Article Type: Articles

Running Head: Distribution of protist cryptic species

Environmental filtering and phylogenetic clustering correlate with the distribution patterns of cryptic protist species

David Singer[1], Anush Kosakyan[1,2,3], Christophe V.W. Seppey[1,4], Amandine Pillonel[1], Leonardo D. Fernández[1,5,6], Diego Fontaneto[7], Edward A.D. Mitchell[1,8], Enrique Lara[1, 9]

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

[2] Department of Zoology, Institute of Biosciences, University of Sao Paulo, Sao Paulo, Brazil

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

[4]Microorganisms and Plants Group, Department of Arctic and Marine Biology, Faculty of Biosciences, Fisheries and Economics; University of Tromso; Framstredet 39, 9037 Tromso; Norway.

[5]Laboratorio de Ecología Evolutiva y Filoinformática, Departamento de Zoología, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Barrio Universitario s/n, Casilla 160-C, Concepción, Chile

[6]Centro de Investigación en Recursos Naturales y Sustentabilidad (CIRENYS), Universidad Bernardo O'Higgins, Avenida Viel 1497, Santiago, Chile

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ecy.02161

[7]National Research Council of Italy, Institute of Ecosystem Study, 28922 Verbania Pallanza, Italy

[8]Jardin Botanique de Neuchâtel, Chemin du Perthuis-du-Sault 58, CH-2000 Neuchâtel, Switzerland

[9]Real Jardín Botánico, CSIC, Plaza Murillo 2, ES 28014 Madrid, Spain

Corresponding author: David Singer (David.singer@unine.ch)

Abstract

The community composition of any group of organisms should theoretically be determined by a combination of assembly processes including resource partitioning, competition, environmental filtering, and phylogenetic legacy. Environmental DNA studies have revealed a huge diversity of protists in all environments, raising questions about the ecological significance of such diversity and the degree to which they obey to the same rules as macroscopic organisms. The fast-growing cultivable protist species on which hypotheses are usually experimentally tested represent only a minority of the protist diversity. Addressing these questions for the lesser known majority can only be inferred through observational studies.

We conducted an environmental DNA survey of the genus *Nebela*, a group of closely related testate (shelled) amoeba species, in different habitats within *Sphagnum*-dominated peatlands. Identification based on the mitochondrial cytochrome c oxidase 1 gene, allowed species-level resolution as well as phylogenetic reconstruction.

Community composition varied strongly across habitats and associated environmental gradients. Species showed little overlap in their realized niche, suggesting resource partitioning, and a strong influence of environmental filtering driving community composition. Furthermore, phylogenetic clustering was observed in the most nitrogen-poor samples, supporting phylogenetic inheritance of adaptations in the group of *N. guttata.*

This study showed that the studied free-living unicellular eukaryotes follow to community assembly rules similar to those known to determine plant and animal communities; the same may be true for much of the huge functional and taxonomic diversity of protists.

Keywords

Environmental filtering, phylogenetic clustering/over-dispersion, Nearest Taxon Index (NTI), niche partitioning, uncultivable protists, *Sphagnum*-dominated peatlands

Introduction

Understanding the rules that determine the composition of communities is critical to the assessment and conservation of biodiversity (Colwell and Coddington 1994). Organismal diversity and the theoretical concepts to understand the drivers of biodiversity have historically been based on macroscopic organisms (Wilkinson 1998) but microorganism models are now increasingly used in ecology (Fukami and Morin 2003, Altermatt et al. 2015).

High throughput sequencing (HTS) studies of aquatic and terrestrial habitats have revealed the existence of hundreds of thousands of previously unknown phylotypes (or operational taxonomic units OTUs) within known clades as well as revealing highly diverse novel "environmental" clades (de Vargas et al. 2015, Grossmann et al. 2016, Mahé et al. 2017). These new molecular tools challenge our vision of global biodiversity and its drivers (Van Hannen et al. 1999, López-García et al. 2001). This is leading to a re-assessment of the diversity and community ecology of micro-organisms (e.g. (Ofiteru et al. 2010, Lentendu et al. 2014, Zhou et al. 2014, Geisen et al. 2015, Worden et al. 2015) and provides opportunities for testing the broader validity of theoretical concepts across all domains of life.

At the fine taxonomic level, closely-related species, which are often identified only by DNA barcoding (i.e. cryptic species) (Hebert et al. 2004), offer valuable test cases to assess how species traits (e.g. ecological optima, behavior) contribute to speciation mechanisms. The taxonomic validity of cryptic species can be further assessed by comparing phylogenetic and ecological data, such as shown for Amazonian trees (Esteves and Vicentini 2013), bats (Santos et al. 2014) and salamanders (Rissler and Apodaca 2007).

Cryptic diversity has also been found in many protist taxa (Kooistra et al. 2005, Kosakyan et al. 2012). Unrecognized cryptic species lead to erroneous biodiversity estimates and mask potentially contrasted distribution patterns from the global (biogeography) to local (community assembly processes) scales. Molecular assessment of protist diversity may shed new light on their community ecology. With these tools, it is now possible to assess not only the fine-scale distribution patterns at the relevant taxonomic resolution but also to explore the mechanisms driving speciation as hitherto done only for macroscopic organisms (Rissler and Apodaca 2007, Esteves and Vicentini 2013, Santos et al. 2014, Janzen et al. 2017).

Historically, the coexistence in a given homogeneous environment of a higher protist diversity than would be predicted based on the available resources was formulated by Hutchinson (Hutchinson 1961) as the “paradox of the plankton”. Several explanations have been given, most often based on chaotic oscillations of the different populations, which rely on fast population growth and high mortality rates typically caused by viruses or parasites (Scheffer et al. 2003). However, few protists are fast growers, and they rather represent a minority in the protist realm (del Campo et al. 2016). Thus, chaotic effects are not likely to shape the composition of communities of slow-growing protists. The composition of such communities can be rather seen as the product of immigration, with persistence then selected by deterministic forces due to both biotic (interactions among organisms) and abiotic (environmental filters) processes, plus random drift and extinction (Carlson et al. 2010, Hanson et al. 2012), although some studies suggest that patterns are consistent with neutral models (Wanner and Xylander 2005).

Niche breadth as well as the stability and selectivity of ecosystems are known drivers of community assembly (Weiner et al. 2012). Ecological tolerance, which can be considered as a reliable proxy for niche breadth (Futuyma and Moreno 1988), varies considerably amongst microbial eukaryotes, including within a single genus. For instance, contrasted tolerance to low pH values has been shown between genetically different (but morphologically identical) strains of the fast growing ciliate *Meseres corlissi* (Weisse et al. 2007).

Sphagnum-dominated peatlands are characterized by stable but harsh conditions (oligotrophy, low pH, recalcitrant organic matter), causing a strong selective pressure resulting in the existence of highly specific biotic communities (Dedysh et al. 2006, Steinberg et al. 2006) characterized by low growth and turnover. Spatