Kündig, Genève, 714 pp.

Sudzuki M (1964) Zur biologischen Analyse der mikroskopischen Süsswassertierwelt geringster Wassermengen I. Zoological Magazine [Tokyo] 73: 165–174.

Taranek KJ (1881) Beiträge zur Kenntniss der Süsswasser-Rhizopoden Böhmens. Sitzungsberichte der Königlichen Böhmischen Gesellschaft der Wissenschaften in Prag 1881: 220–235.

Taranek KJ (1882) Monographie der Böhmen's. Ein Beitrag zur Kenntniss der Süsswasser-Monothalamien. Abhandlungen der Königlichen Böhmischen Gesellschaft der Wissenschaften 2: 1–56.

Tsyganov AN, Babeshko KV, Mazei YA (2016) A guide to testate amoebae with the keys to genera. Publishing house PSU, Penza, 132 pp.

Vejdovsk F (1882). Thierische Organismen der Brunnenwasser von Prag. Self-published, Prag, 1-70

Appendix C
Spatial modelling of microeukaryotes

Soil protist diversity in the Swiss western Alps is better predicted by topo-climatic than by edaphic variables

Christophe V.W. Seppey[1,2], Olivier Broennimann[3,4], Aline Buri[4], Erika Yashiro[3,5], Eric Pinto-Figueroa[3,5], David Singer[1,6], Quentin Blandenier[1,7], Edward A.D. Mitchell[1,8], Hélène Niculita-Hirzel[9], Antoine Guisan[3,4*], Enrique Lara[1,7*]

[1]Laboratory of Soil Biodiversity, University of Neuchâtel, Émile-Argand 11, 2000 Neuchâtel, Switzerland
[2]Department of Arctic and Marine Biology, University of Tromsø, Framstredet 39, 9019 Tromsø, Norway
[3]Department of Ecology and Evolution, University of Lausanne, Biophore, 1015 Lausanne, Switzerland
[4]Institute of Earth Surface Dynamics, University of Lausanne, Géopolis, 1015 Lausanne, Switzerland
[5]Department of Fundamental Microbiology, University of Lausanne, Biophore, 1015 Lausanne, Switzerland
[6]Department of Zoology, Institute of Biosciences, University of São Paulo, Matão 14, 05508-090 São Paulo, Brazil
[7]Real Jardín Botánico de Madrid, CSIC, Department of Mycology, Plaza de Murillo 2, 28014 Madrid
[8]Botanical Garden of Neuchâtel, Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland
[9] Department of Occupational Health and Environment, Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne, Switzerland
* Co-last authors

Journal of Biogeography, 2020, Vol 47, Issue 4

DOI: 10.1111/jbi.13755

Abstract.

Aim: Trends in spatial patterns of diversity in macroscopic organisms can be well predicted from correlative models, using topo-climatic variables for plants and animals allowing inference over large scales. By contrast, diversity in soil microorganisms is generally considered as mostly driven by edaphic variables and, therefore, difficult to extrapolate on a large spatial scale based on predictive models. Here, we compared the power of topo-climatic versus edaphic variables for predicting the diversity of various soil protist groups at the regional scale.

Location: Swiss western Alps.

Taxa: Full protist community and nine clades belonging respectively to three functional groups: parasites (Apicomplexa, Peronosporomycetes and Phytomyxea), phagotrophs (Sarcomonadea, Tubulinea and Spirotrichea) and phototrophs (Chlorophyta, Trebouxiophyceae and Diatomeae).

Methods: We extracted soil DNA from 178 sites along a wide range of elevations with a random-stratified sampling design. We defined protist Operational Taxonomic Units assemblages by metabarcoding of the V4 region of the rRNA small subunit gene. We assessed and modelled the diversity (Shannon index) patterns of all above-mentioned taxonomic groups based on topo-climatic (topography, slope southness, slope steepness and average summer temperature) and edaphic (soil temperature, relative humidity, pH, electroconductivity, phosphorus percentage, carbon/nitrogen, loss on ignition and shale percentage) variables in Generalized Additive Models (GAM).

Results: The respective significance of topo-climatic and edaphic variables varied among taxonomic and—to a certain extent—functional groups: while many variables explained significantly the diversity of the three phototrophs this was less the case for the three parasites. Topo-climatic variables had a better predictive power than edaphic variables, yet predictive power varied among taxonomic groups.

Main conclusions: Topo-climatic variables (particularly slope steepness and summer temperature if we consider their significance in the GAMs) were, on average, better predictors of protist diversity at the landscape scale than edaphic variables. However, the predictive power of these variables on diversity differed considerably among taxonomic groups; such relationships may be due to direct and/or indirect (e.g. biotic) influences (like with parasitic taxa, where low predictive power is most likely explained by the absence of information on the hosts' distribution). Future prospects include using such spatial models to predict hotspots of diversity and disease outbreaks.

Keywords. Algae; functional groups; macroecology; meadow soils; microbial ecology; parasites; predictive modelling; protozoa; spatial modelling.

Introduction

Protists, i.e. all eukaryotes with the exception of fungi, plants and animals are hyper-diverse in soil systems (Geisen et al., 2018; Mahé et al., 2017), where they play many ecological roles as primary producers, saprotrophs, predators, or parasites (Adl & Gupta, 2006; Geisen et al., 2016), and, thus, play a key role in ecosystem functioning. Photosynthetic groups are essential components of cryptogamic crusts (Elbert et al., 2012; Pushkareva, Johansen, & Elste, 2016) and constitute a significant source of organic carbon for soil organisms (Schmidt, Dyckmans, & Schrader, 2016; Seppey et al., 2017). Predatory protists occupy different levels of the microbial food web, as primary consumers of algae (cyanobacteria or eukaryotic), fungi and bacteria (Bonkowski & Clarholm, 2012; Dumack, Mueller, & Bonkowski, 2016; Hess & Melkonian, 2014), but also occupy higher trophic levels by predating on phagotrophic protists or even micro-Metazoa (e.g. nematodes) (Geisen et al., 2015; Gilbert, Amblard, Bourdier, Francez, & Mitchell, 2000). Parasites are thought to regulate natural populations, notably of animals (Mahé et al., 2017) and can be either very specific such as between the parasitic Gregarines and their animal hosts (Clopton, 2009), or generalist as for Phytomyxea species which can infect hosts from different eukaryotic kingdoms (Neuhauser, Kirchmair, Bulman, & Bass, 2014). Characterizing such complex communities is essential to understand the main on-going ecological processes in soil, and represents a first step towards predicting the effects of environmental changes on communities and, consequently, on ecosystem functioning.

As a whole, soil protist communities have been shown to respond to edaphic conditions, such as gradients of pH (Dupont, Griffiths, Bell, & Bass, 2016), nutrients and moisture (Singer et al., 2018), as well as pesticide amounts (Ekelund, 1999; Foissner, 1999; Nesbitt & Adl, 2014) and other perturbations (Foissner, 1997). These variables are rarely integrated in spatial modelling of biodiversity in general (Mod, Scherrer, Luoto, & Guisan, (2016) for plant communities), especially at broad spatial scales, because they are most often not available at the sites of species observations and not easily generalizable in a spatially-explicit way (Buri et al., 2017; Cianfrani, Buri, Verrecchia, & Guisan 2018; Dubuis et al., 2013). On the other hand, topo-climatic variables (such as slope steepness or air temperature) can be more easily modelled at large spatial scales using digital elevation models based on interpolations of weather stations and/or remote sensing methods. These variables have already proved themselves to be useful to model the spatial distribution of plants and animals (Franklin, 2010; Guisan, Thuiller, & Zimmermann, 2017; Peterson et al., 2011) but much more rarely applied to micro-organisms. As a consequence, spatial modelling of the distribution of microorganisms has been restricted to small areas or aquatic environments (Bulit, 2014; Fraile, Schulz, Mulitza, & Kucera, 2008; King et al., 2010; Langer, Weinmann, Loetters, Bernhard, & Roedder, 2013; Mitchell et al., 2000; Zaric, Schulz,

& Mulitza, 2006; Zinger, Shahnavaz, Baptist, Geremia, & Choler, 2009). Nevertheless, soil protists show broad spatial patterns in their distributions from very different environments and spatial scales (Fernández, 2015; Lara, Roussel-Delif, Fournier, Wilkinson, & Mitchell, 2016; Lentendu et al., 2018; Schiaffino et al., 2016) and the understanding of their eco-geographic requirements could benefits from spatial modelling as much as it benefited macro-organisms. The development of such models at the landscape scale would, if repeated across many regions, allows assessing at a much broader scale the processes driven by micro-organisms, such as nutrient cycling or greenhouse gases fluxes and help improve climatic models. In addition, economic and sanitary management could benefit from microbes spatial modelling, for instance by predicting zones at risk of disease outbreaks and therefore make the use of a potential treatment more parsimonious. A third outcome of spatial modelling of soil microbes could also focus on their conservation by identifying microbe diversity hotspots or refine distribution zones of endemic microorganisms (Cotterill, Al-Rasheid, & Foissner, 2008).

Here, we built spatial predictive models of protist diversity, focusing on general communities as well as on nine broad protist taxa chosen within three functional groups - phototrophs, phagotrophs and parasites - along a wide elevation gradient in the western Swiss Alps. We assessed the diversity of protists in 178 meadow soil samples, resulting from a robust random-stratified field survey by metabarcoding of the V4 regions of the small sub-unit rRNA gene. This study assessed the extent of protist diversity in mountainous meadows and determined to what extent two sets of environmental variables (edaphic and topo-climatic) can predict this diversity over the whole Swiss western Alps of the Vaud state. In addition, we brought an interpretation of the patterns observed based on knowledge of the lifestyles of the different groups surveyed.

Material and Methods

Sampling

Meadow soils were sampled from 194 plots distributed across the Swiss western Alps; of these plots, 178 samples successfully yielded sequencing data and were used in the current study (see Appendix S1.1 in Supporting Information). Sampling was performed from July 4th to September 1st 2013 according to a random stratified sampling design. From each plot, five soil cores (100 g per core between the depths of 0-5 cm after removing plants, mosses, and insects) were taken from the four corners and the center of a 2 m^2 plot. The five cores, were then pooled in a sterile plastic bag and kept in an icebox or at 4 C until DNA extraction and soil analyses were done. A subsample of the pooled soil was also flash frozen at each sampling site and kept frozen until further soil analyses. For more details, see Yashiro et al. (2016).

Edaphic variables

We selected eight edaphic variables, one measured directly on site – the soil temperature at a depth of 5 cm (Soil_temp) – and seven in the lab from the soil samples collected. The soil relative humidity (rh) was assessed by weighing the mass of the soil sample before and after drying at 105 C during 2 days. Soil organic Carbon content was determined by loss of ignition (LOI) at 1050 C. The percentage of shale was determined by laser granulometry. The pH and electro-conductivity (EC) were measured from a soil and Milli-Q water slurry in a 1:2.5 and 1:5 (wt/vol) ratio respectively. Total phosphorus amount (P) was determined by colorimetric analysis after a mineralisation at 550 C with $Mg(NO_3)_2$. The C/N ratio was calculated from the total organic carbon and nitrogen percentages measured by ROCK EVAL pyrolysis (Vinci Technologies, Ruell-Malmaison, France) and combustion infrared spectroscopy (Carlo Erba CNS2500 CHN), respectively. All methods were described in details in Yashiro et al. (2016) and Buri et al. (2017).

Topo-climatic variables

Values for seven topo-climatic variables were retrieved from maps of 25 square meter resolution for each sample location. We used the number of growing degree-days above 0 C (gdd), potential evapotranspiration (etp), topography (topo), slope southness (asp) and slope steepness (slp) (Zimmermann & Kienast, 1999; Zimmermann, Edwards, Moisen, Frescino, & Blackard, 2007). In addition, we calculated the summer temperature average (tmean678) and precipitation sum (psum678) for the months of June to August with values of monthly temperature means and precipitation sums from 1981 to 2010. See Buri et al. (2017) for more details.

Molecular analysis

DNA was extracted from the soil samples using the MoBio PowerSoil DNA extraction kit (Calsbad, CA, USA) following the manufacturer instructions. The V4 region of the 18S rRNA gene was then amplified using the general eukaryotic primers TAReuk454FWD1 and TAReukREV3 (Stoeck et al., 2010). The PCR mix was composed of 3 µL DNA extract, 0.4 µL of 10 mg/mL BSA, 4 µL of PCR buffer (Promega GoTaq M7845), 0.2 µL of Taq polymerase (Promega GoTaq M7845), 0.6 µL of dNTPs (Promega kit U1420), 0.6 µL of each primer (MicroSynth, Balgach, Switzerland), and 10.6 µL of ultra-pure water. The PCR reactions started with a denaturation step at 95 C for 5 min followed by 45 cycles of 94 C for 30 s, 47 C for 45 s and 72 C for 1 min, and terminated with an elongation step of 72 C for 10 min. For each DNA sample, the amplifications were performed in triplicate with a PTC-200 Peltier Thermo Cycler (BioConcept, Allschwil, Switzerland). DNA was then quantified with a Qubit® 2.0 Fluorometer (Invitrogen) and 20 ng of each triplicate were pooled. A DNA library was prepared from the pools using the TruSeq Nano PCR-free Library Preparation kit and the paired-end 2x300 bp sequencing was done on an Illumina® MiSeq at the University of Geneva (Molecular Systematics & Environmental Genomics Laboratory). Sequences are available on European Nucleotide Archive via the project number PRJEB30010 (ERP112373).

Bioinformatics pipeline

Good quality sequences were selected based on their nucleotides phred scores. Every sequence with a phred score average below 20 for a 50 nucleotides window was discarded. The chimeras were then removed using the program VSEARCH 1.11.1 (Rognes, Flouri, Nichols, Quince, & Mahé, 2016) by comparing the environmental sequences 1) with each other for each replicate and 2) against the PR^2 database trimmed according to the V4 primers (downloaded on the 12 September 2016; Guillou et al. 2013). To reduce the noise caused by very rare sequences, we then discarded every singleton. Triplicates were then pooled according to their respective samples and OTUs were built with the program SWARM 2.1.8 (Mahé, Rognes, Quince, de Vargas, & Dunthorn, 2015) with the default options (d=1). The dominant sequence of each OTU was taxonomically assigned by aligning it to the trimmed PR^2 database using the global pairwise alignment program GGSEARCH 36.3.6 (Pearson, 2000).

We removed every OTU that did not belong to protists, namely Metazoa, Embryophyceae and Fungi. We also discarded OTUs with a percentage of identity (PID) below 65% with the database PR^2 as sequences with such low PID are usually of prokaryotic origin (threshold verified manually by aligning low PID environmental sequences on GenBank database). From the 178 plots, 4 were sampled twice and 13 were sampled three times during the sampling period. For each of these 17 plots we took the average (2 samples) or median (3 samples) sequence abundance of each OTU for the samples from the same plot. In addition of the total protist community matrix, we also selected nine broad taxonomic groups (I.e. clades, low taxonomic resolution: Adl et al. 2019) from three functional groups (1) parasites: Apicomplexa, Peronosporomycetes, Phytomyxea; (2) phagotrophs: Sarcomonadea (sensu Howe et al., 2011), Tubulinea, Spirotrichea and (3) phototrophs: Chlorophyceae, Trebouxiophyceae, Diatomeae). These taxa were selected because they are abundant and diverse in soils and are functionally homogeneous. For each of these taxa, we established a PID threshold verified manually on GenBank to discard potential misidentification (Apicomplexa: 80%, Peronosporomycetes: 80%, Phytomyxea: 75%, Sarcomonadea: 80%, Tubulinea: 75%, Spirotrichea: 90%, Chlorophyceae: 90%, Trebouxiophyceae: 85%, Diatomeae: 77%).

Richness and diversity analyses

For each of the ten taxonomic data sets (all protists plus nine broad groups), OTU richness and Shannon diversity (H) were calculated, and the differences between their statistical distributions tested by a multiple comparisons of mean rank sums test (Nemenyi test; Hollander, Wolfe, & Chicken, 2015, posthoc.kruskal.nemenyi.test function, 'PMCMR' package 4.1; Pohlert, 2014). The relation between H and the proportion of sequences kept (non-Metazoa/Embryophyceae/Fungi) was also measured to verify if the percentage of non-wanted taxa were biasing the diversity estimate (Spearman correlation tests).

To assess how much predictors impact the protist diversity and distribution, a nonmetric multidimentional scaling (NMDS) was calculated on the ten Bray-Curtis distance matrices and environmental predictors were fitted to the ordinations (envfit function, 'vegan' package 2.5-2; Oksanen et al., 2018). We ran the analyses on all samples with at least two OTUs.

For each of the ten data sets, H was modelled as a function of the environmental variables using a Generalized Additive Model (GAM; assuming Gaussian residuals and identity link function). For each data set, three models were calibrated; the first with topo-climatic variables only, the second with edaphic variables only, and the third with both sets of variables. All models were iterated 100 times based on bootstraps composed of 80% of the 178 original samples. In total 10x3x100 models were fitted. For each model, the predictive power was estimated as the Root Mean Square Error (RMSE) calculated on the independent samples not included to build the model (20% left-out samples). The effect of taxonomic group and the set of predictors on predictive power (RMSE) was tested by a Nemenyi test. Finally, the diversity values of the nine broad taxa and total protist diversity were extrapolated across the full area of the western Swiss Alps based on a GAM including the topo-climatic variables (i.e. the only spatially-explicit variables).

Results

Observed diversity patterns

We retrieved a total of 24'322'487 good quality sequences of which 97% were not chimeric and 71% were not singletons. The 17'110'114 remaining sequences were clustered into 41'048 OTUs of which 19'260 were assigned to protists (see Appendix S2.4 in Supporting Information). Protist diversity was dominated (proportion of sequences) by Cercozoa, (principally Sarcomonadea and Thecofilosea), and Alveolata of which more than half were assigned to Apicomplexa and ca. 45% to Ciliophora (mostly from classes Spirotrichea, Oligohymenophorea, Litostomatea and Colpodea) (see Appendix S2.6 in Supporting Information). The three other dominant groups were the Stramenopiles (including Peronosporomycetes and Diatomeae), Amoebozoa (including Tubulinea) and Archaeplastida (with Chlorophyceae and Trebouxiophyceae) (see Appendix S2.6).

The nine taxa selected jointly contributed to more than half (54%) of all retained sequences and represented over 35% of the total OTU richness (see Appendix S2.4). The average richness per sample of these clades varied from 7 (Phytomyxea) to 249 (Sarcomonadea). Richness was on average lowest for phototrophs (15 OTUs / sample) and highest for phagotrophs (122 OTUs / sample; Fig. 1). Shannon diversity indices followed the same trend, varying from an average value of 1.1 (Phytomyxea) to 4.3 (Sarcomonadea). The correlation tests revealed that only Spirotrichea, Trebouxiophyceae and Diatomeae were showing a relation between diversity and the proportion of sequences kept in a sample (see Appendix S2.5 in Supporting Information).

Environmental models of diversity

A total of 15 edaphic and topo-climatic parameters were determined for each site. Correlation analysis of topo-climatic and edaphic variables indicated clustering and interdependence for tmean678, gdd, etp and psum678 ($|r| > 0.7$; Dormann et al., 2013). Consequently, only tmean678 was kept from these four topo-climatic variables for further analyses (see Appendix S1.2 and Appendix S1.3 in Supporting Information).

We found that the distribution and diversity of the overall protist community and of each taxon were explained by different environmental factors. The overall protists and nine taxa community distribution was structured mostly by pH and rh as well as, on a lesser extent, by Soil_temp, shale, topo, and tmean678 (see Appendix S2.8 and Appendix S2.9 in Supporting Information). In contrast, distinct taxonomic groups were structured by specific profiles of environmental predictors when diversity was modelled by GLM (Table1, see Appendix S3.10 in Supporting information). One example was the Diatomeae where the topo-climatic predictors (slp and tmean678) seemed to be as important in explaining diversity as edaphic predictors (pH and rh) for the overall community (Fig.2, see Appendix S3.10). Similarly, the diversity was much significantly explained by slp for Tubulinea, Sarcomonadea and Phytomyxea, tmean678 for Spirotrichea or pH for Sarcomonadea. The significance of certain environmental predictors was even more accentuated when only topo-climatic predictors were taken into account as for the tmean678 (see Appendix S4.11 and Appendix S4.12 in Supporting Information).

The predictive power showed lower RMSE values (i.e. a better power) for the topo-climatic than for the edaphic variables for all taxa except for the Chlorophyceae, Trebouxiophyceae and Sarcomonadea for which the values were higher or similar (Fig. 2). In addition, the RMSE of the models calculated on the edaphic and topo-climatic variables together were never significantly lower than the RMSE calculated for the topo-climatic variable alone. The RMSE also varied among taxonomic groups when a given set of variables was considered and the diversity of certain taxa were significantly better predicted (e.g. Peronosporomycetes) than others (e.g. Apicomplexa) (Fig. 3). The predictive power of the overall community was in general lower in comparison to the specific taxonomic groups for all sets of variables even if some taxa were less adequately predicted (e.g. Apicomplexa with the topo-climatic predictors or Diatomeae when all predictors).

Discussion

General patterns of protist communities in soils

Our study revealed several important findings on patterns of protist communities across temperate mountain landscapes. Phagotrophs (e.g. Sarcomonadea & Tubulinea) and parasites (Apicomplexa) were the most abundant functional groups in terms of read abundance. Apicomplexan sequences, albeit numerous, were proportionally much less abundant and diversified than in Neotropical soils: as arthropods are less abundant and diversified in temperate regions, this brings further support to the hypothesis that soil apicomplexan communities mirror that of arthropods in the ecosystem (Mahé et al., 2017). Another abundant parasitic group is the Peronosporomycetes (until recently refered as Oomycota: Stramenopiles), which contains many plant parasites, but also animal pathogens and a few free-living, saprotrophic forms (Beakes, Glockling, & Sekimoto, 2012; Lara & Belbahri, 2011). Peronosporomycetes are shown to be common and diverse in temperate soil systems (Seppey et al., 2017; Singer et al., 2016). By contrast they are less abundant and diverse in neotropical forest soil ecosystems, where they comprise mostly animal parasites (Mahé et al., 2017).

Within phagotrophs, the high proportion of sequences from Cercozoa (mostly to Sarcomonadea) was in line with previous soil eukaryotic DNA surveys (Bates et al., 2013; Harder et al., 2016; Seppey et al., 2017). Furthermore, earlier studies based on microscopy observations showed the prevalence of these groups in soils (Adl & Gupta, 2006). Ciliates were also a well-represented phagotrophic group, and were dominated by Spirotrichea, which corroborates also other findings on soil protist molecular diversity (Lara, Berney, Ekelund, Harms, & Chatzinotas, 2007). In summary, the protist communities found in the Swiss western Alps were typical for temperate soil ecosystems and the findings can likely be extrapolated to other climatically similar regions. However, soil communities have been shown to differ in their composition in contrasted climates such as neotropical rainforests (Mahé et al., 2017); therefore, it can be expected that communities from desert, hypersaline soils and other such extreme ecosystems may differ in their structure and may also be controlled by other sets of predictors.

Our data are in accordance with previous studies related to the impact of edaphic variables on protist communities (Dupont, Griffiths, Bell, & Bass, 2016; Ekelund, 1999; Foissner, 1997; Foissner, 1999; Nesbitt & Adl, 2014; Singer et al., 2018) but we also show that topo-climatic predictors explain equally well soil protists distributions. Therefore, the method of measurement of the predictors (in situ for edaphic variables or remote sensing / modelling for topo-climatic) did not seem to affect our capacity to explain protist community distribution.

Model fit and predictive power of topo-climatic and edaphic variables on protist diversity

Slope steepness and pH were the two variables most often found to significantly contribute to the fit of our different protist diversity models. Slope steepness affects drainage and leaching of nutrients and is generally inversely correlated to soil depth. Nevertheless, an enhanced drainage reduces the likelihood of water-logging, which would select for very specialized protists tolerating anoxia and generally would lead to lower diversity. Soil pH is well known as a major driver of microbial diversity, including protists (Bates et al., 2013; Dupont et al., 2016; Shen, et al., 2014), but also bacteria (Santoyo, Hernández-Pacheco, Hernández-Salmeron, & Hernández-Leon, 2017; Yashiro et al., 2016), and fungi (Noyce et al., 2016; Pellissier et al., 2014; Zhang, Jia, & Yu, 2016). The relationship between pH and protist diversity was significant only for three groups, being negative for two groups of phagotrophs (Spirotrichea and Sarcomonadida) and positive for Chlorophyceae. It is unclear if these relationships reflect a direct effect of pH or rather indirect effects such as biotic effects (e.g. impact on bacterial or fungal food sources), the availability of nutrients for the growth of autotrophs (correlation coefficient between pH and EC = -0.14; P = 0.055), shifts in plant-microbial interactions and root exudate composition (Yashiro et al., 2018), or other drivers.

Predictability varies also to a large extent among taxonomic groups. Indeed, while many variables explained significantly the diversity of the three groups of phototrophs and phagotrophs, it was less so for parasites (Apicomplexa and Peronosporomycetes) particularly when only topo-climatic variables were taken into account (see Appendix S4.11). The latter functional group depends directly on the availability of host species and only indirectly on environmental values for which the influence will be indirect and the fraction of variance explained by these variables and their significance will thus be lower. Additionally, the contrasted predictive power among taxonomic groups may be due to the fact that certain taxa are better discriminated by metabarcoding than others because of differences in taxonomic resolution of the 18S rRNA gene. In our dataset, many OTUs assigned to Apicomplexa and Tubulinea were considered as undetermined as their identity with the best match in the database did not reach 80% (see Appendix S2.7 in Supporting Information). Nevertheless, while the diversity of Apicomplexa was poorly predicted in comparison with other taxa, the models predicted the diversity of Tubulinea with an accuracy that was comparable with other taxa. For nine out of the ten taxonomic group tested, the predictive power of the topo-climatic variables was either significantly better, or at least not different than the ones including the edaphic variables. Moreover, it was never lower than the predictive power of the models including both sets of variables. This suggests that, within the levels of predictability achieved, predictive models built solely on topo-climatic variables are as accurate, or possibly even better, than the models built with the addition of edaphic variables. These variables are available at large scales and are already largely used for modelling the spatial distribution of macro-organisms (Guisan & Zimmermann, 2000), to the contrary of local edaphic values that are always tedious and costly to measure in the landscape across large regions and environmental gradients. These findings open the way to larger sampling designs that could further increase the performance of models.

Interpretation of the spatial patterns of protist diversity modelled with topo-climatic variables

As for macro-organisms (D'Amen, Pradervand, & Guisan, 2015; Dubuis et al., 2011; McCain, 2005; Reymond, Purcell, Cherix, Guisan, & Pellissier, 2013), and increasingly reported for other soil micro-organisms (Geml, Morgado, Semenova-Nelsen & Schilthuizen, 2017; Pellissier et al., 2014), protists diversity showed clear spatial and elevational patterns when only topo-climatic variables were taken into account to build the model (Fig. 4). This pattern seemed to be driven by summer temperature in most cases (see Appendix S4.11 and Appendix S4.12), either in a positive (Diatomeae, Phytomyxea and Tubulinea), unimodal (Apicomplexa, Sarcomonadea and Spirotrichea) or negative way (Chlorophyceae, Peronosporomycetes). A positive correlation of diversity with temperature (and, thus, productivity) is a typical pattern in macroecology that can be related to the species-energy hypothesis as long as moisture is not a limiting factor. This pattern has already been demonstrated for protist communities, more exactly testate amoebae (Fernández et al., 2016; Lara et al., 2016), a paraphyletic group which comprises, interestingly, many Tubulinea (Adl et al., 2019). Other related models for diversity patterns, like elevation gradients (Huston, 1994; see Spehn & Körner, (2009)) have been also shown in testate amoebae (Heger et al., 2016). On the other hand, if moisture is limiting, unimodal patterns are to be expected, and diversity peaks where both moisture and energy are optimal (water energy model: Fernández et al. (2016)) intermediate disturbance hypothesis or mid-domain effect (discussed for the same area in Dubuis et al. (2011)). Finally, Chlorophyceae and Peronosporomycetes are typically sensitive to high temperatures and desiccation, both including often flagellated life stages for dispersal that needs at least a thin water film to disperse (Jeger & Pautasso, 2008). In addition, Chlorophyceae high diversity in the lowest temperature zone (Fig.4, see Appendix S4.12) could be explained by the fact that micro-eukaryotic algae have a higher growth rate at low temperatures, favouring diversification in cold environments (Rose & Caron, 2007) or possibly reduced competition from vascular plants. However, while these patterns can be observed in some groups, they cannot be extended to the whole protist community; indeed, another study showed no significant correlation between elevation and diversity when considering entire microbial eukaryotic communities (Shen et al., 2014).

Technical and methodological issues

The correspondence between OTUs and biological species has always been a hot topic in eukaryotic environmental microbiology. The V4 region of the gene coding for the RNA molecule of the small subunit of the RNA (SSU rRNA or SSU for short) has been listed among a handful DNA fragment for protist barcoding (Pawlowski et al., 2012). However, a single SSU rRNA gene sequence may include, in certain groups, a wide diversity of species with different lifestyles and ecological preferences. This has been shown for different soil protists such as ciliates (Lara, & Acosta-Mercado, 2012). In contrast, in Myxomycetes (Amoebozoa), SSU sequences are truly hypervariable and discriminate relatively accurately between species; intragenomic polymorphism of SSU sequences has been even detected (Dahl et al., 2018), which may artificially inflate interpretations on environmental diversity. However, the accuracy of the estimation can be expected to increase with the narrowing of the taxonomic range of the investigated organisms, as evolutionary drivers become more homogeneous. In other words, comparing the diversity of e.g. Apicomplexans between two localities can be reasonably expected to be more accurate than comparing the whole eukaryotic diversity.

Computation of H indices includes quantitative data, classically the proportion of a given species in a given sample, which can be reasonably inferred by numbers of reads in High Throughput Sequencing data. Indeed, there is a correspondence between this number of reads and the biovolume (calculated from simple measurements of the cell and assuming geometrical shapes; Charrière et al., 2006) of individual organisms that has been shown for many groups of protists (Giner et al., 2016; Kosakyan, Mulot, Mitchell, & Lara, 2015;). De Vargas et al. (2015) showed a linear relationship between the logarithm of organisms length versus the logarithm of 18S rRNA copy numbers (see Supplementary figure W4 in that article). Thus, H indices provide a satisfactory cell quantification based on sequence data in species that display a stable rDNA copy number (Rodriguez-Martinez et al.2009). Nevertheless, if an organism violates the correspondence by inducing more reads per biovolume, its presence in a sample would be translated by OTUs covering a large proportion of the community which would result in lower diversity. For example, Foraminifera are particularly prone to biases in inferring the abundance from rDNA sequences due to alternation of generation, variation in ploidy, and variation in number of nuclei (Weber & Pawlowski, 2013). Similar biases have been shown for ciliates, which are known for having highly polyploid macronuclei, and sometimes smaller species may have higher rDNA copy number than larger cells (Dunthorn et al., 2014). To avoid such biases, the sequence abundance of each species needs to be normalized by rDNA copy number. However, such an approach requires a previous characterization of the species rRNA genes, and therefore cannot be applied to unknown biodiversity. It is therefore important to keep this in mind when assessing the diversity of groups with a heterogeneous number of reads per biovolume.

In some cases, even when avoiding collinearity, it is possible that some edaphic and topo-climatic predictors still depend on each other (e.g. soil temperature and tmean678; Yashiro et al., 2016 see Appendix S1.2). Nevertheless, the comparison between two dependent variables coming from different measurement methods (in situ for edaphic and remote sensing for topo-climatic predictors respectively) is still interesting because we aim to assess if topo-climatic predictors achieve at least as accurate models as edaphic ones.

Conclusion

We showed that the diversity of some taxa belonging to major functional groups in the Swiss western Alps, was explained up to >30% by topo-climatic and edaphic conditions. A somewhat surprising result was that topography and climate predicted protist diversity as well or better than the edaphic variables more commonly used in soil microbial studies. This implies that soil protist diversity patterns could be at least partly inferred, for some groups (e.g. Chlorophyceae) and to some extent (22%), based on topo-climatic spatial models only. The applicability of spatial modelling of protists diversity to soil under other climates than temperate is still to be established. Nevertheless, considering that spatial patterns of micro-organisms is increasingly recognized, it is likely that spatial modelling will become a powerful tool in microbial ecology in the near future.

Such an approach could be applied at finer taxonomic levels to predict the distribution of individual species, which would be of high socio-economic relevance in the case of invasive agricultural or forestry pests of economic importance such as certain Peronosporomycetes. The models could be improved by refining the taxonomic groups, as taxa responding more homogeneously to environmental conditions may show stronger correlation with abiotic variables than the broad group classification we used. For instance, the Peronosporomycetes contain organisms belonging to other functional groups than parasites (e.g. saprotroph; Beakes et al., 2012; Lara & Belbahri, 2011) or able to target a wide range of hosts (e.g. *Phytophothora cinnamomi*; Hardham, 2005). Therefore, other modelling techniques, such as calculating the diversity after modelling the abundance of individual OTUs and stacking then (i.e. staked-SDMs; Guisah & Rahbek, 2011), could enhance the predictive power on certain taxa. These improvements would pave the way toward extrapolation of protists diversity across large spatial scales and provide useful tools to identify biodiversity hotspots, predict spatially the risk of pathogen infection or model soil protist diversity according to future environmental change scenarios.

Acknowledgements

The authors would like to thank all fields and laboratory assistants and technicians who participate in the alpine soil project, with a particular thank to Amandine Pillonel, Laura Desponds and Dessislava Savova Bianchi for the laboratory work. Special thanks to the Transports Publics du Chablais and Glacier 3000 (https://www.glacier3000.ch/en) who provide their installations for free to carry the soils from the sampling sites. We would also like to thanks all anonymous farmers who allowed sampling on their lands. The study was funded by the Swiss National Science Foundation under the projects 310003A 143960 to E.L., 31003A-152866 (SESAM'ALP), PDFMP3-135129 (MICROBIAL BIOGEOGRAPHY) and CR23I2-162754 (INTEGRALP) to A.G. and P2NEP3-178543 to D.S., as well as the internal funding of the Universities of Neuchâtel and Lausanne. E.L. would also like to thanks the program "Atracción de talentos" from the Community of Madrid project 2017-T1/AMB-5210 and the project MYXOTROPIC VI (PGC2018-094660-B-I00) from the Spanish Governement. E.Y. also thanks the European Community FP7-PEOPLE-2010-IIF program (MP-Alps, grant agreement 273965), the Agassiz Foundation, and the Pro-Femmes Fellowship program from the Faculty of Biology and Medicine of the University of Lausanne. C.V.W.S. also thanks the Norwegian Research Council projects 270252 (BiodivERsA-Climate change impacts on Arctic soil and lake microbiomes) and 256132 (ERAnet-LAC, METHAnogenic Biodiversity and activity in Arctic and Subantarctic Ecosystems affected by climate change). A significant part of the computations was also performed on resources provided by the Calculations Center of the Faculty of Science of the University of Neuchâtel and by UNINETT Sigma2 – the National Infrastructure for High Performance Computing and Data Storage in Norway.

References

Adl, S. M. , & Gupta, V. V. S. R. (2006). Protists in soil ecology and forest nutrient cycling. *Canadian Journal of Forest Research*, 36(7), 1805–1817.

Adl, S. M., Bass, D., Lane, C. E., Lukes, 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.

Bates, S. T., Clemente, J. C., Flores, G. E., Walters, W. A., Parfrey, L. W., Knight, R., & Fierer, N. (2013). Global biogeography of highly diverse protistan communities in soil. *Isme Journal*, 7(3), 652–659.

Beakes, G. W., Glockling, S. L., & Sekimoto, S. (2012). The evolutionary phylogeny of the oomycete "fungi". *Protoplasma*, 249 (1), 3–19.

Bonkowski, M., & Clarholm, M. (2012). Stimulation of Plant Growth through Interactions of Bacteria and Protozoa: Testing the Auxiliary Microbial Loop Hypothesis. *Acta Protozoologica*, 51(3), 237–247.

Bulit, C. (2014). Good Reasons and Guidance for Mapping Planktonic Protist Distributions. *Acta Protozoologica*, 53(1), 13–27.

Buri, A., Cianfrani, C., Pinto-Figueroa, E., Yashiro, E., Spangenberg, J. E., Adatte, T., ... Pradervand, J. N. (2017). Soil factors improve predictions of plant species distribution in a mountain environment. *Progress in Physical Geography*, 41(6), 703–722.

Charrière, F., Pavillon, N., Colomb, T., Depeursinge, C., and Heger, T. J., Mitchell, E. A. D., … Rappaz, B. (2006). Living specimen tomography by digital holographic microscopy: morphometry of testate amoeba. *Optics Express*, 14(16), 7005-7013.

Cianfrani, C., Buri A., Verrecchia E., Guisan A. (2018). Generalizing soil properties in geographic space: Approaches used and ways forward. *PloS one*, 13(12), e0208823.

Clopton, R. E. (2009). Phylogenetic relationships, evolution, and systematic revision of the septate Gregarines (Apicomplexa: Eugregarinorida: Septatorina). *Comparative Parasitology*, 76(2), 167–190.

Cotterill, F. P. D., Al-Rasheid, K., Foissner, W. (2008). Conservation of protists: is it needed at all? *Biodiversity and Conservation*, 17(2), 427-443.

Dahl, M. B., Brejnrod, A. D., Unterseher, M., Hoppe, T., Feng, Y., Novozhilov, Y., ... Schnittler, M. (2018). Genetic barcoding of dark-spored myxomycetes (Amoebozoa)-Identification, evaluation and application of a sequence similarity threshold for species differentiation in NGS studies. *Molecular Ecology Resources*, 18(2), 306–318.

D'Amen, M., Pradervand, J. N., & Guisan, A. (2015). Predicting richness and composition in mountain insect communities at high resolution: a new test of the SESAM framework. *Global Ecology and Biogeography*, 24(12), 1443–1453.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carre, G., ... Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27–46.

Dubuis, A., Pottier, J., Rion, V., Pellissier, L., Theurillat, J. P., & Guisan, A. (2011). Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. *Diversity and Distributions*, 17 (6), 1122–1131.

Dubuis, A., Giovanettina, S., Pellissier, L., Pottier, J., Vittoz, P., & Guisan, A. (2013). Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables. *Journal of Vegetation Science*, 24(4), 593–606.

Dumack, K., Mueller, M. E. H., & Bonkowski, M. (2016). Description of *Lecythium terrestris* sp nov (Chlamydophryidae, Cercozoa), a Soil dwelling protist feeding on fungi and algae. *Protist*, 167(2), 93–105.

Dupont, A. O. C., Griffiths, R. I., Bell, T., & Bass, D. (2016). Differences in soil micro-eukaryotic communities over soil pH gradients are strongly driven by parasites and saprotrophs. *Environmental Microbiology*, 18(6), 2010–2024.

Ekelund, F. (1999). The impact of the fungicide fenpropimorph (Corbel (R)) on bacterivorous and fungivorous protozoa in soil. *Journal of Applied Ecology*, 36(2), 233–243.

Elbert, W., Weber, B., Burrows, S., Steinkamp, J., Buedel, B., Andreae, M. O., & Poeschl, U. (2012). Contribution of cryptogamic covers to the global cycles of carbon and nitrogen. *Nature Geoscience*, 5(7), 459–462.

Fernández, L. (2015). Source–sink dynamics shapes the spatial distribution of soil protists in an arid shrubland of northern Chile. *Journal of Arid Environments*, 113, 121-125

Fernández, L. D., Fournier, B., Rivera, R., Lara, E., Mitchell, E. A. D., & Hernandez, C. E. (2016). Water-energy balance, past ecological perturbations and evolutionary constraints shape the latitudinal diversity gradient of soil testate amoebae in south-western South America. *Global Ecology and Biogeography*, 25(10), 1216–1227.

Foissner, W. (1997). Protozoa as bioindicators in agroecosystems, with emphasis on farming practices, biocides, and biodiversity. *Agriculture Ecosystems & Environment*, 62(2-3), 93–103.

Foissner, W. (1999). Soil protozoa as bioindicators: pros and cons, methods, diversity, representative examples. *Agriculture Ecosystems & Environment*, 74(1), 95–112.

Fraile, I., Schulz, M., Mulitza, S., & Kucera, M. (2008). Predicting the global distribution of planktonic foraminifera using a dynamic ecosystem model. *Biogeosciences*, 5(3), 891–911.

Franklin, J. (2010). *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge: Cambridge University Press.

Geml, J., Morgado, L. N., Semenova-Nelsen, T. A., & Schilthuizen, M. (2017). Changes in richness and community composition of ectomycorrhizal fungi among altitudinal vegetation types on Mount Kinabalu in Borneo. New Phytologist, 215(1), 454–468.

Geisen, S., Rosengarten, J., Koller, R., Mulder, C., Urich, T., & Bonkowski, M. (2015). Pack hunting by a common soil amoeba on nematodes. *Environmental Microbiology*, 17(11), 4538–4546.

Geisen, S., Koller, R., Huenninghaus, M., Dumack, K., Urich, T., & Bonkowski, M. (2016). The soil food web revisited: Diverse and widespread mycophagous soil protists. *Soil Biology & Biochemistry*, 94, 10–18.

Geisen, S., Mitchell, E. A. D., Adl, S., Bonkowski, M., Dunthorn, M., Ekelund, F., ... Lara, E. (2018). Soil protists: a fertile frontier in soil biology research. *Fems Microbiology Reviews*, 42(3), 293–323.

Gilbert, D., Amblard, C., Bourdier, G., Francez, A. J., & Mitchell, E. A. D. (2000). Le régime alimentaire des Thécamoebiens (Protista, Sarcodina). *L´Annee Biologique*, (39), 57–68.

Giner, C. R., Forn, I., Romac, S., Logares, R., de Vargas, C., & Massana, R. (2016). Environmental sequencing provides reasonable estimates of the relative abundance of specific picoeukaryotes. *Applied and Environmental Microbiology*, 82, 4757–4766.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., ... Christen, R. (2013). The Protist Ribosomal Reference database (PR^2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, 41(1), D597–D604.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2-3), 147–186.

Guisan, A., & Rahbek, C. (2011). SESAM - a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography*, 38, 1433-1444.

Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). *Habitat suitability and distribution models*. Cambridge: Cambridge, University Press.

Harder, C. B., Ronn, R., Brejnrod, A., Bass, D., Abu Al-Soud, W., & Ekelund, F. (2016). Local diversity of heathland Cercozoa explored by in-depth sequencing. *Isme Journal*, 10(10), 2488–2497.

Hardham, A. R. (2005). *Phytophthora cinnamomi*. *Molecular Plant Pathology*, 6(6), 589–604.

Heger, T. J., Derungs, N., Theurillat, J. P., & Mitchell, E. A. D. (2016). Testate amoebae like it hot: species richness decreases along a subalpine-alpine altitudinal gradient in both natural *Calluna vulgaris* litter and transplanted *Minuartia sedoides* cushions. *Microbial Ecology*, 71, 725-734.

Hess, S., & Melkonian, M. (2014). Ultrastructure of the Algivorous Amoeboflagellate *Viridiraptor invadens* (Glissomonadida, Cercozoa). *Protist*, 165(5), 605–635.

Hollander, M., Wolfe, D. A., & Chicken, E. (2015). *Nonparametric statistical methods* (3rd ed.). Hoboken, NJ: Wiley.

Howe, A., Bass, D., Scoble, J., Lewis, R., Vickerman, K., Arndt, H., & Cavalier Smith, T. (2011) Novel cultured protists identify deep-branching environmental DNA clades of cercozoa: New Genera *Tremula*, *Micrometopion*, *Minimassisteria*, *Nudifila*, *Peregrinia*. *Protist*, 162, 332-72.

Huston, M. A. (1994). *Biological diversity: the coexistence of species on changing landscapes*. Cambridge: Cambridge University Press.

Jeger, M. J., & Pautasso, M. (2008). Comparative epidemiology of zoosporic plant pathogens. *European Journal of Plant Pathology*, 122(1), 111–126.

King, A. J., Freeman, K. R., McCormick, K. F., Lynch, R. C., Lozupone, C., Knight, R., & Schmidt, S. K. (2010). Biogeography and habitat modelling of high-alpine bacteria. *Nature Communications*, 1.

Kosakyan, A., Mulot, M., Mitchell, E. A. D., & Lara, E. (2015). Environmental DNA COI barcoding for quantitative analysis of protists communities: A test using the *Nebela collaris* complex (Amoebozoa; Arcellinida; Hyalospheniidae). *European Journal of Protistology*, 51(4), 311–320.

Langer, M. R., Weinmann, A. E., Loetters, S., Bernhard, J. M., & Roedder, D. (2013). Climate-driven range extension of *Amphistegina* (Protista, Foraminiferida): models of current and predicted future ranges. *Plos One*, 8(2).

Lara, E., & Acosta-Mercado, D. (2012). A molecular perspective on ciliates as soil bioindicators. *European Journal of Soil Biology*, 49(SI), 107–111.

Lara, E., & Belbahri, L. (2011). SSU rRNA reveals major trends in oomycete evolution. *Fungal Diversity*, 49(1), 93–100.

Lara, E., Berney, C., Ekelund, F., Harms, H., & Chatzinotas, A. (2007). Molecular comparison of cultivable protozoa from a pristine and a polycyclic aromatic hydrocarbon polluted site. *Soil Biology & Biochemistry*, 39(1), 139–148.

Lara, E., Roussel-Delif, L., Fournier, B., Wilkinson, D. M., & Mitchell, E. A. D. (2016). Soil microorganisms behave like macroscopic organisms: patterns in the global distribution of soil euglyphid testate amoebae. *Journal of Biogeography*, 43(3), 520-532

Lentendu, G., Mahé, F., Bass, D., Rueckert, S., Stoeck, T., Dunthorn, M. (2018). Consistent patterns of high alpha and low beta diversity in tropical parasitic and free-living protists. *Molecular Ecology*, 27(13), 2846-2875

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

Mahé, F., de Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., ... Dunthorn, M. (2017). Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nature Ecology and Evolution*, 1(4).

McCain, C. M. (2005). Elevational gradients in diversity of small mammals. *Ecology*, 86(2), 366–372.

Mitchell, E. A. D., Borcard, D., Buttler, A. J., Grosvernier, P., Gilbert, D., & Gobat, J. M. (2000). Horizontal distribution patterns of testate amoebae (Protozoa) in a *Sphagnum magellanicum* carpet. *Microbial Ecology*, 39(4), 290–300.

Mod, H. K., Scherrer, D., Luoto, M., & Guisan, A. (2016). What we use is not what we know: environmental predictors in plant distribution models. *Journal of Vegetation Science*, 27(6), 1308–1322.

Nesbitt, J. E., & Adl, S. M. (2014). Differences in soil quality indicators between organic and sustainably managed potato fields in Eastern Canada. *Ecological indicators*, 37(A), 119–130.

Neuhauser, S., Kirchmair, M., Bulman, S., & Bass, D. (2014). Cross-kingdom host shifts of phytomyxid parasites. *Bmc Evolutionary Biology*, 14.

Noyce, G. L., Fulthorpe, R., Gorgolewski, A., Hazlett, P., Honghi, T., & Basiliko, N. (2016). Soil microbial responses to wood ash addition and forest fire in managed Ontario forests. *Applied Soil Ecology*, 107, 368–380.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2018). vegan: Community Ecology Package (Version 2.5-2). Retreived from "https://CRAN.R-project.org/package=vegan".

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(11).

Pearson, W. R. (2000). Flexible Sequence Similarity Searching with the FASTA3 Program Package. In S. Misener & S. A. Krawetz (Eds.), *Bioinformatics Methods and Protocols* (pp. 185–218). New York, NY: Humana Press.

Pellissier, L., Niculita-Hirzel, H., Dubuis, A., Pagni, M., Guex, N., Ndiribe, C., ... Guisan, A. (2014). Soil fungal communities of grasslands are environmentally structured at a regional scale in the Alps. *Molecular Ecology*, 23(17), 4274–4290.

Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R., Martínez-Meyer, E., Nakamura, M., & Araújo, M. P. (2011). *Ecological niches and geographic distributions*. Princeton: Princeton University Press.

Pohlert, T. (2014). The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR) (Version 4.1). Retrieved from "http://CRAN.R-project.org/package=PMCMR".

Pushkareva, E., Johansen, J. R., & Elster, J. (2016). A review of the ecology, ecophysiology and biodiversity of microalgae in Arctic soil crusts. *Polar Biology*, 39(12), 2227–2240.

Reymond, A., Purcell, J., Cherix, D., Guisan, A., & Pellissier, L. (2013). Functional diversity decreases with temperature in high elevation ant fauna. *Ecological Entomology*, 38(4), 364–373.

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *Peerj*, 4.

Rose, J. M., & Caron, D. A. (2007). Does low temperature constrain the growth rates of heterotrophic protists? Evidence and implications for algal blooms in cold waters. *Limnology and Oceanography*, 52(2), 886–895.

Santoyo, G., Hernandez-Pacheco, C., Hernandez-Salmeron, J., & Hernandez-Leon, R. (2017). The role of abiotic factors modulating the plant-microbe-soil interactions: toward sustainable agriculture. A review. *Spanish Journal of Agricultural Research*, 15(1).

Schiaffino, M. R., Lara, E., Fernández, L. D., Balagué, V., Singer, D., Seppey, C. V. W., ... Izaguirre, I. (2016). Microbial eukaryote communities exhibit robust biogeographical patterns along a gradient of Patagonian and Antarctic lakes. *Environmental Microbiology*, 18, 5249-5264.

Schmidt, O., Dyckmans, J., & Schrader, S. (2016). Photoautotrophic microorganisms as a carbon source for temperate soil invertebrates. *Biology Letters*, 12(1).

Seppey, C. V. W., Singer, D., Dumack, K., Fournier, B., Belbahri, L., Mitchell, E. A. D., & Lara, E. (2017). Distribution patterns of soil microbial eukaryotes suggests widespread algivory by phagotrophic protists as an alternative pathway for nutrient cycling. *Soil Biology & Biochemistry*, 112, 68–76.

Shen, C. C., Liang, W. J., Shi, Y., Lin, X. G., Zhang, H. Y., Wu, X., … Chu, H. Y. (2014). Contrasting elevational diversity patterns between eukaryotic soil microbes and plants. *Ecology*, 95, 3190-3202.

Shen, C., Shi, Y., Fan, K., He, J. S., Adams, J. M., Ge, Y., & Chu, H. (2019). Soil pH dominates elevational diversity pattern for bacteria in high elevation alkaline soils on the Tibetan Plateau. Fems Microbiology Ecology, 95(2).

Singer, D., Lara, E., Steciow, M. M., Seppey, C. V. W., Paredes, N., Pillonel, A., ... Belbahri, L. (2016). High-throughput sequencing reveals diverse oomycete communities in oligotrophic peat bog micro-habitat. *Fungal Ecology*, 23, 42–47.

Singer, D., Kosakyan, A., Seppey, C. V. W., Pillonel, A., Fernandez, L. D., Fontaneto, D., ... Lara, E. (2018). Environmental filtering and phylogenetic clustering correlate with the distribution patterns of cryptic protist species. *Ecology*, 99(4), 904–914.

Spehn, M., & Körner, C. (2009). *Data mining for global trends in mountain biodiversity*. Boca Raton, FL: CRC Press.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H., & Richards, T. A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology*, 19(s1), 21–31.

Weber, A. A. T., & Pawlowski, J. (2013). Can abundance of protists be inferred from sequence data: a case study of Foraminifera. *Plos One*, 8(2).

Yashiro, E., Pinto-Figueroa, E., Buri, A., Spangenberg, J. E., Adatte, T., Niculita-Hirzel, H., ... van der Meer, J. R. (2016). Local environmental factors drive divergent grassland soil bacterial communities in the western Swiss Alps. *Applied and Environmental Microbiology*, 82(21), 6303–6316.

Yashiro, E., Pinto-Figueroa, E., Buri, A., Spangenberg, J. E., Adatte, T., Niculita-Hirzel, H., … van der Meer, J. R. (2018). Meta-scale mountain grassland observatories uncover commonalities as well as specific interactions among plant and non-rhizosphere soil bacterial communities. *Scientific Reports*, 8.

Zaric, S., Schulz, M., & Mulitza, S. (2006). Global prediction of planktic foraminiferal fluxes from hydrographic and productivity data. *Biogeosciences*, 3(2), 187–207.

Zhang, T., Jia, R., & Yu, L. Y. (2016). Diversity and distribution of soil fungal communities associated with biological soil crusts in the southeastern Tengger Desert (China) as revealed by 454 pyrosequencing. *Fungal Ecology*, 23, 156–163.

Zimmermann, N. E., & Kienast, F. (1999). Predictive mapping of alpine grasslands in Switzerland: Species versus community approach. *Journal of Vegetation Science*, 10(4), 469–482.

Zimmermann, N. E., Edwards, T. C., Moisen, G. G., Frescino, T. S., & Blackard, J. A. (2007). Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of Applied Ecology*, 44(5), 1057–1067.

Zinger, L., Shahnavaz, B., Baptist, F., Geremia, R. A., & Choler, P. (2009). Microbial diversity in alpine tundra soils correlates with snow cover dynamics. *Isme Journal*, 3(7), 850–859.

Figures

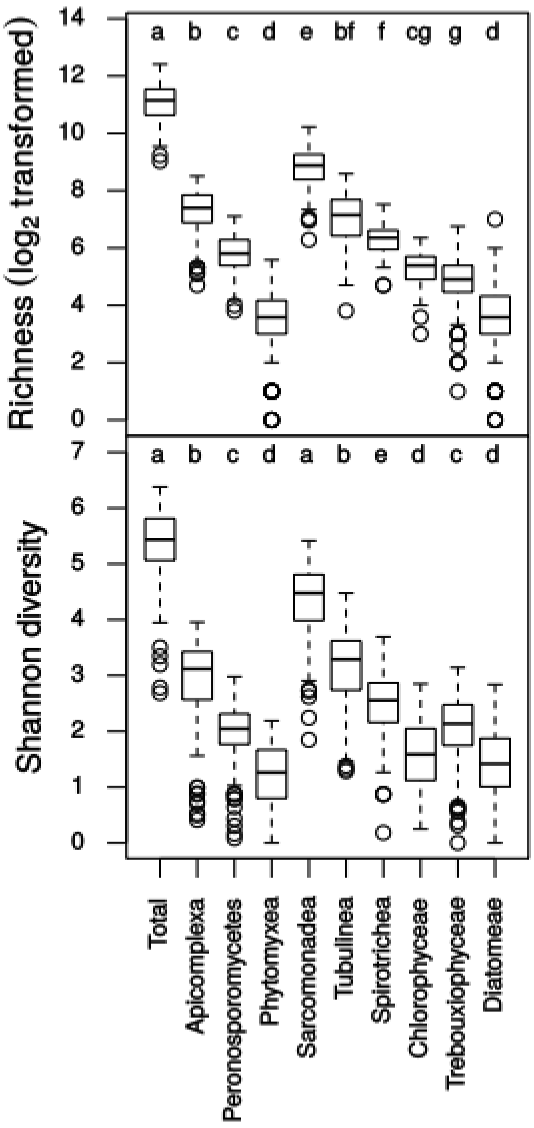

Figure 1: Shannon diversity and richness (log transformed) distributions of protist operational taxonomic units communities retrieved from 178 plots in the Swiss western Alps. The distributions are shown for the total community as well as for nine broad taxa. The letters above the boxplots represent groups according to a multiple comparison mean rank sums test (Nemenyi test $P < 0.05$).

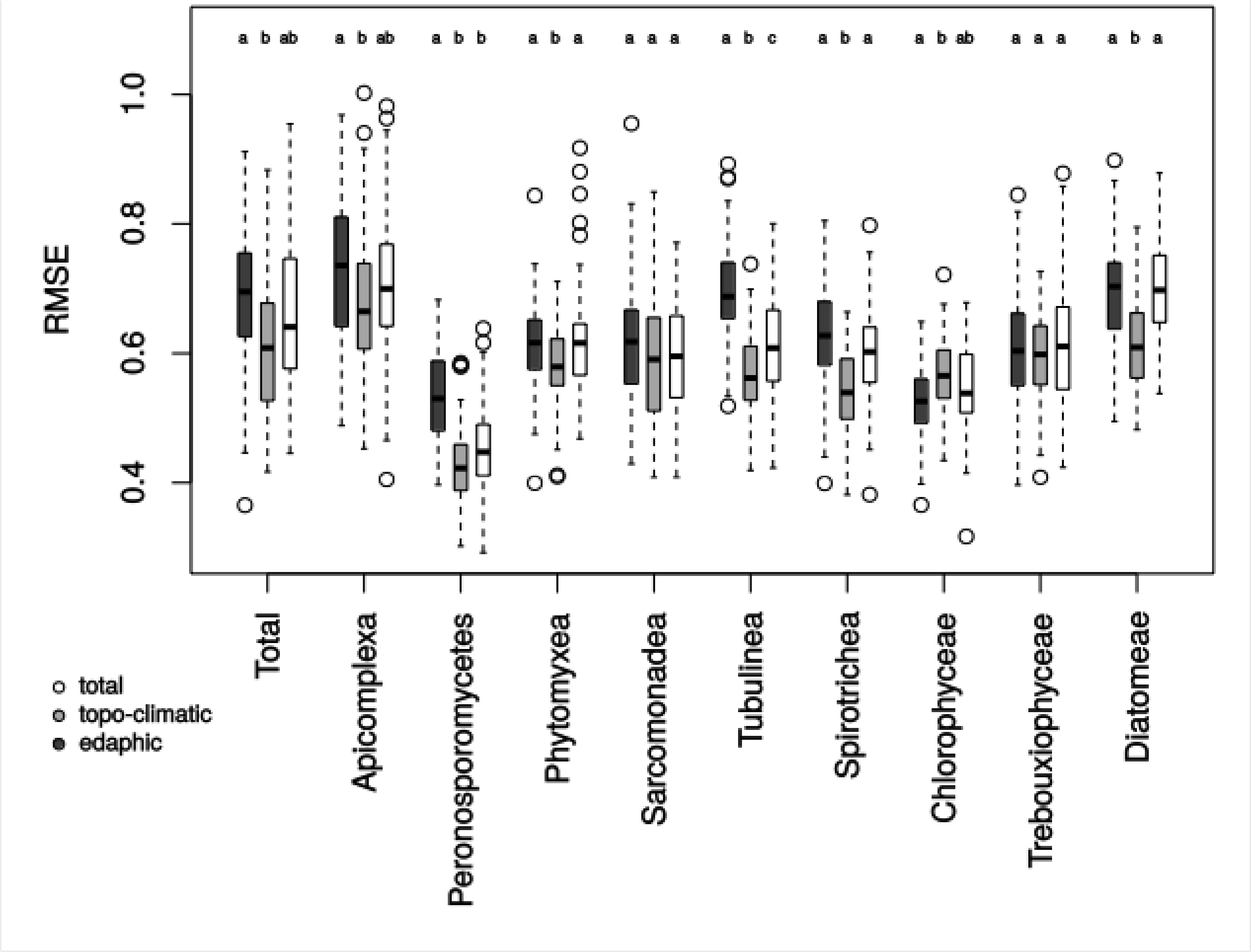

Figure 2: Predictive power (Root Mean Square Error: RMSE) of edaphic (dark grey), topo-climatic (pale gray) and overall (white) predictors calculated on the diversity of protist operational taxonomic units from the total community and nine broad taxa retrieved from 178 meadow soils in the Swiss western Alps. The RMSE were calculated on 100 cross validation of Generalized Additive Models performed with 20% of the samples as test dataset. The letters on the right of the barplot represent significantly different groups according to a multiple comparison mean rank sums test (Nemenyi test $P < 0.05$) for each of the total communities and nine broad taxa.

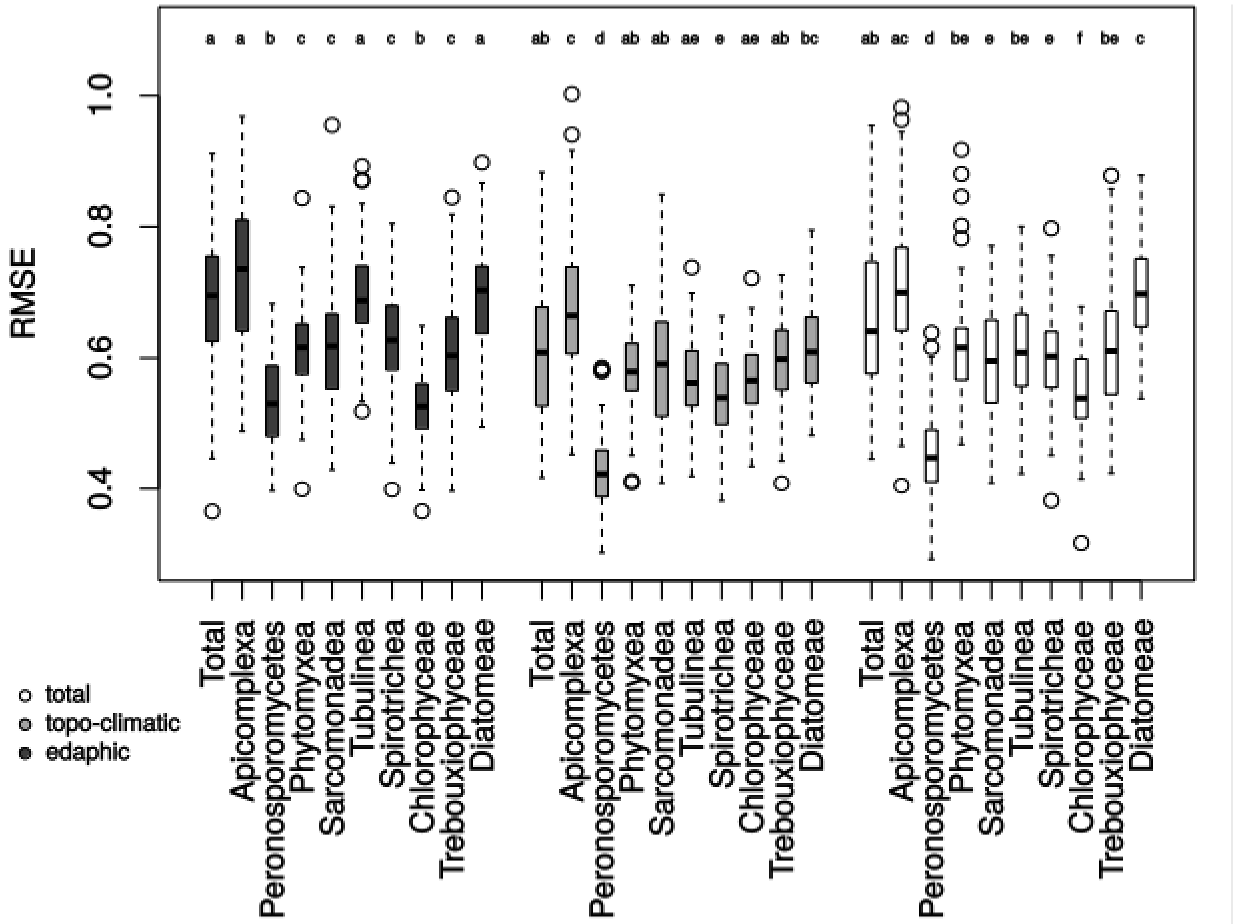

Figure 3: Predictive power (Root Mean Square Error: RMSE) of edaphic (dark grey), topo-climatic (pale gray) and overall (white) predictors calculated on the diversity of protist operational taxonomic units from the overall community and nine broad taxa retrieved from 178 meadow soils in the Swiss western Alps. The RMSE were calculated on 100 cross validation of Generalized Additive Models performed with 20% of the samples as test dataset. The letters on the top of the boxplots represent significantly different groups according to a multiple comparison mean rank sums test (Nemenyi test $P < 0.05$) for each of the edaphic, topo-climatic and overall variables.

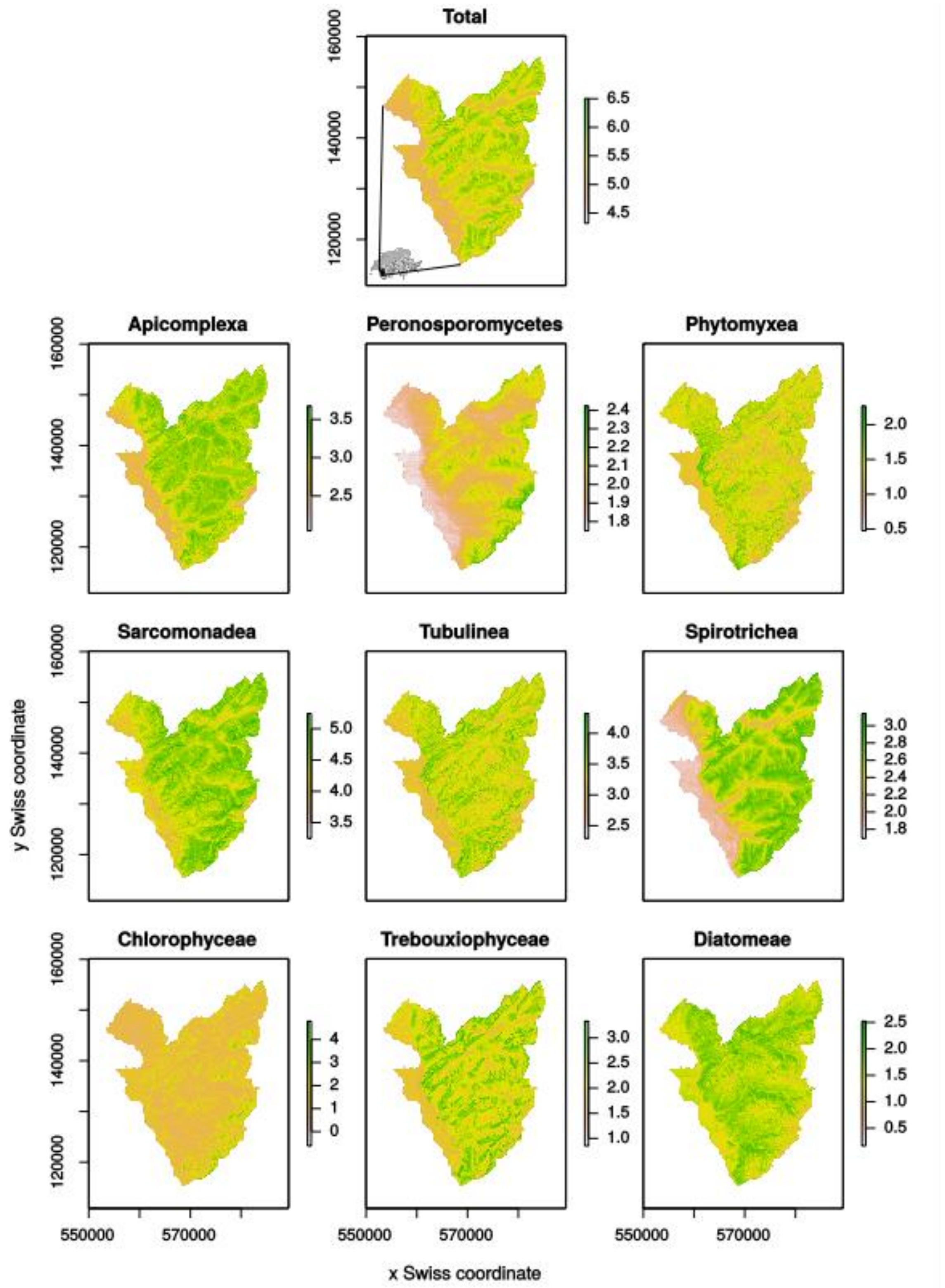

Figure 4: Diversity of the total protist community and nine broad taxa predicted from Generalized Additive Model through the Swiss western Alps based on the topography, slope southness, slope steepness and average temperature from June to August.

TABLE 1 Significance of edaphic (soil temperature: Soil temp, relative humidity: rh, pH, electroconductivity: EC, total phosphorus amount: P, carbon/nitrogen ratio: C/N, loss of ignition: LOI, shale percentage) and topo-climatic (topography: topo, slope southness: asp, slope steepness: slp, summer temperature average: tmean678) predictors on the diversity modelled (Generalized Additive Model) from overall protist community and nine broad taxonomic groups from operational taxonomic units gathered from 178 meadow soils in the Swiss western Alps.

	Edaphic								Topo-climatic				
	Soil_temp (°C)	Rh (1/10000)	pH	EC (µS/cm)	P(%)	C/N	LOI (%)	Shale (%)	topos	Asp (%)	Slp (%)	tmean678 (°C * 100)	adj_R2
Total		(+)	(−)										0.181
Apicomplexa		(+)						+−					0.269
Peronosporomycetes							+	(+)					0.113
Phytomyxea					(+)		(+−+)				++		0.164
Sarcomonadea			--						(−)	(−)	++		0.161
Tubulinea							-				+++		0.189
Spirotrichea			-			(−)						---	0.133
Chlorophyceae			+	+−	+								0.329
Trebouxiophyceae				(+−)		+-+					+	−	0.252
Diatomeae											(−)	(+)	0.057

Note: The + and − signs show if the diversity is positively or negatively associated to the predictor and the number of signs inform on the strength of the association (between parenthesis: $p < .1$, one sign: $p < .05$, two signs: $p < .01$, three signs: $p < .001$). The −+ and +− indicate minimum and maximum of diversity at mid-predictor value respectively. Details of the response of each taxonomic group to the different variables can be found in Appendix S10.10.

Appendix D

Holarctic phylogeography survey of *Hyalosphenia papilio*

Dispersal limitations and historical factors determine the biogeography of specialized terrestrial protists

David Singer[1,2], Edward A. D. Mitchell[1,3], Richard J. Payne[4], Quentin Blandenier[1,5], Clément Duckert[1], Leonardo D. Fernández[1,6], Bertrand Fournier[7], Cristián E. Hernández[8], Gustaf Granath[9], Håkan Rydin[9], Luca Bragazza[10,11,12], Natalia G. Koronatova[13], Irina Goia[14], Lorna I. Harris[15], Katarzyna Kajukało[16], Anush Kosakyan[17], Mariusz Lamentowicz[16], Natalia P. Kosykh[13], Kai Vellak[18], Enrique Lara[1,5]

[1]Laboratory of Soil Biodiversity, Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland
[2]Department of Zoology, Institute of Biosciences, University of São Paulo, São Paulo, Brazil
[3]Jardin Botanique de Neuchâtel, Neuchâtel, Switzerland
[4]Environment, University of York, York, UK
[5]Real Jardín Botánico, CSIC, Madrid, Spain
[6]Centro de Investigación en Recursos Naturales y Sustentabilidad (CIRENYS), Universidad Bernardo O'Higgins, Santiago, Chile
[7]Community and Quantitative Ecology Laboratory, Department of Biology, Concordia University, Montreal, QC, Canada
[8]Facultad de Ciencias Naturales y Oceanográficas, Departamento de Zoología, Universidad de Concepción, Barrio Universitario de Concepción, Chile
[9]Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden
[10]WSL Swiss Federal Institute for Forest, Snow and Landscape Research, Lausanne, Switzerland
[11]Laboratory of Ecological Systems (ECOS), Ecole Polytechnique Féderale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering
(ENAC), Lausanne, Switzerland
[12]Department of Life Science and Biotechnologies, University of Ferrara, Ferrara, Italy
[13]Laboratory of Biogeocenology, Institute of Soil Science and Agrochemistry, Siberian Branch of Russian Academy of Sciences, Novosibirsk, Russia
[14]Department of Taxonomy and Ecology, Faculty of Biology and Geology, Babeș-Bolyai University, Cluj-Napoca, Romania
[15]School of Geography and Earth Sciences, McMaster University, Hamilton, ON, Canada
[16]Laboratory of Wetland Ecology and Monitoring, Faculty of Geographical and Geological Sciences and Department of Biogeography and Paleoecology, Adam
Mickiewicz University, Poznań, Poland
[17]Institute of Parasitology, Biology Center, Czech Academy of Sciences, České Budějovice, Czech Republic
[18]Institute of Ecology and Earth Sciences, Natural History Museum, University of Tartu, Tartu, Estonia

Molecular Ecology, 2019, Vol 28, Issue 12

DOI:10.1111/mec.15117

Abstract. Recent studies show that soil eukaryotic diversity is immense and dominated by micro-organisms. However, it is unclear to what extent the processes that shape the distribution of diversity in plants and animals also apply to micro-organisms. Major diversification events in multicellular organisms have often been attributed to long-term climatic and geological processes, but the impact of such processes on protist diversity has received much less attention as their distribution has often been believed to be largely cosmopolitan. Here, we quantified phylogeographical patterns in *Hyalosphenia papilio*, a large testate amoeba restricted to Holarctic *Sphagnum*-dominated peatlands, to test if the current distribution of its genetic diversity can be explained by historical factors or by the current distribution of suitable habitats. Phylogenetic diversity was higher in Western North America, corresponding to the inferred geographical origin of the *H. papilio* complex, and was lower in Eurasia despite extensive suitable habitats. These results suggest that patterns of phylogenetic diversity and distribution can be explained by the history of Holarctic *Sphagnum* peatland range expansions and contractions in response to Quaternary glaciations that promoted cladogenetic range evolution, rather than the contemporary distribution of suitable habitats. Species distributions were positively correlated with climatic niche breadth, suggesting that climatic tolerance is key to dispersal ability in *H. papilio*. This implies that, at least for large and specialized terrestrial micro-organisms, propagule dispersal is slow enough that historical processes may contribute to their diversification and phylogeographical patterns and may partly explain their very high overall diversity.

Keywords. distribution; Holarctic; *Hyalosphenia papilio*; phylogeography; protists; *Sphagnum* peatland

Introduction

The question of whether the same rules structure the diversity of all eukaryotes, micro- and macroscopic alike, has been the subject of heated debate since the early 2000s. The classical paradigm is that "everything is everywhere, but, the environment selects" (Baas-Becking, 1934). Defenders of this paradigm have argued that geographical barriers are ineffective in preventing the dispersal of microbes (Fenchel, 2005; Finlay, 1998). Other researchers, while accepting that some microbes do indeed occur worldwide, have argued that others are clearly restricted to certain regions: the "moderate endemicity" model (Foissner, 1999). This argument is based particularly on a limited number of so-called "biogeographical flagship species," with conspicuous morphology.

The application of barcoding to protists has brought new nuance to the debate (Pawlowski et al., 2012). Single-cell DNA barcoding studies (Pawlowski et al., 2012) of individual "morphospecies" are now revealing the existence of numerous "cryptic" biological species (Singer et al., 2018). Barcoding studies have recently demonstrated geographically limited distributions in soil (Ryšánek, Hrčková, & Škaloud, 2015), freshwater (Škaloud et al., 2019) and marine organisms (Santoferrara, Rubin, & Mcmanus, 2018), although cases of cosmopolitan distribution have also been reported (Geisen, Fiore-Donno, Walochnik, & Bonkowski, 2014; Šlapeta, López-García, & Moreira, 2005). The development of microbial phylogeography, combining biogeography and molecular phylogeny, in turn has allowed the evaluation of possible drivers of diversity patterns, and comparison with those known to drive plant and animal diversity (Martiny et al., 2006).

Among terrestrial micro-organisms, testate amoebae are particularly useful models for phylogeographical studies. Testate amoebae are conspicuous and relatively easy to identify and are also large enough to be isolated individually for DNA barcoding. Many species have narrow ecological tolerances and thus can only colonize specific, often geographically discontinuous habitats (Singer et al., 2018). Furthermore, some species have well-documented distributions and ecology. A good example is *Hyalosphenia papilio*, a widely recorded and morphologically distinctive testate amoeba taxon (Figure 1). Of particular interest for phylogeographical studies is that, based on single-cell barcoding and the variable molecular marker cytochrome *c* oxidase subunit I (*COI*), *H. papilio* is known to represent a species complex of at least 12 lineages (Heger, Mitchell, & Leander, 2013).

Hyalosphenia papilio is found exclusively in Holarctic *Sphagnum*-dominated peatlands (Amesbury et al., 2018, 2016) and it is known to be absent from similar southern hemisphere sites despite extensive study

(Fernández, Lara, & Mitchell, 2015; Smith, Bobrov, & Lara, 2007). *Sphagnum*- dominated peatlands are comparatively young ecosystems, dating back to the expansion of boreal and subarctic environments near the Pliocene (Shaw et al., 2010). *Sphagnum* is an ecosystem engineer that modifies habitats by increasing soil wetness and decreasing pH and available nutrient content, producing decay-resistant litter rich in phenols and sphagnan (van Breemen, 1995) and hosting very distinctive prokaryotic, algal and fungal communities (Kostka et al., 2016; Mutinová, Neustupa, Bevilacqua, & Terlizzi, 2016). Thus, *Sphagnum* represents a highly selective habitat for macro- and micro-organisms. This explains why *Sphagnum*-dominated ecosystems are species-poor and these same factors are likely to also drive evolutionary adaptations in testate amoebae (Kosakyan et al., 2016; Singer et al., 2018). Hence, it is likely that this taxon does not pre-date the radiation of peat-forming *Sphagnum* species between 17 and 7 Ma (Shaw et al., 2010). Large extents of *Sphagnum* comparable to modern *Sphagnum*-dominated peatlands probably appeared during the late Miocene/early Pliocene, concomitantly with global cooling, that is, between 7 and 5.5 Ma (Herbert et al., 2016). While *Sphagnum* occurs at low as well as high latitudes it is only a dominant component of peatlands at higher latitudes (e.g., Tierra-del-Fuego and the boreal zone of the Holarctic). The taxonomic richness of the genus is low in the Southern Hemisphere high latitudes and high in the Northern Hemisphere high latitudes which correspond to its inferred origin (Shaw, Carter, Aguero, da Costa, & Crowl, 2019).

Holarctic *Sphagnum*-dominated peatlands have experienced considerable changes in their extent due to the repeated advances and retreats of ice sheets during the Quaternary. Many of the largest areas covered by peatlands today were under ice during the Last Glacial Maximum (e.g., Fennoscandia, boreal Canada), while peatlands may have persisted in others (e.g., Pacific coast of Canada) (Treat et al., 2019). These successive glacial expansions and contractions are known to have shaped genetic diversity in multicellular taxa (Schönswetter, Stehlik, Holderegger, & Tribsch, 2005), whose dispersal is assumed to be slow in contrast to eukaryotic micro-organisms (Bahram et al., 2016). If, like plants and animals, protist dispersal is relatively slow, the genetic structure of their populations will bear traces of such range expansions and contractions, and the origin of taxa can potentially be inferred and the timing of phylogenetic events estimated based on molecular clocks (Arbogast et al., 2006; Arbogast, Edwards, Wakeley, Beerli, & Slowinski, 2002). By contrast, fast dispersal in protists would blur any such signature, and taxonomic or phylogenetic diversity would tend to be distributed randomly and peak in areas with the largest extent of favourable habitats (Forest, Colville, & Cowling, 2018).

It follows that the phylogeographical pattern of a given taxon, here *H. papilio*, can be used to test two alternative hypotheses. (a) Dispersal is low and/or slow enough so that traces of glacial cycles are reflected in its extant diversity. The highest diversity, and the probable geographical origin of *H. papilio* would be expected to occur in refugia corresponding to the margins of ice sheets during the Last Glacial Maximum where *Sphagnum* peatlands could survive. (b) Dispersal is high and/or fast, and diversity would be expected to be maximal where the largest expanses of *Sphagnum* peatlands are found today (e.g., western Siberia). Empirical evidence demonstrates a relationship between testate amoeba shell size and geographical range (Wilkinson, 2001). Population genetics analyses (Lara, Heger, Scheihing, & Mitchell, 2011) and modelling (Wilkinson, Koumoutsaris, Mitchell, & Bey, 2011) show a decline in dispersal potential for testate amoebae and theoretical organisms of smaller sizes (*c.* 60 µm) than that reported for *H. papilio* (size range 90–175 µm). The ability of entering a dormant stage (cysts) which can withstand desiccation and other stresses is considered to be a key dispersal trait in protists (Geisen et al., 2018); however, such structures have never been reported in *H. papilio*. We therefore predict that the first hypothesis is more likely to be supported.

Material and methods

Data set preparation

We retrieved all 360 existing *COI* gene sequences of *Hyalosphenia papilio* from GenBank, together with information on the origin of the cells from four studies (Gomaa et al., 2014; Heger et al., 2013; Kosakyan et al., 2012; Oliverio, Lahr, Grant, & Katz, 2015). In addition, we isolated 57 single cells of *H. papilio* from *Sphagnum* samples collected at 13 new sites, targeting under-sampled regions to compile a global data set (Table S1). Briefly, the cells were washed three times in autoclaved distilled water before DNA extraction, which was performed using the guanidine thiocyanate-base protocol (Duckert et al., 2018). Amplifications of *COI* gene fragments were performed in two steps: a first polymerase chain reaction (PCR) was undertaken with the general *COI* primers LCO1490 and HCO2198 (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994), which was followed by a nested PCR, using *H. papilio*-specific primers HPcoiF and HPcoiR (Gomaa et al., 2014). The first DNA amplification profiles consisted of an initial denaturation step for 3 min at 95 C, followed by 39 cycles of 15 s of denaturation at 95 C, 15 s of annealing at 43 C and 1 min of elongation at 72 C with an additional final elongation step at 72 C for 10 min. The procedure for the second PCR profile was the same except that the annealing temperature was increased to 55 C. Sequencing was carried out using a BigDye197 Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) and analysed with an ABI-3130XL DNA sequencer (Applied Biosystems). Sequences were deposited in GenBank with the following accession numbers: MK823130–MK823186. *COI* sequences were edited and aligned (clustalw algorithm; Thompson, Gibson, & Higgins, 2003) using bioedit (version 7.2.3; Hall, 1999). The final data set including published and new sequences consisted of 418 sequences from 61 sites (Table S1).

Lineage delineation

We delimited genetic lineages following the approach described by Heger et al. (2013). Briefly, to obtain a general overview of the existing lineages, we first constructed the phylogenetic tree based on the matrix of the unique sequences (haplotypes) among the 418 considered in this study. The sequence lengths of the data set vary from 430 to 620 bp (depending on the primers used to barcode the isolated cells). We constructed both a maximum likelihood (ML) and a Bayesian tree, with the RAxML algorithm (Stamatakis, Hoover, Rougemont, & Renner, 2008) and mrbayes version 3.2.6 (Ronquist & Huelsenbeck, 2003), using in both cases a GTR + model. We then tested if the haplotypes were distributed randomly by comparing our observed distribution with a null model obtained by haplotypes randomly attributed to lineages (10,000 replicates). The tree root was placed between two major clades (clade I contains lineages L, K, M and J and clade II contains lineages C, DE, F, B, A, G, H and I) that appeared well supported in previous studies (Heger et al., 2013; Kosakyan et al., 2012). Bipartition support values were evaluated with 1,000 bootstrap replicates. Bayesian Markov chain Monte Carlo analysis was carried out with two simultaneous chains and 50,000,000 generations. Trees were sampled every 1,000th generation and the burn-in was set at 25%. The trees were rooted internally based on the topology of trees obtained in previous studies (Heger et al., 2013; Kosakyan et al., 2016), which showed two major clades with maximum support; we rooted the tree between these two clades. As both trees were congruent, we presented only the ML tree and used the Bayesian analysis to evaluate each node's posterior probability (pp). We used three independent methods of lineage delimitation to compare our assignments with those of Heger et al. (2013): (a) Automatic Barcode Gap Discovery (ABGD) (Fontaneto, Flot, & Tang, 2015; Puillandre, Lambert, Brouillet, & Achaz, 2011) using the ABGD web-server http://wwwabi.snv. jussi eu.fr/publi c/abgd/abgdw eb.html; (b) sequence divergences using the Kimura 2-parameter method (Kimura, 1980; Nassonova, Smirnov, Fahrni, & Pawlowski, 2010) using the "ape" package (version 3.2; Paradis, Claude, & Strimmer, 2004) in r (version 3.0.1; R Core Team, 2014); and (c) generalized mixed yule coalescent (GMYC) analysis performed with the splits package, version 1.0-19 (Fujisawa & Barraclough, 2013) coded in r, version 3.1.2 (R Core Team, 2014).

Haplotype and lineage network

Haplotypes (defined as genetic units separated by at least a single mutation) were assigned to the previously determined lineages. Haplotype networks were constructed using minimum spanning network analysis as implemented in the software popart (version 1.7; Leigh & Bryant, 2015 [Table S2]). Four main geographical zones (Eastern North America, Western North America, Europe and Asia) were defined to highlight the distribution of the haplotypes.

Historical biogeography

We first tested whether the observed patterns could be due to chance or to a sampling bias by calculating the observed beta diversity (following Legendre & De Cáceres, 2013). We then tested if the observed beta diversity was higher than expected by chance. We simulated beta diversity values under null expectations and compared them to observed beta diversity values to obtain *p*-values and standardized effect sizes (SESs). Simulated beta diversity values were calculated using the same approach (Legendre & De Cáceres, 2013) on a permuted site by species matrix. Permutations were conducted using the permatswap algorithm of the r-package "vegan" (Oksanen, Blanchet, & Kindt, 2015), which preserves column sums. This allows us to randomly attribute species to station while preserving species total abundance.

To determinate whether species distribution areas were correlated with ecological tolerance, we determined the climatic niche breadth for each species using the tolerance index (Dolédec, Chessel, & Gimaret-Carpentier, 2000), with the r package "ade4" (Dray & Dufour, 2007). This index estimates niche breadth based on environmental tolerance (i.e., climate) (Hurlbert, 1978; Thuiller, 2004) using the dispersion of geographical cells that contain the target species in the climatic multivariate space. Low values of the index suggest narrow tolerance while high values correspond to generalists. These indices were inferred based on geographical coordinates for each occurrence (Table S1) and interpolated climate data sets (Bioclim, 19 variables) that were generated at 2.5 arcmin resolution from meteorological data (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). We then estimated distribution areas of the species based on the number of plots in which a given species was observed (i.e., a rough estimate of the spatial range) and plotted it against the estimated climatic niche breadths.

To evaluate the evolutionary events that may explain the current distribution of *H. papilio* (e.g., dispersal, extinction, rangeswitching, sympatry, vicariance and founder effect), we estimated the ancestral distribution and the frequency of event counts in each of 1,000 biogeographic stochastic mapping (BSM) analyses, using the biogeobears package (Matzke, 2013) in r (version 3.0.1; R Core Team, 2014). biogeobears allows for the estimation of ancestral geographical ranges on dated phylogeny, comparing several models of range evolution. We used the DEC model (Ree, Smith, & Baker, 2008) with two free parameters: "d" (dispersal rate) and "e" (extinction rate), and a fixed cladogenetic model (cladogenetic event allowed: vicariance, sympatric-subset speciation and sympatric range-copying). We also used a DEC model with an extra parameter, "j," which represents the founder-event speciation, where the new species "jumps" to a range outside of the ancestral range (DEC + j model). The comparison of these two models was performed using Akaike's information criterion (AIC). The age of the nodes of the rescaled beast tree of *H. papilio* was estimated by constraining its root to 7 Ma, which corresponds to the documented origin of *Sphagnum* peatlands (Stenøien, Shaw, Shaw, Hassel, & Gunnarsson, 2010), as all known lineages of *H. papilio* are restricted to these environments. This type of approach has been used to identify the geographical origin of multicellular taxa of different ages, dispersion strategies and lifestyles, including hyacinthoid monocots (Ali, Yu, Pfosser, & Wetschnig, 2012), chameleons (Tolley, Townsend, & Vences, 2013) and bees (Trunz, Packer, Vieu, Arrigo, & Praz, 2016), and is used here, to our knowledge, for the first time in micro-organisms. This approach allowed us to test the two hypotheses: if the frequency of event counts in each of 1,000 BSM sustains a low frequency of dispersal events, related to other biogeographical events, the first hypothesis is supported ([a] dispersal is low and/or slow enough so that traces of glacial cycles are reflected in its extant diversity), and otherwise, the second hypothesis is supported ([b] dispersal is high and/or fast, and diversity would be expected to be maximal where the largest expanses of *Sphagnum* peatlands are found today).

Results

Lineage delineation and diversity

The 418 mitochondrial *COI* sequences of *Hyalosphenia papilio* revealed the existence of 13 or 14 distinct lineages (Figure 2). The Kimura 2-parameter test suggested the existence of 14 lineages based on a threshold of 1% sequence divergence. The GMYC method yielded 13 lineages (lower and upper confidence intervals: 10 and 29 lineages, respectively; $p = 0.046$) based on single threshold methods. Finally, the ABGD method identified 13 lineages, using a distinctive barcoding gap of 7%. One of these lineages, called here "M," has not been previously recorded. This lineage was recovered from localities not included in previous studies (Gomaa et al., 2014; Heger et al., 2013; Kosakyan et al., 2012; Oliverio et al., 2015). It was supported by all analyses, although the Kimura 2-parameter test suggested dividing it into two (Figure 2).

Phylogenetic reconstruction

Sequences from the previously overlooked lineage M diverged from all others (pp = 1) and branched as a sister group to lineages J and K (Figure 2). Only a single haplotype was retrieved from lineage E, and five from lineage D (sensu Heger et al., 2013). Here again the genetic divergence was low (i.e., at most six nucleotides difference between the sequence of lineage E, and the five sequences of lineage D, all of which were separated by a single nucleotide [Figure 3]).

Haplotype network

The haplotype network (Figure 3) showed that some lineages (B, H and L) were composed of only a single haplotype, whereas others included several haplotypes, independently of the number of individual cells barcoded. Some lineages were relatively rare (e.g., B, L and M with seven, two and seven individuals, respectively) whilst others were extensively recorded (e.g., lineage A was identified more than 100 times). Null model analyses show that such a pattern is not expected under random assembly of lineages (Figure S1).

Spatial patterns of phylogenetic richness

We found that the observed beta diversity was significantly higher than expected by chance (SES = 1.71; $p = 0.99$), showing a strong spatial structuring of diversity. We also found a strong positive correlation between niche breadth as estimated using Dolédec tolerance indices and distribution areas ($R2 = 0.75$; $p = 0.001$). We also found that lineages differed in their climatic niches with some lineages preferring colder and drier conditions (Figure S2, lineage H) and others preferring warmer conditions with abundant precipitation (Figure S2, lineage I).

The geographical distribution of phylogenetic richness showed a clear contrast (Figures 1 and 3). Only four lineages (A, C, J and G) were recovered from all of Eurasia, five from Eastern North America (A, F, K, J and M), and nine from Western North America (six in Alaska and five in the Pacific Northwest, only two being shared between these two regions). Thus, regional as well as overall diversity and diversity turnover were all higher in North America than in Eurasia.

The distribution of the different lineages (Figure 3) suggests that several haplotypes are specific to certain geographical areas (B, DE, H, I and L occur only in Western North America, while K and M occurred only in Eastern North America), whereas others were geographically widespread (e.g., J is found throughout the Holarctic realm). Null model analyses show that such a pattern is not expected under random assembly of lineages (Figure S1). This structure in lineage distribution suggests that geographical dispersal has occurred comparatively slowly, allowing it to be recovered with a genetic marker such as mitochondrial *COI* used for species-level delineation in this group of organisms (Kosakyan et al., 2012).

Origin of lineages and evaluation of the diversification processes

The AIC selection of biogeographical models implemented in biogeobears indicated that a DEC model was the best supported (Table S3). Based on this model, the most likely ancestral areas for *H. papilio* are in Western North America (Figure 4). The dispersal summary extracted from the 1,000 BSM maps showed that most of the dispersal events occurred from Western North America and Asia to the other biogeographical areas, and from Asia to Europe (Table S4). The results of the ancestral area estimation and number of dispersal events analyses showed that the most frequent process during the historical biogeography of *H. papilio* was narrow sympatry (i.e., when the ancestral range contains one area, and both daughter lineages inherit that area), followed by a low frequency of dispersal events (range expansion) (Figure S1). The importance of vicariance and founder events was comparatively limited (Figure S1).

Discussion

Diversity and geographical distribution of the lineages and haplotypes

The *Hyalosphenia papilio* species complex is represented by at least 13 lineages in the Holarctic region, one of which had not been previously detected. Although it is possible that some lineages remain to be discovered, our globally extensive sampling retrieved only one additional lineage (M), suggesting that we have now captured most of the group's diversity. The genetic distances determined by our taxon delineating approaches are consistent with the barcoding gap (<4%) used to discriminate species in other related testate amoebae lineages (e.g., genus *Nebela*, Hyalospheniidae). The above-mentioned lineages were defined as species under multiple and independent concepts, including ecological, morphological and evolutionary (Singer et al., 2018). This might imply that the lineages retrieved in the present study can all be considered as separate species (Kosakyan, Gomaa, Mitchell, Heger, & Lara, 2013; Kosakyan et al., 2012; Singer, Kosakyan, Pillonel, Mitchell, & Lara, 2015; Singer et al., 2018). The accuracy of a species tree built on a single locus may be still questioned, especially in the case of recent radiations, as the existence of several caveats (e.g., sequencing pseudogenes, ongoing hybridization processes) cannot be ruled out and may distort the tree's topology. In Amoebozoa, *COI* has been chosen as the most accurate marker notably because of its sensitivity and lack of intra-individual variability (Nassonova et al., 2010) and we therefore consider it reliable.

Lineages of *H. papilio* show different distribution patterns over the Holarctic realm. Four lineages (J, A, C, G) were found in several regions with contrasting climates (Figure 1) suggesting that they have a greater ecological tolerance. This is corroborated by the strong correlation between climatic niche breadth and estimated distribution ranges (Figure 5), suggesting that colonization capacity is constrained by specific tolerance to climates. If these distributional patterns reflect evolutionary adaptation to long-distance dispersal, it would then imply that the required physiological/ lifestyle adaptations to long-range migration have appeared independently at least four times in the history of the *H. papilio* species complex (Figure 4).

The existence of restricted distributions is even clearer at the haplotype level. Of the 74 total *H. papilio* haplotypes, only seven (9.2%) were present in two zones, two in three zones (2.6%) and no single haplotype was found in all four zones. This indicates that even widespread lineages (e.g., lineage J) show high infraspecific genetic structuring, which suggests limited gene flow among sites, and thus geographical isolation (Fernández, Hernández, Schiaffino, Izaguirre, & Lara, 2017; Lara et al., 2011).

Hypothesis 2, that diversity is maximal where the largest expanses of *Sphagnum* peatlands are found today, cannot be supported by our data and analyses. Under this hypothesis, highest diversity would be expected in regions such as Western Siberia where peatlands are at their most extensive and cover more than 20% of the landscape (Peregon, Maksyutov, Kosykh, & Mironycheva-Tokareva, 2008). However, only four “far travelled” lineages were found in all of Eurasia, as compared to 13 in North America. This is despite a larger overall area, greater extent of *Sphagnum* peatland and an extensive range of climatic conditions. Six lineages, 50% more than the entire of Eurasia, were found only in Alaska. Moreover, most genetic diversity seems to be located along the Western North American coast, a region where peatlands are typically small and scattered today. This fact, together with the strong spatial patterns in lineage distribution observed (Figures 1 and 4), advocates against our hypothesis 2 (fast dispersal).

Geographical origin and influence of historical events

All Eurasian lineages identified were also present in North America, while several lineages were restricted to North America. This observation alone suggests an American origin for *H. papilio*. Our ancestral range reconstruction corroborates this inference, placing the most probable origin of the *H. papilio* complex in

Western North America (Figure 4).

Dating speciation events is difficult in testate amoebae as their lineages cannot be morphologically distinguished (Mulot et al., 2017); testate amoeba shell records in peat are rare before the Holocene. Nevertheless, it is still possible to infer a time window for the radiation of the lineages indirectly based on the very strict habitat specificity of this taxon. All lineages of *H. papilio* thus far identified are restricted to *Sphagnum* peatlands. It is therefore reasonable to assume that this highly adapted taxon evolved within these ecosystems. The oldest fossils of the genus *Hyalosphenia* were described from the Triassic (*H. baueri*, 220 Ma) (Schönborn, Dörfelt, Foissner, Krienitz, & Schäffer, 1999). *H. baueri* shares some traits like an "indistinctly vase shape" and the presence of an organic lip surrounding the aperture with *H. papilio*. However, it is far from clear that the two taxa are directly related. First, it has been shown that the genus *Hyalosphenia* is paraphyletic, as *H. papilio* and *H. elegans* are only distantly related (Lahr et al., 2019; Lara, Heger, Ekelund, Lamentowicz, & Mitchell, 2008). Furthermore, a rough calculation can rule out the possibility of a very old age for *H. papilio*. "Standard" *COI* (estimated for animals) mutation rates are typically in the range of a few per cent per million years (Ho & Lo, 2013; Papadopoulou, Anastasiou, & Vogler, 2010), sometimes much higher (Ney, Frederick, & Schul, 2018). The most divergent *H. papilio* sequences are separated by roughly 10%, which implies that, in order for the deepest branching in the complex (see Figure 4) to be 100 million years old, the mutation rate would need to be of 0.01% Ma–1. This is far below all rates known to date, and even lower than the mutation rate of cnidarians, which are known for their extremely slow evolving mitochondria (Park et al., 2012). By contrast, to obtain an age of 7 million years the mutation rate would need to be 1.3% Ma–1, which is similar to the mutation rate of many arthropods and thus more parsimonious than the alternative.

During the Pleistocene, large areas of North America were intermittently covered by ice although ice- free refugia remained. The area of *Sphagnum* peatlands probably repeatedly expanded during interglacial periods and contracted in response to glacial periods when ice masses covered most of the landscape (Shaw et al., 2014). This period is also contemporary with most cladogenesis in the *H. papilio* phylogenetic history, which suggests a series of speciation events by cladogenetic range evolution that may have occurred during interglacial periods (Figure 4). Indeed, our analyses show that at least eight out of 12 cladogenesis events occurred during the Pleistocene, immediately after the 2.5 Ma boundary (Figure 4).

This hypothesis is also in line with the fact that the biogeobears analyses designated narrow sympatry or the inheritance of the ancestral area of a range by both daughter lineages, as a key process explaining the distribution of *H. papilio* lineages. At the onset of Quaternary glaciations (2.58 Ma), one lineage probably existed in Eastern and two in Western North America (Figure 4). While the first lineage probably survived south of the ice sheet, where conditions were wet enough to allow the development of peatlands (Shaw et al., 2010), the two others were probably confined to refugia in Western North America.

The location of these refugia is known to have shaped the distribution of plants (Eidesen et al., 2013) and animals (Klütsch, Manseau, Anderson, Sinkins, & Wilson, 2017). In particular, Eastern Beringia (today Alaska and Yukon Territory) was wet enough to support the growth of *Sphagnum* mosses and *Sphagnum* peatlands (Shaw et al., 2013, 2014). These peatlands allowed the survival of associated organisms, probably including the lineages of *H. papilio*. In contrast, Western Beringia (today far eastern Russia) was too dry to support large expanses of *Sphagnum* peatlands (Shaw et al., 2013, 2014) and probably constituted a barrier for the migration of *H. papilio* westwards. Our data suggest that colonization of the Palaearctic region occurred recently, possibly after the last glaciation (Figure 1). Western Siberia, which was a cold desert during the Last Glacial Maximum, became covered with peatlands after 11,000 years ago (Velichko, Timireva, Kremenetski, MacDonald, & Smith, 2011) and could have constituted a bridge that facilitated the invasion of the Western Palaearctic by "far travelled" lineages of *H. papilio*. Interestingly, a similar pattern has been suggested for *Sphagnum angermanicum* (Stenøien et al., 2010).

The present-day distribution of lineages and the local palaeogeographical context designates Eastern Beringia or the Pacific Coast as the most probable origin for all extant *H. papilio* lineages. The higher diversity of *H. papilio* haplotypes in North America as compared to Europe mirrors the higher diversity of vascular plants (Earl Latham & Ricklefs, 1993; Svenning, 2003), and both were probably similarly driven by glaciations. The phylogeographical history of *H. papilio*, used here as a convenient model taxon for protists lacking specialized morphological adaptation for dispersal, thus highlights the importance of historical processes in explaining the distribution of extant microbial diversity.

Therefore, following a dispersal event, sympatric diversification could indeed have played a major role in shaping the current phylogeography of *H. papilio* (Figure S1). It remains to be determined if *H. papilio* is representative of free-living micro-organisms in general. *H. papilio* is large by microbial standards; testate amoebae mostly range between 20 and 200 μm and many other protists and most fungi and prokaryotes are smaller. *H. papilio* is also restricted to *Sphagnum* mosses, which, although widespread across the Holarctic, nevertheless constitute a very specific habitat. More generalist, smaller species and/or species possessing structures adapted to dispersal (e.g., fruiting bodies as in many other Amoebozoa; Shadwick, Spiegel, Shadwick, Brown, & Silberman, 2009) may show patterns which agree better with the second hypothesis. Elucidating the historical processes shaping the diversity of protists with different dispersal strategies, and comparing patterns with better known macroscopic organisms, will open the way to understanding the processes of diversification that produced the immense diversity existing today.

Acknowledgements

This work was funded by the Swiss NSF (310003A_143960 and 31003A_163254), the intramural project 201730E063 (CSIC) and the project PGC2018-094660-BI00 from the Spanish Government to E.L., the Swiss NSF (P2NEP3_178543) to D.S., the Swedish Research Council (VR) (2015-05174) to G.G. and H.R., FONDECYT (11170927) and UBO/VRIP (170201) to L.D.F., FONDECYT (1170815) to C.E.H, the Russian Foundation for Basic Research, N 16-55-16007, and the Ministry of Science and Higher Education of the Russian Federation (according to the state assignment of ISSA SB RAS) to N.G.K. and N.P.K. We thank Indrek Hiiesalu for help during fieldwork, Boris Droz and Christopher Niewoehner (University of North California) and Matt McGlone (Manaaki Whenua/Landcare Research, Lincoln, NZ) as well as Diego Fontaneto (National Research Council of Italy, Water Research Institute, Verbania Pallanza, Italy) for useful comments on the manuscript.

References

Ali, S. S., Yu, Y., Pfosser, M., & Wetschnig, W. (2012). Inferences of biogeographical histories within subfamily Hyacinthoideae using SDIVA and Bayesian binary MCMC analysis implemented in RASP (Reconstruct Ancestral State in Phylogenies). *Annals of Botany*, *109*(1), 95–107. https://doi.org/10.1093/aob/mcr274

Amesbury, M. J., Booth, R. K., Roland, T. P., Bunbury, J., Clifford, M. J., Charman, D. J., … Warner, B. G. (2018). Towards a Holarctic synthesis of peatland testate amoeba ecology: Development of a new continental- scale palaeohydrological transfer function for North America and comparison to European data. *Quaternary Science Reviews*, *201*, 483–500. https ://doi.org/10.1016/j.quasc irev.2018.10.034

Amesbury, M. J., Swindles, G. T., Bobrov, A., Charman, D. J., Holden, J., Lamentowicz, M., … Warner, B. G. (2016). Development of a new pan-European testate amoeba transfer function for reconstructing peatland palaeohydrology. *Quaternary Science Reviews*, *152*, 132– 151. https://doi.org/10.1016/j.quasc irev.2016.09.024

Arbogast, B. S., Drovetski, S. V., Curry, R. L., Boag, P. T., Seutin, G., Grant, P. R., … Anderson, D. J. (2006). The origin and diversification of Galapagos mockingbirds. *Evolution*, *60*(2), 370–382. https ://doi. org/10.1111/j.0014-3820.2006.tb011 13.x

Arbogast, B. S., Edwards, S. V., Wakeley, J., Beerli, P., & Slowinski, J. B. (2002). Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics*, *33*(1), 707–740. https ://doi.org/10.1146/annur ev.ecols ys.33.010802.150500

Baas-Becking, L. G. M. (1934). *Geobiologie; of inleiding tot de milieukunde*. Den Haag, the Netherlands: WP Van Stockum & Zoon NV.

Bahram, M., Kohout, P., Anslan, S., Harend, H., Abarenkov, K., & Tedersoo, L. (2016). Stochastic distribution of small soil eukaryotes resulting from high dispersal and drift in a local environment.

The ISME Journal, *10*(4), 885. https ://doi.org/10.1038/ismej.2015.164

Dolédec, S., Chessel, D., & Gimaret-Carpentier, C. (2000). Niche separation in community analysis: A new method. *Ecology*, *81*(10), 2914–2927. https ://doi.org/10.1890/0012-9658(2000) 081[2914:NSICAA]2.0.CO;2

Dray, S., & Dufour, A.-B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1–20.

Duckert, C., Blandenier, Q., Kupferschmid, F. A. L., Kosakyan, A., Mitchell, E. A. D., Lara, E., & Singer, D. (2018). En garde! Redefinition of *Nebela militaris* (Arcellinida, Hyalospheniidae) and erection of *Alabasta* gen. nov. *European Journal of Protistology*, *66*, 156–165. https ://doi. org/10.1016/j.ejop.2018.08.005

Earl Latham, R., & Ricklefs, R. E. (1993). Global patterns of tree species richness in moist forests: Energy-diversity theory does not account for variation in species richness. *Oikos*, *67*(2), 325– 333. https ://doi. org/10.2307/3545479

Eidesen, P. B., Ehrich, D., Bakkestuen, V., Alsos, I. G., Gilg, O., Taberlet, P., & Brochmann, C. (2013). Genetic roadmap of the Arctic: Plant dispersal highways, traffic barriers and capitals of diversity. *New Phytologist*, *200*(3), 898–910. https ://doi.org/10.1111/nph.12412

Fenchel, T. (2005). Cosmopolitan microbes and their 'cryptic'species. *Aquatic Microbial Ecology*, *41*(1), 49–54. https ://doi.org/10.3354/ ame04 1049

Fernández, L. D., Hernández, C. E., Schiaffino, M. R., Izaguirre, I., & Lara, E. (2017). Geographical distance and local environmental conditions drive the genetic population structure of a freshwater microalga (Bathycoccaceae; Chlorophyta) in Patagonian lakes. *FEMS Microbiology Ecology*, *93*(10). https ://doi.org/10.1093/femse c/fix125

Fernández, L. D., Lara, E., & Mitchell, E. A. D. (2015). Checklist, diversity and distribution of testate amoebae in Chile. *European Journal of Protistology*, *51*(5), 409–424. https ://doi.org/10.1016/j. ejop.2015.07.001

Finlay, B. J. (1998). The global diversity of protozoa and other small species. *International Journal for Parasitology*, *28*(1), 29–48. https ://doi. org/10.1016/S0020-7519(97)00167-7

Foissner, W. (1999). Protist diversity: Estimates of the near-imponderable. *Protist*, *150*(4), 363–368. https ://doi.org/10.1016/ S1434-4610(99)70037-4

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit 1 from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, (5). Retrieved from http://www.vliz.be/en/imis?refxml:id=64543

Fontaneto, D., Flot, J.-F., & Tang, C. Q. (2015). Guidelines for DNA taxonomy, with a focus on the meiofauna. *Marine Biodiversity*, *45*(3), 433–451. https ://doi.org/10.1007/s12526-015-0319-7

Forest, F., Colville, J. F., & Cowling, R. M. (2018). Evolutionary diversity patterns in the Cape flora of South Africa. In R. Scherson & D. Faith (Eds.), *Phylogenetic diversity*. Cham, Switzerland: Springer.

Fujisawa, T., & Barraclough, T. G. (2013). Delimiting species using single-locus data and the generalized mixed yule coalescent approach: A revised method and evaluation on simulated data sets. *Systematic Biology*, *62*(5), 707–724. https ://doi.org/10.1093/sysbi o/syt033

Geisen, S., Fiore-Donno, A. M., Walochnik, J., & Bonkowski, M. (2014). *Acanthamoeba* everywhere: High diversity of *Acanthamoeba* in soils. *Parasitology Research*, *113*(9), 3151–3158. https://doi.org/10.1007/s00436-014-3976-8

Geisen, S., Mitchell, E. A. D., Adl, S., Bonkowski, M., Dunthorn, M., Ekelund, F., … Lara, E. (2018). Soil protists: A fertile frontier in soil biology research. *FEMS Microbiology Reviews*, *42*(3), 293–323. https://doi.org/10.1093/femsr e/fuy006

Gomaa, F., Kosakyan, A., Heger, T. J., Corsaro, D., Mitchell, E. A. D., & Lara, E. (2014). One alga to rule them all: Unrelated mixotrophic testate amoebae (Amoebozoa, Rhizaria and Stramenopiles) share the same symbiont (Trebouxiophyceae). *Protist*, *165*(2), 161–176. https://doi.org/10.1016/j.protis.2014.01.002

Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, *41*, 95–98.

Heger, T. J., Mitchell, E. A. D., & Leander, B. S. (2013). Holarctic phylogeography of the testate amoeba *Hyalosphenia papilio* (Amoebozoa: Arcellinida) reveals extensive genetic diversity explained more by environment than dispersal limitation. *Molecular Ecology*, *22*(20), 5172– 5184. https ://doi.org/10.1111/mec.12449

Herbert, T. D., Lawrence, K. T., Tzanova, A., Cleaveland Peterson, L., Caballero-Gill, R., & Kelly, C. S. (2016). Late Miocene global cooling and the rise of modern ecosystems. *Nature Geoscience*, *9*(11), 843–847. https ://doi.org/10.1038/ngeo2813

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276

Ho, S. Y. W., & Lo, N. (2013). The insect molecular clock. *Australian Journal of Entomology*, *52*(2), 101–105. https ://doi.org/10.1111/aen.12018

Hurlbert, S. H. (1978). The measurement of niche overlap and some relatives. *Ecology*, *59*(1), 67–77. https ://doi.org/10.2307/1936632

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*(2), 111–120. https ://doi.org/10.1007/BF017 31581

Klütsch, C. F. C., Manseau, M., Anderson, M., Sinkins, P., & Wilson, P. J. (2017). Evolutionary reconstruction supports the presence of a Pleistocene Arctic refugium for a large mammal species. *Journal of Biogeography*, *44*(12), 2729–2739. https ://doi.org/10.1111/jbi.13090

Kosakyan, A., Gomaa, F., Mitchell, E. A. D., Heger, T. J., & Lara, E. (2013). Using DNA-barcoding for sorting out protist species complexes: A case study of the *Nebela tincta–collaris–bohemica* group (Amoebozoa; Arcellinida, Hyalospheniidae). *European Journal of Protistology*, *49*(2), 222–237. https ://doi.org/10.1016/j.ejop.2012.08.006

Kosakyan, A., Heger, T. J., Leander, B. S., Todorov, M., Mitchell, E. A. D., & Lara, E. (2012). COI barcoding of Nebelid testate amoebae (Amoebozoa: Arcellinida): Extensive cryptic diversity and redefinition of the Hyalospheniidae Schultze. *Protist*, *163*(3), 415–434. https://doi.org/10.1016/j.protis.2011.10.003

Kosakyan, A., Lahr, D. J. G., Mulot, M., Meisterfeld, R., Mitchell, E. A. D., & Lara, E. (2016). Phylogenetic reconstruction based on COI reshuffles the taxonomy of hyalosphenid shelled (testate) amoebae and reveals the convoluted evolution of shell plate shapes. *Cladistics*, *32*(6), 606–623. https ://doi.org/10.1111/cla.12167

Kostka, J. E., Weston, D. J., Glass, J. B., Lilleskov, E. A., Shaw, A. J., & Turetsky, M. R. (2016). The *Sphagnum* microbiome: New insights from an ancient plant lineage. *New Phytologist*, *211*(1), 57–64.

Lahr, D. J. G., Kosakyan, A., Lara, E., Mitchell, E. A. D., Morais, L., Porfirio-Sousa, A. L., … Brown, M. W. (2019). Phylogenomics and morphological reconstruction of Arcellinida testate amoebae highlight diversity of microbial eukaryotes in the Neoproterozoic. *Current Biology*, *29*(6), 991– 1001.e3. https ://doi.org/10.1016/j.cub.2019.01.078

Lara, E., Heger, T. J., Ekelund, F., Lamentowicz, M., & Mitchell, E. A. D. (2008). Ribosomal RNA genes challenge the monophyly of the Hyalospheniidae (Amoebozoa: Arcellinida). *Protist*, *159*(2), 165–176. https ://doi.org/10.1016/j.protis.2007.09.003

Lara, E., Heger, T. J., Scheihing, R., & Mitchell, E. A. D. (2011). COI gene and ecological data suggest size-dependent high dispersal and low intra-specific diversity in free-living terrestrial protists (Euglyphida: *Assulina*). *Journal of Biogeography*, *38*(4), 640–650. https ://doi.org/10.1111/j.1365-2699.2010.02426.x

Legendre, P., & De Cáceres, M. (2013). Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. *Ecology Letters*, *16*(8), 951–963. https ://doi.org/10.1111/ele.12141

Leigh, J. W., & Bryant, D. (2015). popart: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116. https ://doi.org/10.1111/2041- 210X.12410

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., … Staley, J. T. (2006). Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology*, *4*(2), 102. https ://doi.org/10.1038/nrmic ro1341

Matzke, N. J. (2013). *Probabilistic historical niogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing*. University of California Berkeley. Retrieved from https ://escho larsh ip.org/uc/item/8227p52c

Mulot, M., Marcisz, K., Grandgirard, L., Lara, E., Kosakyan, A., Robroek, B. J. M., … Mitchell, E. A. D. (2017). Genetic determinism vs. phenotypic plasticity in protist morphology. *Journal of Eukaryotic Microbiology*, *64*(6), 729–739. https ://doi.org/10.1111/jeu.12406

Mutinová, P. T., Neustupa, J., Bevilacqua, S., & Terlizzi, A. (2016). Host specificity of epiphytic diatom (Bacillariophyceae) and desmid (Desmidiales) communities. *Aquatic Ecology*, *50*(4), 697–709. https ://doi.org/10.1007/s10452-016-9587-y

Nassonova, E., Smirnov, A., Fahrni, J., & Pawlowski, J. (2010). Barcoding amoebae: Comparison of SSU, ITS and COI genes as tools for molecular identification of naked lobose amoebae. *Protist*, *161*(1), 102–115. https ://doi.org/10.1016/j.protis.2009.07.003

Ney, G., Frederick, K., & Schul, J. (2018). A post-Pleistocene calibrated mutation rate from insect museum specimens. *PLoS Currents*, *10*. https ://doi.org/10.1371/curre nts.tol.aba55 7de56 be881 79326 1f7e1 565cf35

Oksanen, J., Blanchet, F. G., & Kindt, R. (2015). *Vegan: Commity ecology package. R package version 2.3-0.*

http://cran.r-project.org

Oliverio, A. M., Lahr, D. J. G., Grant, J., & Katz, L. A. (2015). Are microbes fundamentally different than macroorganisms? Convergence and a possible case for neutral phenotypic evolution in testate amoeba (Amoebozoa: Arcellinida). *Royal Society Open Science*, *2*(12), 150414. https ://doi.org/10.1098/rsos.150414

Papadopoulou, A., Anastasiou, I., & Vogler, A. P. (2010). Revisiting the insect mitochondrial molecular clock: The mid-Aegean trench calibration. *Molecular Biology and Evolution*, *27*(7), 1659–1672. https ://doi.org/10.1093/molbe v/msq051

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, *20*(2), 289–290. https ://doi.org/10.1093/bioin forma tics/btg412

Park, E., Hwang, D.-S., Lee, J.-S., Song, J.-I., Seo, T.-K., & Won, Y.-J. (2012). Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. *Molecular Phylogenetics and Evolution*, *62* (1), 329–345. https ://doi.org/10.1016/j.ympev.2011.10.008

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*(11), e1001419. https ://doi.org/10.1371/journal.pbio.1001419

Peregon, A., Maksyutov, S., Kosykh, N. P., & Mironycheva-Tokareva, N. P. (2008). Map-based inventory of wetland biomass and net primary production in western Siberia. *Journal of Geophysical Research: Biogeosciences*, *113*(G1). https ://doi.org/10.1029/2007J G000441

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2011). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, *21*(8), 1864–1877. https ://doi.org/10.1111/ j.1365-294X.2011.05239.x

R Core Team (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ree, R. H., Smith, S. A., & Baker, A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, *57*(1), 4–14. https ://doi.org/10.1080/10635 15070 1883881

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*(12), 1572–1574. https ://doi.org/10.1093/bioin forma tics/btg180

Ryšánek, D., Hr ková, K., & Škaloud, P. (2015). Global ubiquity and local endemism of free-living terrestrial protists: Phylogeographic assessment of the streptophyte alga *Klebsormidium*. *Environmental Microbiology*, *17*(3), 689–698.

Santoferrara, L. F., Rubin, E., & Mcmanus, G. B. (2018). Global and local DNA (meta) barcoding reveal new biogeography patterns in tintinnid ciliates. *Journal of Plankton Research*, *40*(3), 209–221. https ://doi.org/10.1093/plank t/fby011

Schönborn, W., Dörfelt, H., Foissner, W., Krienitz, L., & Schäffer, U. (1999). A fossilized microcenosis in Triassic amber. *Journal of Eukaryotic Microbiology*, *46*(6), 571–584. https://doi.org/10.1111/ j.1550-7408.1999.tb051 33.x

Schönswetter, P., Stehlik, I., Holderegger, R., & Tribsch, A. (2005). Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, *14*(11), 3547–3555. https://doi.org/10.1111/j.1365-294X.2005.02683.x

Shadwick, L. L., Spiegel, F. W., Shadwick, J. D. L., Brown, M. W., & Silberman, J. D. (2009). *Eumycetozoa* = *Amoebozoa*?: SSUrDNA phylogeny of protosteloid slime molds and its significance for the amoebozoans supergroup. *PLoS ONE*, *4*(8), e6754.

Shaw, A. J., Carter, B. E., Aguero, B., da Costa, D. P., & Crowl, A. A. (2019). Range change evolution of peat mosses (*Sphagnum*) within and between climate zones. *Global Change Biology*, *25*(1), 108–120.

Shaw, A. J., Devos, N., Cox, C. J., Boles, S. B., Shaw, B., Buchanan, A. M., … Seppelt, R. (2010). Peatmoss (*Sphagnum*) diversification associated with Miocene Northern Hemisphere climatic cooling? *Molecular Phylogenetics and Evolution*, *55*(3), 1139–1145. https://doi.org/10.1016/j.ympev.2010.01.020

Shaw, A. J., Golinski, G. K., Clark, E. G., Shaw, B., Stenøien, H. K., & Flatberg, K. I. (2013). Intercontinental genetic structure in the amphi-Pacific peatmoss *Sphagnum miyabeanum* (Bryophyta: Sphagnaceae). *Biological Journal of the Linnean Society*, *111*(1), 17–37. https://doi.org/10.1111/bij.12200

Shaw, A. J., Shaw, B., Stenøien, H. K., Golinski, G. K., Hassel, K., & Flatberg, K. I. (2014). Pleistocene survival, regional genetic structure and interspecific gene flow among three northern peatmosses: *Sphagnum inexspectatum, S. orientale and S. miyabeanum*. *Journal of Biogeography*, *42*(2), 364–376. https ://doi.org/10.1111/jbi.12399

Singer, D., Kosakyan, A., Pillonel, A., Mitchell, E. A. D., & Lara, E. (2015). Eight species in the *Nebela collaris* complex: *Nebela gimlii* (Arcellinida, Hyalospheniidae), a new species described from a Swiss raised bog. *European Journal of Protistology*, *51*(1), 79–85.

https://doi.org/10.1016/j.ejop.2014.11.004

Singer, D., Kosakyan, A., Seppey, C. V. W., Pillonel, A., Fernández, L. D., Fontaneto, D., … Lara, E. (2018). Environmental filtering and phylogenetic clustering correlate with the distribution patterns of cryptic protist species. *Ecology*, *99*(4), 904–914. https ://doi.org/10.1002/ecy.2161

Škaloud, P., Škaloudová, M., Dosko ilová, P., Kim, J. I., Shin, W., & Dvo ák, P. (2019). Speciation in protists: Spatial and ecological divergence processes cause rapid species diversification in a freshwater chrysophyte. *Molecular Ecology*, *28*(5), 1084–1095.

Šlapeta, J., López-García, P., & Moreira, D. (2005). Global dispersal and ancient cryptic species in the smallest marine eukaryotes. *Molecular Biology and Evolution*, *23*(1), 23–29. https://doi.org/10.1093/molbe v/msj001

Smith, H. G., Bobrov, A., & Lara, E. (2007). Diversity and biogeography of testate amoebae. In W. Foissner & D. L. Hawksworth (Eds.), *Protist diversity and geographical distribution. Topics in biodiversity and conservation* (Vol. *8*). Dordrecht, the Netherlands: Springer. https ://link.sprin ger.com/chapter/10.1007/978-90-481-2801-3_8#citeas

Stamatakis, A., Hoover, P., Rougemont, J., & Renner, S. (2008). A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology*, *57*(5), 758–771. https ://doi.org/10.1080/10635 15080 2429642

Stenøien, H. K., Shaw, A. J., Shaw, B., Hassel, K., & Gunnarsson, U. (2010). North American origin and recent European establishments of the Amphi-Atlantic peat moss *Sphagnum Angermanicum*. *Evolution*, *65*(4), 1181–1194. https ://doi.org/10.1111/j.1558-5646.2010.01191.x

Svenning, J.-C. (2003). Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. *Ecology Letters*, *6*(7), 646–653. https ://doi.org/10.1046/j.1461-0248.2003.00477.x

Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2003). Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics*, *00*, 2.3.1–2.3.22. https ://doi.org/10.1002/0471250953.bi020 3s00

Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, *10*(12), 2020–2027. https ://doi.org/10.1111/j.1365-2486.2004.00859.x

Tolley, K. A., Townsend, T. M., & Vences, M. (2013). Large-scale phylogeny of chameleons suggests African origins and Eocene diversification. *Proceedings of the Royal Society of London B: Biological Sciences*, *280*(1759), 20130184. https ://doi.org/10.1098/rspb.2013.0184

Treat, C. C., Kleinen, T., Broothaerts, N., Dalton, A. S., Dommain, R., Douglas, T. A., … Brovkin, V. (2019). Widespread global peatland establishment and persistence over the last 130,000 y. *Proceedings of the National Academy of Sciences*, *116*(11), 4822. https://doi.org/10.1073/pnas.18133 05116

Trunz, V., Packer, L., Vieu, J., Arrigo, N., & Praz, C. J. (2016). Comprehensive phylogeny, biogeography and new classification of the diverse bee tribe Megachilini: Can we use DNA barcodes in phylogenies of large genera? *Molecular Phylogenetics and Evolution*, *103*, 245–259. https://doi.org/10.1016/j.ympev.2016.07.004

van Breemen, N. (1995). How *Sphagnum* bogs down other plants. *Trends in Ecology & Evolution*, *10*(7), 270–275. https ://doi.org/10.1016/0169-5347(95)90007-1

Velichko, A. A., Timireva, S. N., Kremenetski, K. V., MacDonald, G. M., & Smith, L. C. (2011). West Siberian Plain as a late glacial desert. *Quaternary International*, *237*(1–2), 45–53. https://doi.org/10.1016/j.quaint.2011.01.013

Wilkinson, D. M. (2001). What is the upper size limit for cosmopolitan distribution in free-living microorganisms? *Journal of Biogeography*, *28*(3), 285–291. https ://doi.org/10.1046/j.1365-2699.2001.00518.x

Wilkinson, D. M., Koumoutsaris, S., Mitchell, E. A. D., & Bey, I. (2011). Modelling the effect of size on the aerial dispersal of microorganisms. *Journal of Biogeography*, *39*(1), 89–97. https://doi.org/10.1111/j.1365-2699.2011.02569.x

Figures

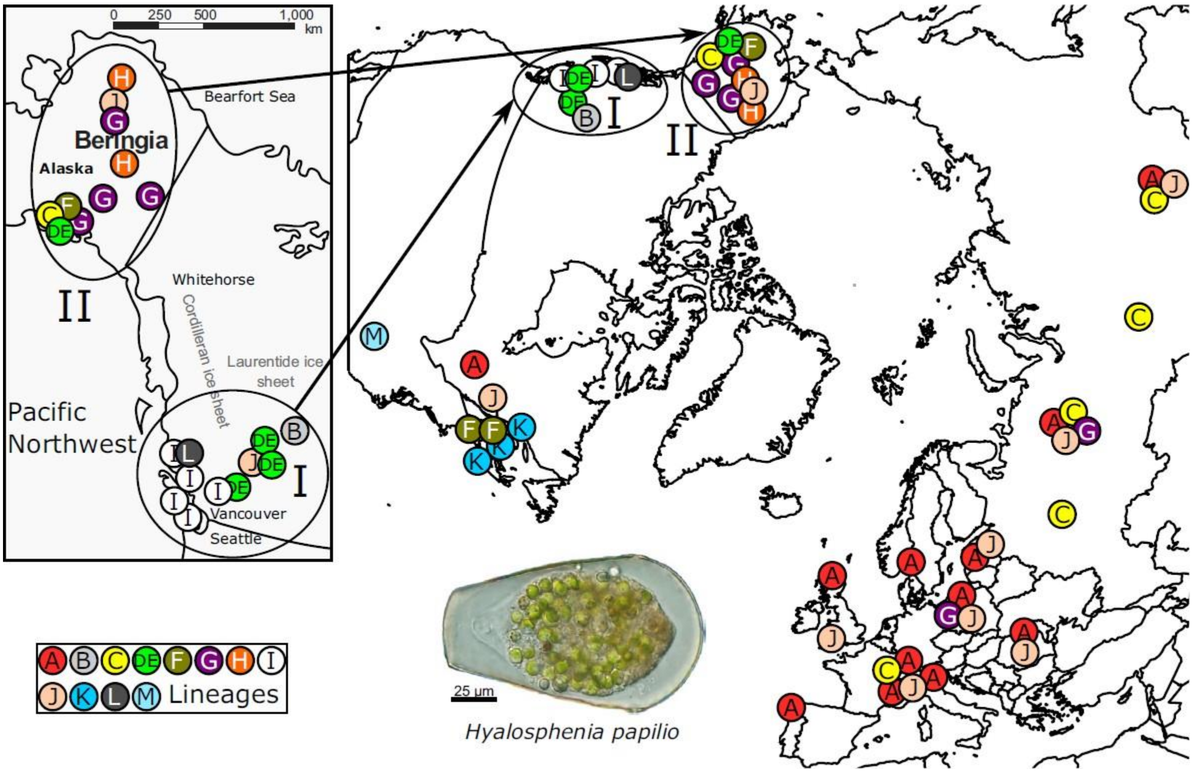

Figure 1: Holarctic distribution of *Hyalosphenia papilio* lineages. Each circle corresponds to a sampling site where the lineage has been detected. Lineage codes correspond to phylogenetic groups, as identified in Heger et al. (2013). I and II provide a detailed representation of the Beringia area. Inset: light micrograph of *H. papilio*. The pyriform outline corresponds to the shell that protects the single-cell body of the organism and its endosymbiotic microalgae (green dots).

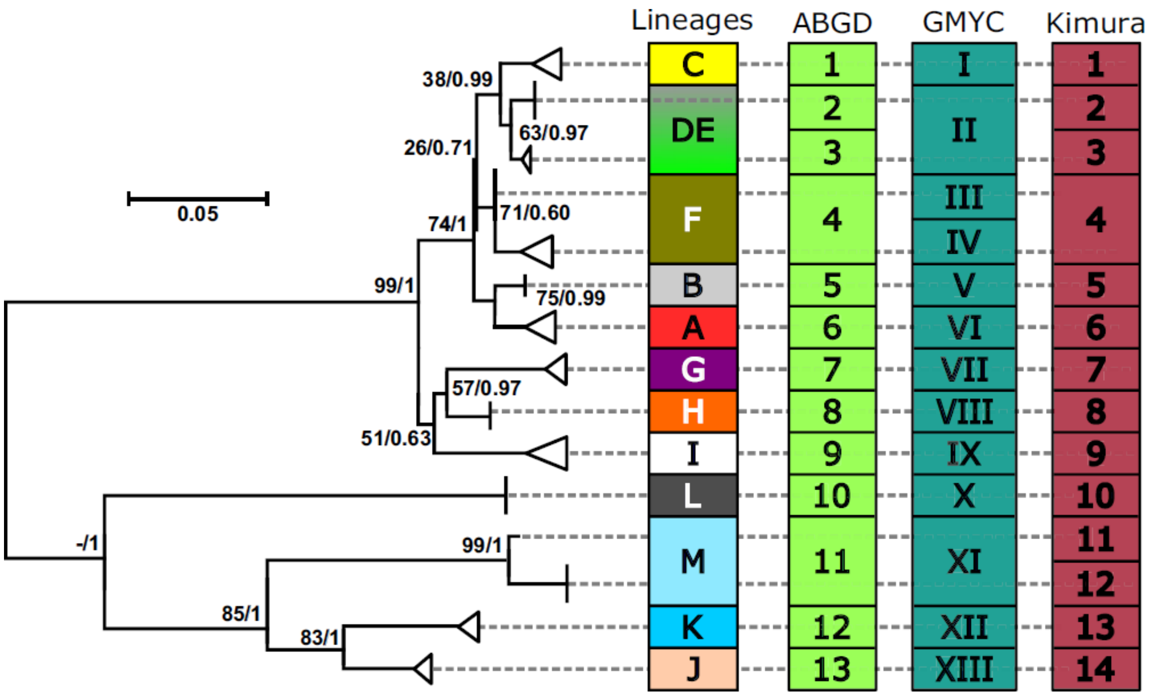

Figure 2: Maximum likelihood (ML) and Bayesian concatenated phylogenetic tree from 76 unique sequences of *Hyalosphenia papilio* isolated from *Sphagnum* peatlands across the Holarctic realm. Numbers along branches represent, respectively, bootstrap support values obtained by ML and posterior probabilities as calculated with Bayesian analyses. Trees were rooted internally based on the topology of trees obtained in previous studies (Heger et al., 2013; Kosakyan et al., 2016), which showed two major clades with maximum support; we rooted the tree between these two clades. The tree also represents the different lineages obtained with the ABGD and GMYC analyses and the Kimura 2-parameter test.

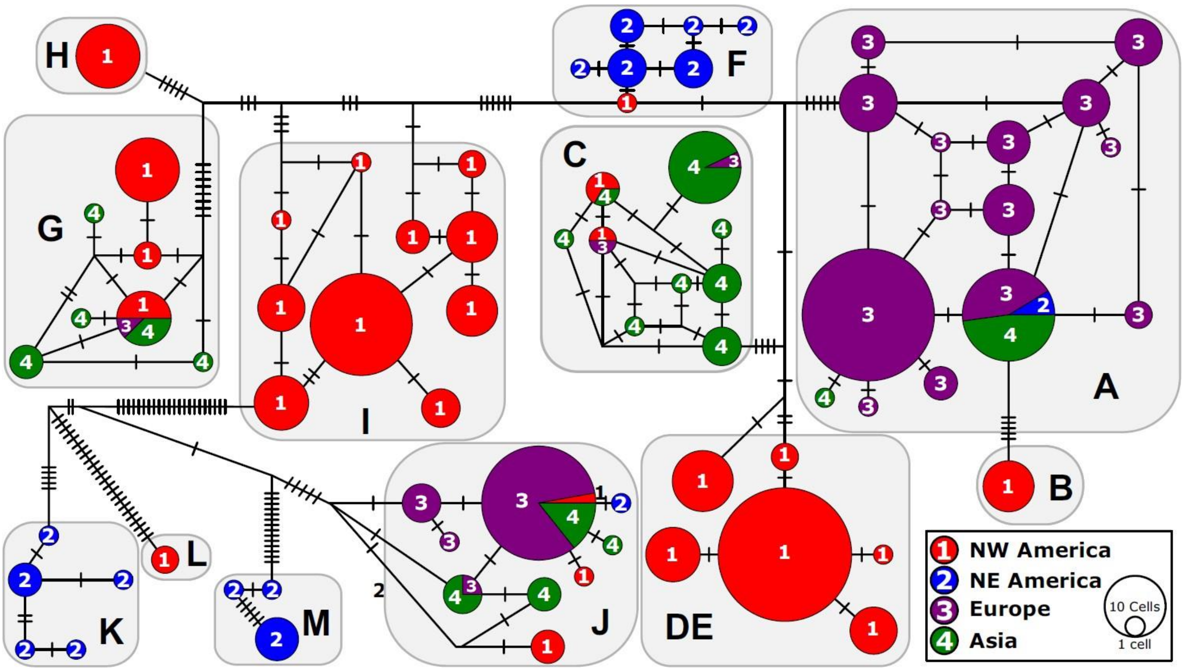

Figure 3: Median joining haplotype network of the cytochrome oxidase subunit 1 (*COI*) gene of *Hyalosphenia papilio* from *Sphagnum* peatlands in the Holarctic realm. Grey boxes and letters represent the different lineages identified in the present study. Colours indicate geographical regions (legend: bottom right inset). Circle sizes are proportional to the number of sequenced single cells of *H. papilio* within each haplotype. Cross lines show the number of mutational steps between haplotypes.

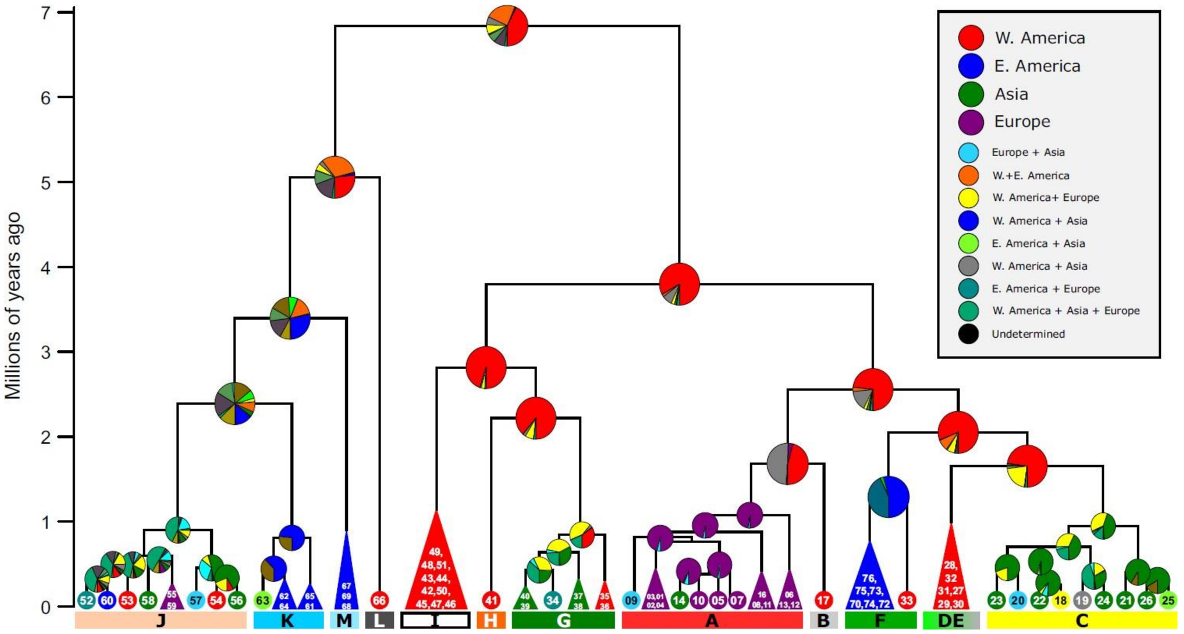

Figure 4: Biogeographical analysis of *Hyalosphenia papilio* from *Sphagnum* peatlands across the Holarctic realm using biogeobears. The four biogeographical areas are: Eastern North America (in blue), Western North America (in red), Europe (in purple) and Asia (in green). Pie charts at nodes indicate support for each area. The tips are labelled with present-day distributions. The secondary colours indicate range combinations of the tip ranges.

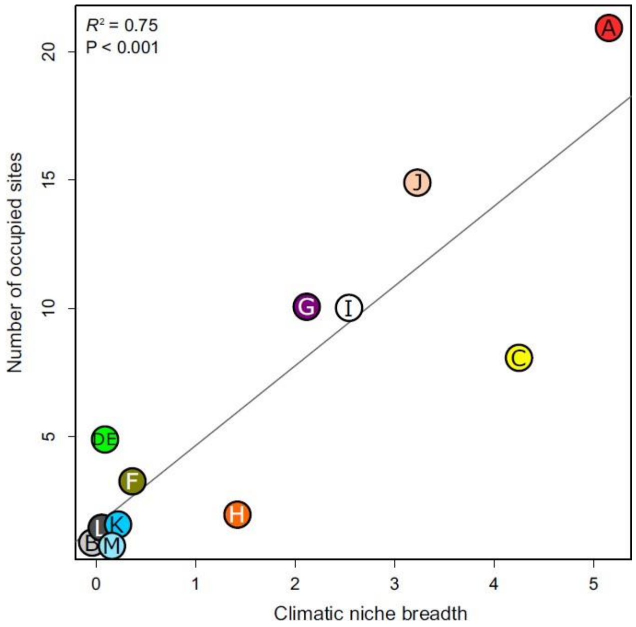

Figure 5: Relationship of lineage climatic niche breadth to lineage range size (estimated as the number of locations where a given lineage is present) of *Hyalosphenia papilio*. Climatic niche breadth was estimating from Bioclim variables using the tolerance index of Dolédec et al. (2000).

Appendix E
30 questions for soil protistology

Soil protistology rebooted: 30 fundamental questions to start with

Stefan Geisen [a, ag], Edward A.D. Mitchell [b, c], David M. Wilkinson [d, aq], Sina Adl [e], Michael Bonkowski [f], Matthew W. Brown [g], Anna Maria Fiore-Donno [f], Thierry J. Heger [b, at], Vincent E.J. Jassey [h, i], Valentyna Krashevska [j], Daniel J.G. Lahr [k], Katarzyna Marcisz [l, ar], Matthieu Mulot [b, as], Richard Payne [m], David Singer [b], O. Roger Anderson [n], Dan J. Charman [o], Flemming Ekelund [p], Bryan S. Griffiths [q], Regin Rønn [p], Alexey Smirnov [r], David Bass [s, t], Lassaâd Belbahri [b], Cédric Berney [u, v], Quentin Blandenier [b], Antonis Chatzinotas [w, x], Marianne Clarholm [y], Micah Dunthorn [z], Alan Feest [aa], Leonardo D. Fernández [ab], Wilhelm Foissner [ac], Bertrand Fournier [ad], Eleni Gentekaki [ae], Michal Hájek [af], Johannes Helder [ag], Alexandre Jousset [ah], Robert Koller [ai], Santosh Kumar [aj, ak], Antonietta La Terza [ak], Mariusz Lamentowicz [l], Yuri Mazei [al, am], Susana S. Santos [an], Christophe V.W. Seppey [b], Frederick W. Spiegel [ao], Julia Walochnik [ap], Anne Winding [an], Enrique Lara [b]

[a]Department of Terrestrial Ecology, Netherlands Institute of Ecology, 6708 PB Wageningen, The Netherlands
[b]Laboratory of Soil Biodiversity, University of Neuchâtel, Rue Emile-Argand 11, Neuchâtel 2000, Switzerland
[c]Jardin Botanique de Neuchâtel, Chemin du Perthuis-du-Sault 58, Neuchâtel 2000, Switzerland
[d]Natural Science and Psychology, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK
[e]Department of Soil Sciences, College of Agriculture and Bioresources, University of Saskatchewan, 51 Campus Drive, Saskatoon, Canada
[f]Department of Terrestrial Ecology, University of Cologne, Zuelpicher Str. 47b, 50674 Cologne, Germany
[g]Department of Biological Sciences, Mississippi State University, MS, 39762, United States
[h]School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne EPFL, Ecological Systems Laboratory (ECOS), Station 2, 1015 Lausanne, Switzerland
[i]Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Site Lausanne, Station 2, 1015 Lausanne, Switzerland
[j]Georg August University Göttingen, J.F. Blumenbach Institute of Zoology and Anthropology, Berliner Str. 28, 37073 Göttingen, Germany
[k]Department of Zoology, Institute of Biosciences, University of São Paulo, 05508-090, Brazil
[l]Laboratory of Wetland Ecology and Monitoring & Department of Biogeography and Paleoecology, Adam Mickiewicz University, Dziegielowa 27, 61-680 Poznan, Poland
[m]Environment, University of York, York YO105DD, United Kingdom
[n]Biology and Paleoenvironment, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, 10964, United States
[o]Department of Geography, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4RJ, United Kingdom
[p]Department of Biology, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen, Denmark
[q]SRUC, Crop and Soil Systems Research Group, West Mains Road, Edinburgh EH9 3JG, United Kingdom
[r]Department of Invertebrate Zoology, Faculty of Biology, Saint Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, 199034, Russia
[s]Department of Life Sciences, The Natural History Museum, London SW7 5BD, United Kingdom
[t]Cefas, Barrack Road, Weymouth, Dorset DT4 8UB, United Kingdom
[u]EPEP-Evolution des Protistes et des Ecosystèmes Pélagiques-team, Sorbonne Universités, UPMC Univ Paris 06, UMR 7144, Station Biologique de Roscoff, Roscoff, France
[v]CNRS, UMR 7144, Station Biologique de Roscoff, Roscoff, France
[w]Helmholtz Centre for Environmental Research e UFZ, Department of Environmental Microbiology, Permoserstr. 15, 04318 Leipzig, Germany
[x]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany
[y]Department of Forest Mycology and Plant Pathology, SLU, Uppsala, Sweden
[z]Department of Ecology, University of Kaiserslautern, 67663 Kaiserslautern, Germany
[aa]Faculty of Engineering, University of Bristol, Bristol BS8 1TR, United Kingdom

[ab]Centro de Investigación en Recursos Naturales y Sustentabilidad (CIRENYS), Universidad Bernardo O'Higgins, Fábrica 1990, 2 piso, Santiago, Chile
[ac]University of Salzburg, Department of Ecology and Evolution, Hellbrunnerstrasse 34, A-5020 Salzburg, Austria
[ad]Department of Biology, Concordia University, 7141 Sherbrooke Street West, Montreal, QC H4B 1R6, Canada
[ae]School of Science, Mae Fah Luang University, Chiang Rai, 57100, Thailand
[af]Department of Botany and Zoology, Masaryk University, Kotlárská 2, 61137 Brno, Czech Republic
[ag]Laboratory of Nematology, Wageningen University, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands
[ah]Department of Ecology and Biodiversity, Utrecht University, 3584 CH Utrecht, The Netherlands
[ai]Forschungszentrum Jülich, IBG-2: Plant Sciences, 52425 Jülich, Germany
[aj]Department of Biological Sciences, College of Natural Sciences, University of Ulsan, Ulsan 44610, South Korea
[ak]School of Bioscience and Veterinary Medicine, University of Camerino, Via Gentile III da Varano, 62032 Camerino (MC), Italy
[al]Department of Hydrobiology, Lomonosov Moscow State University, Leninskiye gory, 1, Moscow 119899, Russia
[am]Department of Zoology and Ecology, Penza State University, Krasnaya str. 40, 440026 Penza, Russia
[an]Department of Environmental Science, Aarhus University, Frederiksborgvej 399, 4000 Roskilde, Denmark
[ao]Department of Biological Sciences, University of Arkansas, Fayetteville, AR, 72701, United States
[ap]Molecular Parasitology, Institute of Tropical Medicine, Medical University, 1090 Vienna, Austria
[aq]School of Life Science, University of Lincoln, United Kingdom
[ar]Institute of Plant Sciences and Oeschger Centre for Climate Change Research, University of Bern, Altenbergrain 21, CH-3013 Bern, Switzerland
[as]Sorbonne Universités, UPMC Univ Paris 06, CNRS UMR 7144, Adaptation et Diversité en Milieu Marin,Equipe EPEP, Station Biologique de Roscoff, 29680 Roscoff, France
[at]The University of Applied Sciences Western Switzerland, Changins, Switzerland

Soil Biology and Biochemistry, 2017, Vol 111

DOI: 10.1016/j.soilbio.2017.04.001

Abstract. Protists are the most diverse eukaryotes. These microbes are keystone organisms of soil ecosystems and regulate essential processes of soil fertility such as nutrient cycling and plant growth. Despite this, protists have received little scientific attention, especially compared to bacteria, fungi and nematodes in soil studies. Recent methodological advances, particularly in molecular biology techniques, have made the study of soil protists more accessible, and have created a resurgence of interest in soil protistology. This ongoing revolution now enables comprehensive investigations of the structure and functioning of soil protist communities, paving the way to a new era in soil biology. Instead of providing an exhaustive review, we provide a synthesis of research gaps that should be prioritized in future studies of soil protistology to guide this rapidly developing research area. Based on a synthesis of expert opinion we propose 30 key questions covering a broad range of topics including evolution, phylogenetics, functional ecology, macroecology, paleoecology, and methodologies. These questions highlight a diversity of topics that will establish soil protistology as a hub discipline connecting different fundamental and applied fields such as ecology, biogeography, evolution, plant-microbe interactions, agronomy, and conservation biology. We are convinced that soil protistology has the potential to be one of the most exciting frontiers in biology.

Keywords. Soil protists; Protozoa; Microbial interactions; Food web; Biodiversity; Functional diversity

Introduction

Protists are everywhere, in aquatic and terrestrial ecosystems, free-living, and as symbionts (including parasites) of many organisms including humans. These usually single-celled or colonial microorganisms are by far the most diverse eukaryotes (Adl et al., 2012) and their species-numbers might easily exceed 10 million (Global Soil Biodiversity Atlas; www.globalsoilbiodiversity.org). Since the term 'protista' was introduced (Haeckel, 1866), profound taxonomic re-orderings have taken place. The vast majority of eukaryotic lineages has been shown to be protists, with the exception of the derived monophyletic multicellular lineages: animals, plants, and some fungi (Burki, 2014). Electron microscopy and molecular phylogenies have revealed that both algal and protozoan lineages are intermingled throughout the eukaryote phylogenies (Delwiche, 1999; Burki, 2014), and hence it is less confusing to use Haeckel's broader category of 'protist'. Similarly, the classical protozoan morphological categories: flagellates, testate and naked amoebae e but not ciliates - are not monophyletic but distributed across the eukaryotic tree of life (Adl et al., 2012). A snapshot of the immense morphological and phylogenetic diversity of soil protists is visualized in Fig. 1. We therefore recommend to use 'protist' as a term for all single celled phototrophic, mixotrophic and heterotrophic eukaryotes, with the exception of fungi.

The huge diversity of protist species has only recently become evident as many morphospecies recognizable under the microscope were shown to hide many cryptic species (Boenigk et al., 2012). This 'dark matter of biodiversity' suggests that protist taxon richness has been considerably underestimated. A recent study of environmental eukaryotic diversity based on state-of-the art high-throughput sequencing (HTS) showed that protists are considerably more diverse than plants and animals in the sunlit zone of oceans (de Vargas et al., 2015). HTS studies of soil protists have shown a wide diversity of non-phagotrophic protists and the diversity of protists in soils is at least as diverse as that in aquatic systems e.g. (Bates et al., 2013; Geisen et al., 2015c). Nevertheless, soil protists are much less well studied than their aquatic counterparts and this gap is increasing (Fig. 2a).

Soil protists have received relatively little attention mainly due to methodological challenges, especially their isolation from the opaque soil matrix. These, however, do not entirely explain why soil protists are relatively less studied than other soil organisms, especially bacteria, fungi and nematodes (Fig. 2b). The volume of work on microbial bacteria and fungi far outweighs protist studies, possibly because of their direct role as primary decomposers, and they represent monophyletic groups that can more easily be studied with various targeted methodological approaches (Foissner, 1987; Mitchell, 2015). Even soil viruses have been subject to more studies than soil protists, despite being extremely challenging to study (Fierer et al., 2007) and their uncertain functional importance in soils. The under-studied nature of soil protists is exemplified by a comparison between research on protists and on soil archaea, a domain erected in 1990 and reported to be functionally important in soil only decade ago (Leininger et al., 2006; Bates et al., 2011). Historically studies mentioning soil protists in the title were eight times more abundant than those including archaea (Fig. 2b, Supplementary Table 2). However, in the last 15 years, this pattern entirely changed; studies on soil protists decreased by 15% while those on other common soil organisms increased by at least 30%, especially soil archaea which increased by 88% (Fig. 2b, Supplementary Table 2).

The relative decline of papers on soil protists strongly contrasts with what we now know about their ubiquity, diversity, and perhaps more importantly, their functional significance. Soil protists can both make an important contribution to primary production (Jassey et al., 2015; Schmidt et al., 2016) and play a key role in the decomposition pathways as consumers of bacteria (Clarholm, 1981; de Ruiter et al., 1995), fungi, other protists, and small invertebrates; they can also act as parasites of plants and animals (Adl and Gupta, 2006; Jassey et al., 2012; Geisen, 2016b). As predators, protists transfer nutrients to higher trophic levels in the soil foodweb (de Ruiter et al., 1995; Crotty et al., 2012). Protist predation also stimulates microbial activity and nutrient cycling via the microbial loop, thus stimulating plant growth (Bonkowski and Clarholm, 2012) and representing an important link between aboveground and belowground components.

The functional significance, abundance, environmental sensitivity, rapid response times and increasing ease of analysis of soil protists also makes them invaluable bioindicators of a variety of aspects of environmental change (Foissner, 1987; Gupta and Yeates, 1997; Payne, 2013). A particular example of this is in paleoecology, where the hard shells of testate amoebae, diatoms and foraminifera are widely used in the reconstruction of past environments and past climate change over a range of timescales (Mitchell et al., 2008; Adl et al., 2011; Charman, 2015).

Most of these applications are, however, based on a few often small-scale studies. Thus significant taxonomic and functional aspects remain largely untouched. Our aim in this report is to pool expert knowledge and opinion across the diverse field of soil protistology and soil microbial ecology to identify major knowledge gaps that need to be addressed and their significance for soil processes and ecosystem services.

Materials and methods

Approach to identify the 30 most relevant questions

Our aim was to review research gaps both in the field of soil protistology and in general soil biology with a special focus on protists. In line with recent studies (Sutherland et al., 2013; Seddon et al., 2014), we aimed to pool community expertise to identify the most important questions in different broad categories. We modified previously-used methods (Sutherland et al., 2013; Seddon et al., 2014) to obtain a list of most interesting questions through a democratic, transparent, multi-step curation process.

The participants in this process are involved in a wide range of research areas, with self-determined primary research area expressed as being ecology (62%), paleoecology (12%), evolution (9%), biogeography (6%), phylogeny (6%), taxonomy (3%), parasitology (3%). Each participant formulated up to 10 questions that they believed were most relevant for their future research. The resulting 368 questions were then compiled via an integrative group effort into consensus questions and placed into six major categories following a discussion by 16 of the participants at the German Society for Protozoology meeting in February 2016. We included very broad, general questions as well as highly specialized topics into similar scaled consensus questions.

The resulting consensus questions were then re-evaluated and groupings adjusted in a vote. These questions (Supplementary Table 1) were sent out to all 47 participants, who individually indicated up to 12 priority questions with at least one being allocated in each of the six following categories: (i) Morphology, Phylogeny, Taxonomy, Evolution and Physiology, (ii) Diversity, Community Composition and Biogeography, (iii) Interactions among Protists and other Organisms, (iv) Functions of Protists, (v) Global Change, Bioindicators and Applications, and (vi) Methodology.

All 47 participants were asked to provide their key scientific expertise and literature references for studies that (partly) addressed individual questions. Finally, minor comments raised by individual participants during the vote were integrated to clarify the questions and give consistent formatting without changing the meaning of the questions that had been voted upon.

All individual votes were combined and five questions per category chosen to result in the final list of 30 key questions. When more than one question received the same number of votes (as present in categories 1, 2, and 4), these questions were sent out to all 47 participants for another vote on the selected questions only.

Potential limitations

Biases in broad-scale studies are impossible to avoid (Sutherland et al., 2013). On the other hand, the more interdisciplinary the panel of authors is in terms of cultural and societal background and specific scientific expertise, the more biases are reduced. Researchers working on soil protists are often ecologists, whereas taxonomists, phylogeneticists, and physiologists more often focus on aquatic taxa that are easier to isolate and cultivate. Indeed, participants who indicated ecology as their first expertise dominated our list of participants (62%). Ecology, however, is a broad field and our division into finer categories such as biogeography, paleoecology, community structure, and interactions resulted in a broad diversification into different subcategories. Additionally, 21% of the participants indicated topics such as taxonomy, phylogeny, evolution, and physiology as their main expertise corroborating the wide diversity of research fields among the co-authors.

The majority of participants are PhDs (Professor: 38%; Graduated scientists: 29%; Post-doc: 24%) with an average number of publications on protists of 43 (minimum = 1; maximum = 230). A high proportion of the participants work on multiple ecosystems (41%). Many focus on testate amoebae (41%) although 26% of them work on multiple morphogroups (ciliates, heterotrophic flagellates, amoebae, etc.). A majority (74%) of participants have a European background, but Asia, North and South America are also well represented, thus reducing potential impacts of geographic origin. Furthermore, most participants have international collaborations that partly compensate for gaps in the geographic distribution of individuals.

Despite these potential limitations, we found few biases in the way participants replied to questions (Supplementary Results 1). Most participants (70%) selected questions evenly distributed across the six categories, except a small group of people mostly constituted of researchers from the same institute and/or with the same kind of expertise (phylogeny, taxonomy and evolutionary; see Supplementary Results 1). This small group allocated 45% of their votes to the category (ii). Except this small bias, most participants selected questions regardless of their experience, age, geographic background, and most importantly, their expertise and group of interest.

Questions were differently formulated, hence we had to make decisions and remove some nuances as we merged similar questions. This resulted in some discussions about how questions should be best stated and consequently combined and grouped into non-predetermined categories. However, we preferred to receive non-restricted questions to stimulate lateral thinking as previously suggested (Sutherland et al., 2013); due to intensive exchange and to a democratic group effort at all steps of the procedure, we are convinced that we have reached a consensus format.

Results and discussion

The 30 most relevant consensus questions

From the 107 questions in the final vote, 94% received at least one, 79% two, 67% three and 50% five votes showing that the preselected questions had a wide general appeal to the scientific experts involved (Supplementary Table 1). Therefore, all questions seem to be relevant for future studies that focus on soil protists. However, as we aimed at providing a highly specific list of the major research gaps and open challenges in soil protistology, we only provide the top-ranked 30 questions classified in six major categories that most researches voted upon.

Linking the individual topics of this category is one of the major tasks confronting soil protistology. The coupling of morphology and phylogeny is crucial to obtain a stable taxonomic framework for protists. This is, for instance, crucial to answer evolutionary questions on the origin of eukaryotes (López-García and Moreira, 2015). Soil protists may have an important role to play in such research as most taxa likely remain unknown and novel higher-level taxonomic groups are continuously being discovered (Berney et al., 2015; Bass et al., 2016; Singer et al., 2016; Tice et al., 2016). Soil protists might fill remaining phylogenetic gaps from better-studied aquatic taxa to improve phylogenetic resolution within and between protist clades, as strict soil protist clades seem to be common (Bass et al., 2016). Sequencing whole genomes will reveal ancient traits of eukaryotes and potential changes in their function during the evolution of eukaryotes. In this respect, soil protists must certainly play a key role for understanding the evolution of the eukaryotic cell and, therefore, of life as a whole.

While the morphological and phylogenetic framework for ciliates is reasonably well established (Lynn, 2008; Foissner, 2016), it remains rudimentary for other morphogroups as well as phylogenetic clades of protists (Kosakyan et al., 2016). The taxonomy of the groups has profoundly benefited from (mainly) 18S rRNA genebased characterisations that have often led to drastic changes in phylogenetic placements of individual species, genera, families or even orders (Boenigk et al., 2012; Berney et al., 2015; Bass et al., 2016). This is exemplified by the morphologically and functionally diverse Cercozoa, which was the first protist clade inferred solely based on molecular phylogenetic information, and has become home to ever more morphologically different organisms (Bass et al., 2016). Therefore, the true extent of morphological and genetic variability in different groups of soil protists remains largely unknown and is a key missing gap for future studies (Q2).

A key feature of soil protist species is their capacity for cyst formation as this allows them to resist constantly changing conditions, especially with respect to moisture and temperature. Furthermore, given that protists can excyst after decades, even millennia (Shmakova et al., 2016), cyst formation may protect species from becoming extinct at local or even at the global scales, influence population dynamics and maintain biodiversity (Corliss and Esser, 1974; Jones and Lennon, 2010) The importance of the cyst bank in ecosystem functioning and resilience remains largely unknown and have consequently been identified as a key element for future studies (Q1). More generally, this applies to all specific (physiological) adaptations of soil protists in comparison to their aquatic relatives (Q3, Q4) and to reproduction (Q5).

We are progressively shedding light into the soil 'black box'; however, knowledge on protists lags behind that of other groups (Fig. 2) (Wilkinson, 2008). Traditional studies have focused exclusively on a few of the 'classic' morphogroups, especially ciliates and testate amoebae, at least partly due to their ease of isolation and feature-rich morphologies (Foissner, 1999). Despite dominating many soil protist communities in terms of numbers and diversity, flagellates and naked amoebae have remained understudied, due to their often smaller sizes, lack of diagnostic features when studied by light microscopy, and the need to establish specialised enrichment cultivation for their isolation from soils (Berthold and Palzenberger, 1995; Foissner, 1999; Smirnov and Brown, 2004; Tikhonenkov et al., 2010). The development of molecular tools such as DNA barcoding and metabarcoding has considerably improved the situation in the last decade (Pawlowski et al., 2012) and allowed a phylogenetically based (morphogroup-independent) and consequently much more detailed analysis of the entirety of soil protist communities. These studies have revealed an enormous diversity of protists inhabiting soils, a phylogenetic diversity that might be similar to that of bacteria e.g., (Bates et al., 2013; Geisen et al., 2015c; Mahé et al., 2017). Also, groups of protists previously almost unknown from soils have been shown to be common e.g. choanoflagellates, foraminifera, dinoflagellates, parasitic apicomplexans and pathogenic oomycetes (Bates et al., 2013; Geisen et al., 2015c; Grossmann et al., 2016; Mahé et al., 2017). Therefore, we are only beginning to understand the diversity of soil protists (Q10), which part is active, how this diversity differs in different soil environments (Q6), how protist communities are structured by, e.g., abiotic factors (Q7, Q8) (Geisen et al., 2014a, 2015c; Lentendu et al., 2014; Dupont et al., 2016).

In addition, the biogeography of (soil) protists has been addressed in some studies, but it is still unclear which and how many groups display a restricted biogeography and what the factors are that shape these distributions (Q8, Q9). Although distribution of bacteria may support the hypothesis that "everything is everywhere, but, the environment selects" (Baas-Becking, 1934), its extrapolation to protists has been countered, particularly, by work on ciliates and testate amoebae (Foissner, 2006, 2008; Fernández et al., 2016). The diversity and biogeographical distribution of protists, is, therefore, one of degree (rather than all cosmopolitan or all limited) and the possibility exists that the investigation of cryptic diversity within morphospecies will allow a finer-scale resolution of these questions.

Soil protists are still predominantly considered as being mainly bacterivorous (Bradford, 2016; Geisen, 2016a). Differential feeding by protists stimulated by bacterial volatiles modifies the community composition of bacteria (Bonkowski, 2004; Glücksman et al., 2010; Schulz-Bohm et al., 2017), which results in functional changes in the bacterial community structure (see next section). Many free-living bacteria can, in turn, defend themselves against certain protist predators and even kill them (Greub and Raoult, 2004; Jousset et al., 2006). Several bacteria, viruses, and even other protists can also parasitize protist hosts (Barker and Brown, 1994; Raoult and Boyer, 2010).

The prey spectrum of protists has, however, repeatedly been shown to be much more diverse than bacteria. Indeed, archaea (Ballen-Segura et al., 2017), fungi (Gupta and Germida, 1988; Ekelund, 1998; Adl and Gupta, 2006; Geisen et al., 2016), other protists (Page, 1977; Jassey et al., 2012), and nematodes (Bjørnlund and Rønn, 2008; Geisen et al., 2015b) constitute prey for diverse protist species. Recently, HTS approaches have revealed the ubiquitous presence and dominant roles of protist parasites and pathogens in soils, and they likely represent a key component controlling other soil organisms including larger soil metazoans (animals) and plants (Geisen et al., 2015a; Dupont et al., 2016; Geisen, 2016b). This draws attention to the enormous complexity and importance of protist interactions with other organisms (Bonkowski, 2004).

Due to our limited knowledge of protist diversity and because most studies have used only one or few protists as models, we lack understanding about most aspects of how soil protist communities interact with other organisms. Disentangling the diverse interactions of protists with other soil organisms (Q11, Q12, Q15), the exact mechanisms (Q14) and the resulting importance for functioning (Q13, Q14), therefore, are key knowledge gaps necessitating future research.

Many acknowledged functions of soil protists are attributed to interactions with other organisms as outlined above. Especially important is the role of protists in driving the microbial loop, i.e. releasing nutrients (particularly nitrogen) bound in bacterial prey. The microbial loop has been demonstrated both in aquatic (Azam et al., 1983) and soil systems (Clarholm, 1985). This ground-breaking research identified protists as important drivers of the global ecosystem. Subsequent work on the microbial loop demonstrated that differential feeding by protists on bacterial prey is beneficial for plant growth (Bonkowski, 2004; Rosenberg et al., 2009). The main focus in earlier studies was, however, mainly on nitrogen cycling, and the importance of protists for cycling of other elements such as carbon and phosphorus has been relatively neglected, with few exceptions (Cole et al., 1977; Gupta and Germida, 1988; Treonis and Lussenhop, 1997; Frey et al., 2001; Murase et al., 2011; Eisenhauer et al., 2012; Jassey et al., 2015). Protists might even play a role in silica cycling as some use Si as reinforcing elements or in an exoskeleton (Aoki et al., 2007; Creevy et al., 2016). More thorough investigations about the functional roles of additional protist species and communities as a whole will likely reveal insights into the importance of protists in biogeochemical nutrient cycling. This was identified by most participants of this study as the most important question for future research (Q16).

In contrast to free-living protists, plant pathogenic protists, such as oomycetes or plasmodiophorids, have, not surprisingly, attracted considerable attention due to their agro-economic impact (Anderson et al., 2004; Bell et al., 2006; Neuhauser et al., 2014). These were, however, until very recently often considered as 'fungi' (Schardl and Craven, 2003; Gams et al., 2011). Similarly, soil protists with immediate relevance for human diseases such as those directly harmful to humans (Schuster, 2002; Siddiqui and Ahmed Khan, 2012; Geisen et al., 2014b) and those that act as "Trojan horses" harbouring human-pathogenic bacteria (Brown and Barker, 1999; Molmeret et al., 2005) have received considerable attention. In turn, the role of protists in plant disease control due to, e.g., increasing bacterial biocontrol agents (Jousset, 2012) or by directly feeding on plant pathogens (Old and Chakraborty, 1986; Geisen et al., 2016) has received comparatively little attention. In line with their importance in nutrient cycling and as biocontrol agents, the role of individual protist species and that of protist diversity for the general functioning of soils and ecosystems (Q17, Q18, Q20), also in comparison to other groups of microbes (Q19), were identified as important questions to be addressed in future studies.

Protist communities are often studied as bioindicators of past and present climatic conditions, land use changes and pollution (Gupta and Yeates, 1997; Mitchell et al., 2008). Abiotic changes affect protists in species-specific ways, thus forming the basis for their use as bioindicators (Fournier et al., 2012). They may, for instance, provide information on soil state in agro-ecosystems (Foissner, 1997, 1999; Bharti et al., 2015). Testate amoebae and their subfossil remains have been used to evaluate wetland hydrological conditions, applied, for instance, in studies of peatland restoration (Marcisz et al., 2014) and reconstruction of Holocene environmental change (Turner et al., 2014; Lamentowicz et al., 2015; Payne et al., 2015). However, more generally, there has been little progress on evaluating protists as bioindicators even though reliable indicators to assess soil quality continue to be of high relevance (Griffiths et al., 2016) as also revealed here

(Q22). Application of protists for stimulating plant performance in terms of nutrition, growth, productivity and disease suppression holds great promise but has received little attention (Q24).

Effects of ongoing global climate change and human impact on the environment are the focus of increasing scientific attention. Global warming has been shown to alter the abundance and community structure of protists (Tsyganov et al., 2011; Jassey et al., 2013) in the limited number of studies that have been done. Predicted changes in precipitation regime will likely affect water availability, which will impact protist communities directly (Clarholm, 1981; Bates et al., 2013; Geisen et al., 2014a). Elevated atmospheric CO2 has also been shown to increase abundance and changes community structure of rhizosphere protists, possibly due to increased plant productivity and enhanced release of root organic exudates (e.g., (Treonis and Lussenhop, 1997; Anderson and Griffin, 2001; Rønn et al., 2002)). Increased air pollution by nitrogen, sulphur, tropospheric ozone and metals are also likely to alter protist abundance and diversity (Meyer et al., 2012; Payne et al., 2012, 2013). Most of these studies focused on testate amoebae, but it is important to study how global environmental changes affect entire protist communities (Q21, Q23, Q25) as these changes are likely to have significant impacts on ecosystem functioning/ services and, consequently, on human welfare, and may provide more informative markers of environmental change.

Diverse methods are used to study community structures of soil protists. Even with respect to more classical culturing and morphological observational techniques, the application of methods of non-protistological disciplines, such as mycology, have the potential of broadening our perspectives on the soil protist community (Spiegel et al., 2004). However, especially recent developments in molecular sequencing technologies, have changed and will continue to change our knowledge about protist diversity and community structure in soils (Bates et al., 2013; Geisen et al., 2015c). However, some issues relating to HTS-based efforts remain as they provide relative abundances of taxa without providing information on absolute abundances. For example PCR-based HTS efforts have been shown to artificially alter the observed community structure of soil protists, a problem which needs to be solved to decipher their real community structure (Geisen et al., 2015a). PCR-free 'omics-approaches', i.e. metagenomics and metatranscriptomics, might resolve some of these issues (Geisen et al., 2015c; Jacquiod et al., 2016). Indeed, these sequence-based omics approaches and sequence-independent metaproteomics provide valuable information not only on taxonomic diversity but also on their potential functions (Prosser, 2015). Calibrating, standardizing and adopting community-defined methodologies to study soil protists will, consequently, be key for cross-study comparisons (Q27) and correct sampling and analyses through different scales need to be defined a priory (Q28, Q29). Furthermore, it is essential to identify the most meaningful taxonomic levels to use in the study of diversity and functioning of soil protists (Q26), but even the definition of a species remains a challenge (Boenigk et al., 2012) and integrating morphology to phylogeny to function remains missing (Q30).

In addition, medical and novel imaging techniques applied to soil are revolutionising in situ work allowing us to study protist species in undisturbed soil and on plant roots. These include applications of NanoSIMS technology to precisely locate isotopic markers and isotopic composition of material in fixed preparations and to study dynamics of nutrient fluxes (Stockdale et al., 2009), which allows tracing nutrient flow from microbial prey to protist predator and further in the food web in high resolution. This will allow detailed investigations how protists selective interact in microsites with their prey, how nutrients become released and where they are translocated. Applications of a variety of X-ray based synchrotron spectroscopy and tomography with undisturbed soil is becoming technically feasible and permits the study of dynamics and fluxes at a very fine resolution without interfering with the matrix (Keyes et al., 2013). The ability to use soils with intact fine roots, and examining undisturbed natural soil communities finally provides access to rhizosphere processes. Techniques to measure and analyse chemically soil community molecular interactions and communications are now only a few steps away.

(Partial) knowledge gaps and future directions

In this paper we provide a guide to 30 highly relevant questions for future studies in soil protistology. Research has already been conducted on many of these questions. Literature searches and personal knowledge of the literature allowed us to identify studies that addressed 91% of the initial and 97% of the final questions. However, many of these studies focus on organisms other than soil protists (e.g. aquatic protists or non-protist microbes), and may not be directly applicable to the situation with soil protists. The fact that these 30 questions have been identified by our pool of experts strongly implies that previous research has been insufficient to provide conclusive answers. In Supplementary Table 1 we provide an extensive bibliography of previous research relevant to addressing these questions. This bibliography will be a valuable literature guide to the current state of the art on soil protistology.

We are beginning to understand many aspects of soil protist biology, as we are identifying the hyperdiverse nature of protist communities, determining their (a)biotic drivers, deciphering interactions with other organisms, and shedding light on their importance in ecosystem dynamics. So far, however, we are only seeing the tip of the iceberg. Addressing many of the 30 questions highlighted here will undoubtedly reveal novel insights, not only into soil protists, but also into other organisms, soils, and fundamental ecological processes. We hope that these questions will be used to catalyse soil protistology and to build research agendas for the future. More specifically, we encourage both protistologists and researchers in closely related fields to consider these questions carefully and to use them to develop new and innovative individual and collaborative projects. With newly available techniques, an increase in knowledge and a growing awareness of the importance of soil protists, we are at the start of a bright future for soil protist research!

Acknowledgements

We thank Alex K. Tice for providing pictures of *Acrasis* and *Rosculus* and Clément Duckert for providing the picture of *Euglypha* included in Fig.1. SG thanks Wim van der Putten and his ERC-Adv grant 260e55290 (SPECIALS). We further acknowledge funding by the Swiss Contribution to the enlarged European Union (Project CLIMPEAT: PSPB-013/2010) to ML, KM, MM, VEJJ and EADM, the Russian Science Foundation (no 14-14-00891 to RJP and YuM and 14-50-00029 to YuM), the São Paulo Research Foundation (FAPESP) (Young Investigator Award no 2013/04585-3) to DJGL, the Danish Council for Independent Research (DFF-4002-00274) to FE, the Swiss National Science Foundation to VEJJ (no 315260_149807), EL, QB and DS (grants no 143960 and 163254), EADM & TH (no 141188 and 163431), the Chilean Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) (doctoral grants no 21110037 and 78130011), the Universidad de Concepcion (Chile) and the Wildlife Conservation Society (Karukinka-grant no. 2012) to LDF, the Russian Science Foundation (no 14-14-00891 to RJP and YuM and 14-50-00029 to YuM, and the University of Neuchâtel to the SNF projects to EL (31003A_163254) and EADW Mitchell (31003A_163431). This is publication of the NIOO-KNAW.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.soilbio.2017.04.001.

References

Adl, M.S., Gupta, V.V.S.R., 2006. Protists in soil ecology and forest nutrient cycling. Canadian Journal of Forest Research 36, 1805e1817.

Adl, S., Girard, V., Breton, G., Lak, M., Maharning, A., Mills, A., Perrichot, V., Trionnaire, M., Vullo, R., Néraudeau, D., 2011. Reconstructing the soil food web of a 100 million-year-old forest: the case of the mid-Cretaceous fossils in the amber of Charentes (SW France). Soil Biology and Biochemistry 43, 726e735.

Adl, S.M., Simpson, A.G.B., Lane, C.E., Lukes, J., Bass, D., Bowser, S.S., Brown, M.W., Burki, F., Dunthorn, M., Hampl, V., Heiss, A., Hoppenrath, M., Lara, E., le Gall, L., Lynn, D.H., McManus, H., Mitchell, E.A.D., Mozley-Stanridge, S.E., Parfrey, L.W., Pawlowski, J., Rueckert, S., Shadwick, L., Schoch, C.L., Smirnov, A., Spiegel, F.W., 2012. The revised classification of eukaryotes. Journal of Eukaryotic Microbiology 59, 429e514.

Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R., Daszak, P., 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends in Ecology and Evolution 19, 535e544.

Anderson, O.R., Griffin, K.L., 2001. Abundances of protozoa in soil of laboratory grown wheat plants cultivated under low and high atmospheric CO2 concentrations. Protistology 2.

Aoki, Y., Hoshino, M., Matsubara, T., 2007. Silica and testate amoebae in a soil under pineeoak forest. Geoderma 142, 29e35.

Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., Thingstad, F., 1983. The ecological role of water-column microbes in the sea. Marine Ecology Progress Series 10, 257e263.

Baas-Becking, L.G.M., 1934. Geobiologie, of inleiding tot de milieukunde. Van Stockum.

Ballen-Segura, M., Felip, M., Catalan, J., 2017. Some mixotrophic flagellate species selectively graze on archaea. Applied and Environmental Microbiology 83.

Barker, J., Brown, M.R.W., 1994. Trojan Horses of the microbial world: protozoa and the survival of bacterial pathogens in the environment. Microbiology 140, 1253e1259.

Bass, D., Silberman, J.D., Brown, M.W., Tice, A.K., Jousset, A., Geisen, S., Hartikainen, H., 2016. Coprophilic amoebae and flagellates, including Guttulinopsis, Rosculus and Helkesimastix, characterise a divergent and diverse rhizarian radiation and contribute to a large diversity of faecal-associated protists. Environmental Microbiology 18, 1604e1619.

Bates, S.T., Berg-Lyons, D., Caporaso, J.G., Walters, W.A., Knight, R., Fierer, N., 2011. Examining the global distribution of dominant archaeal populations in soil. The ISME Journal 5, 908e917.

Bates, S.T., Clemente, J.C., Flores, G.E., Walters, W.A., Parfrey, L.W., Knight, R., Fierer, N., 2013. Global biogeography of highly diverse protistan communities in soil. ISME Journal 7, 652e659.

Bell, T., Freckleton, R.P., Lewis, O.T., 2006. Plant pathogens drive density-dependent seedling mortality in a tropical tree. Ecology Letters 9, 569e574.

Berney, C., Geisen, S., Van Wichelen, J., Nitsche, F., Vanormelingen, P., Bonkowski, M., Bass, D., 2015. Expansion of the ‘Reticulosphere’: diversity of novel branching and network-forming amoebae helps to define Variosea (Amoebozoa). Protist 166, 271e295.

Berthold, A., Palzenberger, M., 1995. Comparison between direct counts of active soil ciliates (Protozoa) and most probable number estimates obtained by Singh's dilution culture method. Biology and Fertility of Soils 19, 348e356.

Bharti, D., Kumar, S., La Terza, A., 2015. Two gonostomatid ciliates from the soil of Lombardia, Italy; including note on the soil mapping project. Journal of Eukaryotic Microbiology 62, 762e772.

Bjørnlund, L., Rønn, R., 2008. ‘David and Goliath’ of the soil food web e flagellates that kill nematodes. Soil Biology and Biochemistry 40, 2032e2039.

Boenigk, J., Ereshefsky, M., Hoef-Emden, K., Mallet, J., Bass, D., 2012. Concepts in protistology: species definitions and boundaries. European Journal of Protistology 48, 96e102.

Bonkowski, M., 2004. Protozoa and plant growth: the microbial loop in soil revisited. New Phytologist 162, 617e631.

Bonkowski, M., Clarholm, M., 2012. Stimulation of plant growth through interactions of bacteria and protozoa: testing the auxiliary microbial loop hypothesis. Acta Protozool 51, 237e247.

Bradford, M.A., 2016. Re-visioning soil food webs. Soil Biology and Biochemistry 102, 1e3.

Brown, M.R.W., Barker, J., 1999. Unexplored reservoirs of pathogenic bacteria: protozoa and biofilms. Trends in Microbiology 7, 46e50.

Burki, F., 2014. The eukaryotic tree of life from a global phylogenomic perspective. Cold Spring Harbor Perspectives in Biology 6, a016147.

Charman, D.J., 2015. Testate amoebae. In: Shennan, I., Long, A.J., Horton, B.P. (Eds.), Handbook of Sea-level Research. Wiley, Chichester, pp. 281e294.

Clarholm, M., 1981. Protozoan grazing of bacteria in soil impact and importance. Microbial Ecology 7, 343e350.

Clarholm, M., 1985. Interactions of bacteria, protozoa and plants leading to mineralization of soil nitrogen. Soil Biology and Biochemistry 17, 181e187.

Cole, C.V., Elliott, E.T., Hunt, H.W., Coleman, D.C., 1977. Trophic interactions in soils as they affect energy and nutrient dynamics. V. Phosphorus transformations. Microbial Ecology 4, 381e387.

Corliss, J.O., Esser, S.C., 1974. Comments on the role of the cyst in the life cycle and survival of free-living protozoa. Transactions of the American Microscopical Society 93, 578e593.

Creevy, A.L., Fisher, J., Puppe, D., Wilkinson, D.M., 2016. Protist diversity on a nature reserve in NW Englanddwith particular reference to their role in soil biogenic silicon pools. Pedobiologia 59, 51e59.

Crotty, F., Adl, S., Blackshaw, R., Murray, P., 2012. Protozoan pulses unveil their pivotal position within the soil food web. Microbial Ecology 63, 905e918.

de Ruiter, P.C., Neutel, A.M., Moore, J.C., 1995. Energetics, patterns of interaction strengths, and stability in real ecosystems. Science 269, 1257e1260.

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., Lukes, 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, 1261605.

Delwiche, C.F., 1999. Tracing the thread of plastid diversity through the tapestry of life. The American Naturalist 154, S164eS177.

Dupont, A.O., Griffiths, R.I., Bell, T., Bass, D., 2016. Differences in soil microeukaryotic communities over soil pH gradients are strongly driven by parasites and saprotrophs. Environmental Microbiology 18, 2010e2024.

Eisenhauer, N., Cesarz, S., Koller, R., Worm, K., Reich, P.B., 2012. Global change belowground: impacts of elevated CO2, nitrogen, and summer drought on soil food webs and biodiversity. Global Change Biology 18, 435e447.

Ekelund, F., 1998. Enumeration and abundance of mycophagous protozoa in soil, with special emphasis on heterotrophic flagellates. Soil Biology and Biochemistry 30, 1343e1347.

Fernández, L.D., Fournier, B., Rivera, R., Lara, E., Mitchell, E.A.D., Hernández, C.E., 2016. Water-energy balance, past ecological perturbations and evolutionary constraints shape the latitudinal diversity gradient of soil testate amoebae in south-western South America. Global Ecology and Biogeography 25, 1216e1227.

Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., Robeson, M., Edwards, R.A., Felts, B., Rayhawk, S., Knight, R., Rohwer, F., Jackson, R.B., 2007. Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. Applied and Environmental Microbiology 73, 7059e7066.

Foissner, W., 1987. Soil protozoa: fundamental problems, ecological significance, adaptations in ciliates and testaceans, bioindicators, and guide to the literature. Progress in Protistology 2, 69e212.

Foissner, W., 1997. Protozoa as bioindicators in agroecosystems, with emphasis on farming practices, biocides, and biodiversity. Agriculture, Ecosystems & Environment 62, 93e103.

Foissner, W., 1999. Soil protozoa as bioindicators: pros and cons, methods, diversity, representative examples. Agriculture, Ecosystems & Environment 74, 95e112.

Foissner, W., 2006. Biogeography and dispersal of micro-organisms: a review emphasizing protists. Acta Protozoologica 45, 111e136.

Foissner, W., 2008. Protist diversity and distribution: some basic considerations. Biodiversity and Conservation 17, 235e242.

Foissner, W., 2016. Terrestrial and semiterrestrial ciliates (Protozoa, Ciliophora) from Venezuela and Galapagos. Denisia 35, 1e912.

Fournier, B., Malysheva, E., Mazei, Y., Moretti, M., Mitchell, E.A.D., 2012. Toward the use of testate amoeba functional traits as indicator of floodplain restoration success. European Journal of Soil Biology 49, 85e91.

Frey, S.D., Gupta, V.V.S.R., Elliott, E.T., Paustian, K., 2001. Protozoan grazing affects estimates of carbon utilization efficiency of the soil microbial community. Soil Biology and Biochemistry 33, 1759e1768.

Gams, W., Diederich, P., Poldmaa, K., 2011. Biodiversity of fungi: inventory and monitoring methods. In: Mueller, G.M., Bills, G.F., Foster, M.S. (Eds.), Biodiversity of Fungi: Inventory and Monitoring Methods. Elsevier.

Geisen, S., 2016a. The bacterial-fungal energy channel concept challenged by enormous functional versatility of soil protists. Soil Biology and Biochemistry 102, 22e25.

Geisen, S., 2016b. Thorough high-throughput sequencing analyses unravels huge diversities of soil parasitic protists. Environmental Microbiology 18, 1669e1672.

Geisen, S., Bandow, C., R€ombke, J., Bonkowski, M., 2014a. Soil water availability strongly alters the community composition of soil protists. Pedobiologia 57, 205e213.

Geisen, S., Fiore-Donno, A.M., Walochnik, J., Bonkowski, M., 2014b. Acanthamoeba everywhere: high diversity of Acanthamoeba in soils. Parasitology Research 113, 3151e3158.

Geisen, S., Koller, R., Hünninghaus, M., Dumack, K., Urich, T., Bonkowski, M., 2016. The soil food web revisited: diverse and widespread mycophagous soil protists. Soil Biology and Biochemistry 94, 10e18.

Geisen, S., Laros, I., Vizcaíno, A., Bonkowski, M., de Groot, G.A., 2015a. Not all are free-living: high-throughput DNA metabarcoding reveals a diverse community of protists parasitizing soil metazoa. Molecular Ecology 24, 4556e4569.

Geisen, S., Rosengarten, J., Koller, R., Mulder, C., Urich, T., Bonkowski, M., 2015b. Pack hunting by a common soil amoeba on nematodes. Environmental Microbiology 17, 4538e4546.

Geisen, S., Tveit, A.T., Clark, I.M., Richter, A., Svenning, M.M., Bonkowski, M., Urich, T., 2015c. Metatranscriptomic census of active protists in soils. The ISME Journal 9, 2178e2190.

Glücksman, E., Bell, T., Griffiths, R.I., Bass, D., 2010. Closely related protist strains have different grazing impacts on natural bacterial communities. Environmental Microbiology 12, 3105e3113.

Greub, G., Raoult, D., 2004. Microorganisms resistant to free-living amoebae. Clinical Microbiology Reviews 17, 413e433.

Griffiths, B.S., Römbke, J., Schmelz, R.M., Scheffczyk, A., Faber, J.H., Bloem, J., Pérès, G., Cluzeau, D., Chabbi, A., Suhadolc, M., Sousa, J.P., Martins da Silva, P., Carvalho, F., Mendes, S., Morais, P., Francisco, R., Pereira, C., Bonkowski, M., Geisen, S., Bardgett, R.D., de Vries, F.T., Bolger, T., Dirilgen, T., Schmidt, O., Winding, A., Hendriksen, N.B., Johansen, A., Philippot, L., Plassart, P., Bru, D., Thomson, B., Griffiths, R.I., Bailey, M.J., Keith, A., Rutgers, M., Mulder, C., Hannula, S.E., Creamer, R., Stone, D., 2016. Selecting cost effective and policy-relevant biological indicators for European monitoring of soil biodiversity and ecosystem function. Ecological Indicators 69, 213e223.

Grossmann, L., Jensen, M., Heider, D., Jost, S., Glucksman, E., Hartikainen, H., Mahamdallie, S.S., Gardner, M., Hoffmann, D., Bass, D., Boenigk, J., 2016. Protistan community analysis: key findings of a large-scale molecular sampling. ISME Journal 10, 2269e2279.

Gupta, V.V., Yeates, G., 1997. Soil Microfauna as Bioindicators of Soil Health. CAB International.

Gupta, V.V.S.R., Germida, J.J., 1988. Distribution of microbial biomass and its activity in different soil aggregate size classes as affected by cultivation. Soil Biology and Biochemistry 20, 777e786.

Haeckel, E., 1866. Generelle Morphologie der Organismen. G. Reimer, Berlin.

Jacquiod, S., Stenbaek, J., Santos, S.S., Winding, A., Sorensen, S.J., Prieme, A., 2016. Metagenomes provide valuable comparative information on soil microeukaryotes. Research in Microbiology 167, 436e450.

Jassey, V.E.J., Chiapusio, G., Binet, P., Buttler, A., Laggoun-Défarge, F., Delarue, F., Bernard, N., Mitchell, E.A.D., Toussaint, M.-L., Francez, A.-J., Gilbert, D., 2013. Above- and belowground linkages in Sphagnum peatland: climate warming affects plant-microbial interactions. Global Change Biology 19, 811e823.

Jassey, V.E.J., Shimano, S., Dupuy, C., Toussaint, M.L., Gilbert, D., 2012. Characterizing the feeding habits of the testate amoebae Hyalosphenia papilio and Nebela tincta along a narrow "fen-bog" gradient using digestive vacuole content and 13C and 15N isotopic analyses. Protist 163, 451e464.

Jassey, V.E.J., Signarbieux, C., H€attenschwiler, S., Bragazza, L., Buttler, A., Delarue, F., Fournier, B., Gilbert, D., Laggoun-Défarge, F., Lara, E., T. E. Mills, R., Mitchell, E.A.D., Payne, R.J., Robroek, B.J.M., 2015. An unexpected role for mixotrophs in the response of peatland carbon cycling to climate warming. Scientific Reports 5, 16931.

Jones, S.E., Lennon, J.T., 2010. Dormancy contributes to the maintenance of microbial diversity. Proceedings of the National Academy of Sciences 107, 5881e5886.

Jousset, A., 2012. Ecological and evolutive implications of bacterial defences against predators. Environmental Microbiology 14, 1830e1843.

Jousset, A., Lara, E., Wall, L.G., Valverde, C., 2006. Secondary metabolites help biocontrol strain Pseudomonas fluorescens CHA0 to escape protozoan grazing. Applied and Environmental Microbiology 72, 7083e7090.

Keyes, S.D., Daly, K.R., Gostling, N.J., Jones, D.L., Talboys, P., Pinzer, B.R., Boardman, R., Sinclair, I., Marchant, A., Roose, T., 2013. High resolution synchrotron imaging of wheat root hairs growing in soil and image based modelling of phosphate uptake. New Phytologist 198, 1023e1029.

Kosakyan, A., Gomaa, F., Lara, E., Lahr, D.J.G., 2016. Current and future perspectives on the systematics, taxonomy and nomenclature of testate amoebae. European Journal of Protistology 55, 105e117.

Lamentowicz, M., Gałka, M., Lamentowicz, Ł., Obremska, M., Kühl, N., Lücke, A., Jassey, V., 2015. Climate change over the last 4000 years in a Baltic bog in northern Poland revealed by a trait-based approach, biotic proxies, and stable isotopes. Palaeogeography Palaeoclimatology Palaeoecology 418, 261e277.

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G.W., Prosser, J.I., Schuster, S.C., Schleper, C., 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442, 806e809.

Lentendu, G., Wubet, T., Chatzinotas, A., Wilhelm, C., Buscot, F., Schlegel, M., 2014. Effects of long-term differential fertilization on eukaryotic microbial communities in an arable soil: a multiple barcoding approach. Molecular Ecology 23, 3341e3355.

López-García, P., Moreira, D., 2015. Open questions on the origin of eukaryotes. Trends in Ecology and Evolution 30, 697e708.

Lynn, D., 2008. The Ciliated protozoa: Characterization, Classification, and Guide to the Literature, third ed. Springer Science & Business Media, Dordrecht.

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.M., Mitchell, E.A.D., Seppey, C.V.W., Egge, E., Wirth, R., Trueba, G., Dunthorn, M., 2017. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. Nature Ecology and Evolution 1, 0091.

Marcisz, K., Lamentowicz, Ł., Słowi_nska, S., Słowi_nski, M., Muszak, W., Lamentowicz, M., 2014. Seasonal changes in Sphagnum peatland testate amoeba communities along a hydrological gradient. European Journal of Protistology 50, 445e455.

Meyer, C., Gilbert, D., Gillet, F., Moskura, M., Franchi, M., Bernard, N., 2012. Using "bryophytes and their associated testate amoeba" microsystems as indicators of atmospheric pollution. Ecological Indicators 13, 144e151.

Mitchell, E.A.D., 2015. Pack hunting by minute soil testate amoebae: nematode hell is a naturalist's paradise. Environmental Microbiology 17, 4145e4147.

Mitchell, E.A.D., Charman, D.J., Warner, B.G., 2008. Testate amoebae analysis in ecological and paleoecological studies of wetlands: past, present and future. Biodiversity & Conservation 17, 2115e2137.

Molmeret, M., Horn, M., Wagner, M., Santic, M., Kwaik, Y.A., 2005. Amoebae as training grounds for intracellular bacterial pathogens. Applied and Environmental Microbiology 71, 20e28.

Murase, J., Hordijk, K., Tayasu, I., Bodelier, P.L.E., 2011. Strain-specific incorporation of methanotrophic biomass into eukaryotic grazers in a rice field soil revealed by PLFA-SIP. FEMS Microbiology Ecology 75, 284e290.

Neuhauser, S., Kirchmair, M., Bulman, S., Bass, D., 2014. Cross-kingdom host shifts of phytomyxid parasites. BMC Evolutionary Biology 14, 33.

Old, K.M., Chakraborty, S., 1986. Mycophagous soil amoebae: their biology and significance in the ecology of soil-borne plant pathogens. Progress in Protistology 1, 163e194.

Page, F.C., 1977. The genus Thecamoeba (Protozoa, Gymnamoebia) species distinctions, locomotive morphology, and protozoan prey. Journal of Natural History 11, 25e63.

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S.S., Cepicka, I., Decelle, J., Dunthorn, M., Fiore-Donno, A.M., Gile, G.H., Holzmann, M., Jahn, R., Jirk_u, M., Keeling, P.J., Kostka, M., Kudryavtsev, A.A., Lara, E., Lukes, J., Mann, D.G., Mitchell, E.A.D., Nitsche, F., Romeralo, M., Saunders, G.W., Simpson, A.G.B., Smirnov, A.V., Spouge, J.L., Stern, R.F., Stoeck, T., Zimmermann, J., Schindel, D., de Vargas, C., 2012. CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. Plos Biology 10, e1001419.

Payne, R.J., 2013. Seven reasons why protists make useful bioindicators. Acta Protozoologica 52, 105.

Payne, R.J., Jassey, V.E., Leith, I.D., Sheppard, L.J., Dise, N.B., Gilbert, D., 2013. Ammonia exposure promotes algal biomass in an ombrotrophic peatland. Soil Biology and Biochemistry 57, 936e938.

Payne, R.J., Malysheva, E., Tsyganov, A., Pampura, T., Novenko, E., Volkova, E., Babeshko, K., Mazei, Y., 2015. A multi-proxy record of Holocene environmental change, peatland development and carbon accumulation from Staroselsky Moch peatland, Russia. The Holocene 26, 314e326.

Payne, R.J., Mitchell, E.A., Nguyen-Viet, H., Gilbert, D., 2012. Can pollution bias peatland paleoclimate reconstruction? Quaternary Research 78, 170e173.

Prosser, J.I., 2015. Dispersing misconceptions and identifying opportunities for the use of 'omics' in soil microbial ecology. Nature Reviews Microbiology 13, 439e446.

Raoult, D., Boyer, M., 2010. Amoebae as genitors and reservoirs of giant viruses. Intervirology 53, 321e329.

Rønn, R., Gavito, M., Larsen, J., Jakobsen, I., Frederiksen, H., Christensen, S., 2002. Response of free-living soil protozoa and microorganisms to elevated atmospheric CO2 and presence of mycorrhiza. Soil Biology and Biochemistry 34, 923e932.

Rosenberg, K., Bertaux, J., Krome, K., Hartmann, A., Scheu, S., Bonkowski, M., 2009. Soil amoebae rapidly change bacterial community composition in the rhizosphere of Arabidopsis thaliana. ISME Journal 3, 675e684.

Schardl, C.L., Craven, K.D., 2003. Interspecific hybridization in plant-associated fungi and oomycetes: a review. Molecular Ecology 12, 2861e2873.

Schmidt, O., Dyckmans, J., Schrader, S., 2016. Photoautotrophic microorganisms as a carbon source for temperate soil invertebrates. Biology Letters 12.

Schulz-Bohm, K., Geisen, S., Wubs, E.R.J., Song, C., de Boer, W., Garbeva, P., 2017. The prey's scent - volatile organic compound mediated interactions between soil bacteria and their protist predators. ISME Journal 11, 817e820.

Schuster, F.L., 2002. Cultivation of pathogenic and opportunistic free-living amebas. Clinical Microbiology Reviews 15, 342e354.

Seddon, A.W.R., Mackay, A.W., Baker, A.G., Birks, H.J.B., Breman, E., Buck, C.E., Ellis, E.C., Froyd, C.A., Gill, J.L., Gillson, L., Johnson, E.A., Jones, V.J., Juggins, S., Macias-Fauria, M., Mills, K., Morris, J.L., Nogués-Bravo, D., Punyasena, S.W., Roland, T.P., Tanentzap, A.J., Willis, K.J., Aberhan, M., van Asperen, E.N., Austin, W.E.N., Battarbee, R.W., Bhagwat, S., Belanger, C.L., Bennett, K.D., Birks, H.H., Bronk Ramsey, C., Brooks, S.J., de Bruyn, M., Butler, P.G., Chambers, F.M., Clarke, S.J., Davies, A.L., Dearing, J.A., Ezard, T.H.G., Feurdean, A., Flower, R.J., Gell, P., Hausmann, S., Hogan, E.J., Hopkins, M.J., Jeffers, E.S., Korhola, A.A., Marchant, R., Kiefer, T., Lamentowicz, M., Larocque-Tobler, I., López-Merino, L., Liow, L.H., McGowan, S., Miller, J.H., Montoya, E., Morton, O., Nogué, S., Onoufriou, C., Boush, L.P., Rodriguez-Sanchez, F., Rose, N.L., Sayer, C.D., Shaw, H.E., Payne, R., Simpson, G., Sohar, K., Whitehouse, N.J., Williams, J.W., Witkowski, A., 2014. Looking forward through the past: identification of 50 priority research questions in palaeoecology. Journal of Ecology 102, 256e267.

Shmakova, L., Bondarenko, N., Smirnov, A., 2016. Viable species of Flamella (Amoebozoa: variosea) isolated from ancient arctic permafrost sediments. Protist 167, 13e30.

Siddiqui, R., Ahmed Khan, N., 2012. Biology and pathogenesis of Acanthamoeba. Parasites and Vectors 5, 1e13.

Singer, D., Lara, E., Steciow, M.M., Seppey, C.V.W., Paredes, N., Pillonel, A., Oszako, T., Belbahri, L., 2016. High-throughput sequencing reveals diverse oomycete communities in oligotrophic peat bog micro-habitat. Fungal Ecology 23, 42e47.

Smirnov, A.V., Brown, S., 2004. Guide to the methods of study and identification of soil gymnamoebae. Protistology 3, 148e190.

Spiegel, F.W., Stephenson, S.L., Keller, H.W., Moore, D.L., Cavender, J.C., 2004. Biodiversity of fungi: inventory and monitoring methods. In: Mueller, G.M., Foster, M.S., Bills, G.F. (Eds.), Biodiversity of Fungi, Inventory and Monitoring Methods. Elsevier Academic Press, Burlington, MA, pp. 547e576.

Stockdale, A., Davison, W., Zhang, H., 2009. Micro-scale biogeochemical heterogeneity in sediments: a review of available technology and observed evidence. Earth-Science Reviews 92, 81e97.

Sutherland, W.J., Freckleton, R.P., Godfray, H.C.J., Beissinger, S.R., Benton, T., Cameron, D.D., Carmel, Y., Coomes, D.A., Coulson, T., Emmerson, M.C., Hails, R.S., Hays, G.C., Hodgson, D.J., Hutchings, M.J., Johnson, D., Jones, J.P.G., Keeling, M.J., Kokko, H., Kunin, W.E., Lambin, X., Lewis, O.T., Malhi, Y., Mieszkowska, N., Milner-Gulland, E.J., Norris, K., Phillimore, A.B., Purves, D.W., Reid, J.M., Reuman, D.C., Thompson, K., Travis, J.M.J., Turnbull, L.A., Wardle, D.A., Wiegand, T., 2013. Identification of 100 fundamental ecological questions. Journal of Ecology 101, 58e67.

Tice, A.K., Shadwick, L.L., Fiore-Donno, A.M., Geisen, S., Kang, S., Schuler, G.A., Spiegel, F.W.,Wilkinson, K.A., Bonkowski, M., Dumack, K., Lahr, D.J., Voelcker, E., Clauss, S., Zhang, J., Brown, M.W., 2016. Expansion of the molecular and morphological diversity of Acanthamoebidae (Centramoebida, Amoebozoa) and identification of a novel life cycle type within the group. Biology Direct 11, 69.

Tikhonenkov, D.V., Mazei, Y.A., Embulaeva, E.A., 2010. Effect of ecosystem type on soil heterotrophic flagellate communities under forest-steppe conditions. Protistology 6.

Treonis, A.M., Lussenhop, J.F., 1997. Rapid response of soil protozoa to elevated CO2. Biology and Fertility of Soils 25, 60e62.

Tsyganov, A.N., Nijs, I., Beyens, L., 2011. Does climate warming stimulate or inhibit soil protist communities? A test on testate amoebae in high-arctic tundra with free-air temperature increase. Protist 162, 237e248.

Turner, T.E., Swindles, G.T., Roucoux, K.H., 2014. Late Holocene ecohydrological and carbon dynamics of a UK raised bog: impact of human activity and climate change. Quaternary Science Reviews 84, 65e85.

Wilkinson, D.M., 2008. Testate amoebae and nutrient cycling: peering into the black box of soil ecology. Trends in Ecology & Evolution 23, 596e599.

Figures

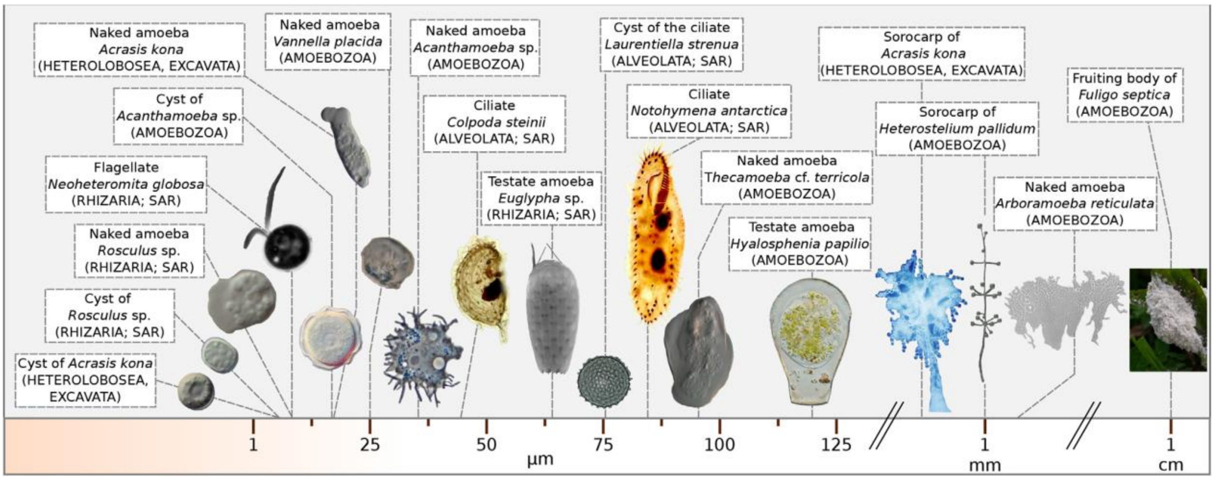

Figure 1: Common free-living soil protists as visualized by size (lengths), morphology and phylogenetic affiliation. Note, soil protists belong to a wide range of supergroups (in brackets), whereas animals are only placed in the supergroup Opisthokonta. Furthermore, soil protists span a much wider size range than commonly assumed. With the exception of ciliates, morphogroups are not phylogenetically conserved and are placed in different eukaryotic supergroups. Most soil protists can occur in different life forms including active form (amoeba, flagellate, ciliate), but most form cysts, while some can form special reproduction structures (sorocarps and fruiting bodies).

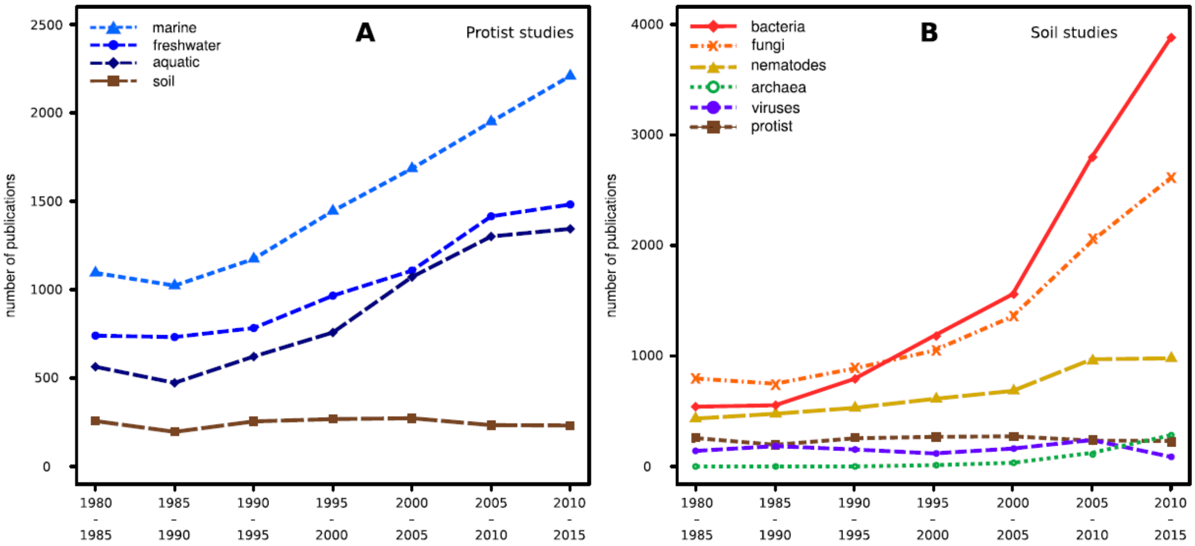

Figure 2: (a) Overview of studies specifically mentioning protists in the title in 5-year intervals since 1980. Soil studies represent only about a fraction of aquatic studies (separated into freshwater, marine and those that more broadly indicate aquatic) showing a strong increase in protist research in aquatic, but not soil environments; (b) Comparison of soil studies specifically mentioning protists in the title with those on other micro-sized organisms including viruses (blue filled circles), archaea (green open circles), bacteria (red diamonds), fungi (orange crosses) and nematodes (green triangles). See Supplementary Methods for details on the search. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Tables

i. I Morphology, Phylogeny, Taxonomy, Evolution and Physiology

1 How long can protists survive in an encysted form? What are the tolerances of (encysted) protists to stress and what is the importance of cysts for ecosystem resilience?
2 How much morphological and genetic variability exists within soil protists?
3 How do species that occur in both aquatic and soil systems adapt to differing demands?
4 What are the phylogenetic relations of true soil to aquatic protist taxa and how often have soils been colonized by aquatic protists and vice versa?
5 How widespread is sex in soil protists?

ii. II Diversity, Community Composition and Biogeography

6 What is the real diversity and community structure of soil protists in different systems (e.g. soils, rhizosphere, (plant) endosphere)?
7 How similar are the diversity patterns of soil protists and other soil biota along ecological gradients, and to what extent do different environmental factors shape their respective diversity?
8 What abiotic environmental factors influence the distribution and community composition of protists, and how?
9 How cosmopolitan are protists and how many endemic soil protist species are there?
10 What are the dominant groups of soil protists in terms of turnover, abundance and biomass?

iii. III Interactions among Protists and other Organisms

11 How do protist taxa affect the composition of the soil microbiome and what other important interactions take place?
12 What are the biotic interactions of soil protists with other taxonomic groups, and how are protists linked within the soil food web?
13 What is the relative contribution of nutrient cycling (i.e. the microbial loop) versus modification of the rhizosphere microbiome in protist-induced stimulation of plant growth?
14 What are the mechanisms by which individual soil protist species affect plant performance, and do those mechanisms differ between plant species?
15 What is the impact of protists on the community functioning of other soil microbes?

iv. IV Functions of Protists

16 What is the importance of soil protists in biogeochemical cycling?
17 How much functional redundancy is there in the soil protist community?
18 Does increased protist diversity affect ecosystem functioning?
19 What is the comparative importance of eukaryotic microbes vs. prokaryotes in driving key soil processes?
20 Which individual functions are performed by distinct groups, and what is the entire functional diversity of soil protists?

v. V Global Change, Bioindicators and Applications

21 How do changing climatic patterns affect the diversity of, community structure of and ecosystem services provided by soil protists?
22 Which protist clades can be used as bioindicators to assess soil properties, ecosystem state, and anthropogenic impacts? How could this be implemented?
23 Why are some species more sensitive to environmental change than others, why do some respond faster to environmental factors?
24 How can protists be used for nutrient mobilization and biocontrol in cropping systems?
25 What is the importance of soil protists for biodiversity conservation and ecosystem management and restoration? Should we protect particular species or habitats?

vi. VI Methodology

26 What is the most practical taxonomic unit to measure protist diversity?
27 How can we standardize and calibrate cultivation based and molecular methods to reliably quantify soil protist abundance, diversity and activity?
28 How should sampling be performed to adequately evaluate soil protist diversity?
29 At what scales (temporal, spatial/physical, morphological, phylogenetic) should we study protists to fully understand their diversity and function in soil; which one should be prioritized?
30 How can we infer functional traits of soil protists based on morphology or phylogenetic affiliation, and what taxonomic resolution is needed?

Curriculum Vitae

Name: Quentin Blandenier
Date of birth: 31 janvier 1991
Nationality: Swiss
Civil status: Single
Place of origin: Villiers (NE), Switzerland
E-mail: quentin.blandenier@gmail.com
Personal adresse: Roche 6, 2012 Auvernier, Switzerland
ORCID: https://orcid.org/0000-0002-4297-0262
ScholaarID: https://scholar.google.com/citations?hl=en&user=xe1pKrUAAAAJ

Formation

Jan. 2016 – Apr. 2020 Doctor of Philosophy in Biological Science (Ph.D), Laboratory of Biodiversity, University of Neuchâtel, Switzerland and Botanical Garden of Madrid, Spain. Director Dr. Enrique Lara. Thesis entitled: *Exploration and characterization of Amoebozoa diversity and investigation of their diversity patterns at regional and global scales*

Sep. 2013 – Sep. 2015 Master of Science in Biogeosciences (M.Sc), Universities of Neuchâtel and Lausanne. Master thesis entitled: *"Design and evaluation of a new mitochondrial marker for Arcellinida (Amoebozoa)."*

Sep. 2010 – Jul. 2013 Bachelor of Science in Biology (B.Sc), University of Neuchâtel, Switzerland

Work experience

Jan. 2016 – Apr. 2020 PhD thesis in microbiology and ecology under the direction of Dr. Enrique Lara and co-supervison of Prof. Edward Mitchell.

Sep. 2014 – Sep. 2015 Master thesis performed in both Laboratory of Soil Biodiversity, University of Neuchâtel, and Laboratory of Evolutionary Protistology, University of Sao Paulo under the joint supervision of Dr. Enrique Lara, Prof. Daniel JG. Lahr and Prof. Edward AD. Mitchell.

Jul. 2012 – Sep. 2014 Intership in the Laboratory of Biodiversity, University of Neuchâtel under the supervision of Dr. Enrique Lara.

Approved research project

Aug 2014 Financial support (1'400 CHF) by the Fonds Wüthrich and Matthey-Dupraz for the master thesis ''Making of a whole mitochondrial genetic map of some amoebozoa species".

Mentoring experience

Jan. 2016 – Jan. 2019	Supervision of several internships for barcoding methods and phylogenetic analysis: Dr. Andrew Macumber, Fanny Kupferschmid, Nikola Bankov, and Angela O'Donnell.

Teaching activities

Sep. 2012 – Jun. 2018	Practical courses for bachelor and master students at the University of Neuchâtel: - Protistes et invertébrés (B.Sc.) - TP de microbiologie (B.Sc.) - Invertébrés II (B.Sc.) - Apprentissage par problèmes (APP) (B.Sc.) - TP Biologie du sol (M.Sc.)

Reviewing activities

- Microbial Ecology, 2x,
- The Journal of Eukaryotic Microbiology, 1x
- European Journal of Protistology, 2x

Society Membership

Since Mar. 2017	Member of the International Society of Protistologists (ISOP)
Since Feb. 2017	Member of the German Society for Protozoology (DGP)

Organisation of conferences

14-16.02.2018	18th annual meeting of Swiss student in organismal biology (Biology18) (in charge of fundraising and of the gala dinner organization)
23-26.02.2016	35th annual meeting of the German Society for Protozoology (main secretary of the comity and in charge of fundraising)

Awards

27.02- 02.03.2018	37th annual meeting of the German Society for Protozoology, Cologne, Germany. Third best oral presentation
21-24.02.2017	36th annual meeting of the German Society for Protozoology, Meissen, Germany. Best oral presentation

Skills

Languages	French – Mother tongue English – Highly proficient in spoken and written German – High-school level Portuguese – Basic knowledge

Laboratory	Microscopical methods – Sample preparation, cells manipulation and isolation, high Light and SEM quality pictures Molecular methods – DNA extraction, PCR, Cloning/sequencing
Computer	Microsoft Office, Phylogenetic analysis, Linux, NGS Data analysis

Published articles

1. Macumber AL, **Blandenier Q**, Todorov M, Duckert C, Lara E, Lahr DJG, Mitchell EAD, Roe HM. Phylogenetic divergence within the Arcellinida (Amoebozoa) is congruent with test size and metabolism type. *European Journal of Protistology (in press).*
2. Seppey CVW, Broennimann O, Buri A, Yashiro E, Pinto-Figueroa E, Singer D, **Blandenier Q**, Mitchell EAD, Niculita Hirzel H, Guisan A, Lara E. Soil protist diversity in the Swiss western Alps is better predicted by topo-climatic than by edaphic variables. *Journal of Biogeography (in press).*
3. Singer D, Mitchell EAD, Payne RJ, **Blandenier Q**, Duckert C, Fernández LD, Fournier B, Hernández CE, Granath G, Rydin H, Bragazza L, Koronatova NG, Goia I, Harris LI, Kajukało K, Kosakyan A, Lamentowicz M, Kosykh NP, Vellak K, Lara E. **2019**. Dispersal limitations and historical factors determine the biogeography of specialized terrestrial protists. *Molecular Ecology* 28: 3089-3100. https://doi.org/10.1111/mec.15117.
4. **Blandenier Q***, & Duckert C*, Kupferschmid FAL, Kosakyan A, Mitchell EAD, Lara E, Singer D. **2018**. En garde! Redefinition of *Nebela militaris* (Arcellinida, Hyalospheniidae) and erection of *Alabasta* gen. nov. *European Journal of Protistology* 66: 156-165. https://doi.org/10.1016/j.ejop.2018.08.005. http://doc.rero.ch/record/323497/ (*These authors contributed equally to the paper.)
5. Geisen S, Mitchell EAD, Wilkinson DM, Adl S, Bonkowski M, Brown MW, Fiore-Donno AM, Heger TJ, Jassey VEJ, Krashevska V, Lahr DJG, Marcisz K, Mulot M, Payne R, Singer D, Anderson OR, Charman DJ, Ekelund F, Griffiths BS, Rønn R, Smirnov A, Bass D, Belbahri L, Berney C, **Blandenier Q**, Chatzinotas A, Clarholm M, Dunthorn M, Feest A, Fernandez LD, Foissner W, Fournier B, Gentekaki E, Hajek M, Helder J, Jousset A, Koller R, Kumar S, La Terza A, Lamentowicz M, Mazei Y, Santos SS, Seppey CVW, Spiegel FW, Walochnik J, Winding A, Lara E. **2017**. Soil protistology rebooted: 30 fundamental questions to start with. *Soil Biology and Biochemistry* 111: 94-103. https://doi.org/10.1016/j.soilbio.2017.04.001. http://doc.rero.ch/record/288493
6. **Blandenier Q**, Lara E, Mitchell EAD, Alcantar DMC, Siemensma FJ, Todorov M, Lahr DJG. **2017**. NAD9/NAD7 (mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene) — A new "Holy Grail" phylogenetic and DNA-barcoding marker for Arcellinida (Amoebozoa)? *European Journal of Protistology* 58: https://doi.org/10.1016/j.ejop.2016.12.002. http://doc.rero.ch/record/288513
7. **Blandenier Q**, Seppey C, Singer D, Vlimant M, Simon A, Duckert C, Lara E. **2016**. *Mycamoeba gemmipara* nov. gen., nov. sp., the First Cultured Member of the Environmental Dermamoebidae Clade LKM74 and its Unusual Life Cycle. *Journal of Eukaryotic Microbiology* 64: 257-65. https://doi.org/10.1111/jeu.12357. http://doc.rero.ch/record/288438/files/?ln=de

Contributions to international meetings

29.07-02.08.**2019**	European Congress of Protistology (ECOP8), Rome, Italia. **(Talk)** *High Throughput Sequencing approaches focusing specifically on genus* Nebela *allow testing the influence of the altitude on the community composition*. **Blandenier Q**, Singer D, Jauslin R, Duckert C, Pillonel A, Baur M, Mitchell EAD, Lara E.
19-22.02.**2019**	Deutsche Gesellschaft für Protozoologie (DGP38), Vienna, Austria. **(Talk)** *High Throughput Sequencing approaches focusing specifically on genus* Nebela *allow testing the influence of the altitude on the community composition*. **Blandenier Q,** Singer D, Jauslin R, Duckert C, Pillonel A, Baur M, Mitchell EAD, Lara E.
10-14.09.**2018**	International Society of Testate Amoeba (ISTA9), Belfast, North Ireland. **(Talk)** *Exploring the diversity of Hyalospheniids testate amoebae (Amoebozoa) from regional to sub-continental scales*. **Blandenier Q**, Singer D, Duckert C, Mitchell EAD, Kosakyan A, Todorov M, Lahr DJG, Lara E.
28.02-02.03.**2018**	Deutsche Gesellschaft für Protozoologie (DGP37), Cologne, Germany. **(Talk)** *Exploring the diversity of Hyalospheniids testate amoebae (Amoebozoa) from regional to sub-continental scales*. **Blandenier Q**, Singer D, Duckert C, Mitchell EAD, Kosakyan A, Todorov M, Lahr DJG, Lara E. **3rd Award.**
30.07-04.08.**2017**	International Congress of Protistology (ISOP15), Prague, Czech Republic. **(Talk)** *A molecular survey of genus* Nebela *s. str. diversity, ecology and geographical distribution*. **Blandenier Q**, Singer D, Duckert C, Mitchell EAD, Kosakyan A, Lahr DJG, Lara E.
	(Poster) Mycamoeba gemmipara *Nov. gen., Nov. sp., The first cultured member of the environmental Dermamoebidae clade LKM74 and its unusual life cycle*. **Blandenier Q**, Seppey C, Singer D, Vlimant M, Simon A, Duckert C, Lara E.
21-24.02.**2017**	Deutsche Gesellschaft für Protozoologie (DGP36), Meissen, Germany. **(Talk)** Mycamoeba gemmipara *Nov. gen., Nov. sp., The first cultured member of the environmental Dermamoebidae clade LKM74 and its unusual life cycle*. **Blandenier Q**, Seppey C, Singer D, Vlimant M, Simon A, Duckert C, Lara E. **1st Award.**
12-15.09.**2016**	International Society of Testate Amoeba (ISTA8), Ilhabela, Brazil. **(Talk)** *Exploring the mitochondrial genomes of Amoebozoa in search of novel molecular markers: the emergence of a new barcode for Arcellinida*. **Blandenier Q**, Lara E, Mitchell EAD, Alcantar DMC, Siemensma FJ, Todorov M, Lahr DJG.

23-26.02.**2016** Deutsche Gesellschaft für Protozoologie (DGP35), Saignelegier, Switzerland. **(Talk)** *Exploring the mitochondrial genomes of Amoebozoa in search of novel molecular markers: the emergence of a new barcode for Arcellinida*. **Blandenier Q**, Lara E, Mitchell EAD, Alcantar DMC, Siemensma FJ, Todorov M, Lahr DJG.

05-10.11.**2015** European Congress of Protistology (ECOP7), Sevilla, Spain. **(Poster)** *NAD9/NAD7 the emergence of a new barcode for Arcellinida*. **Blandenier Q**, Lara E, Mitchell EAD, Alcantar DMC, Siemensma FJ, Todorov M, Lahr DJG.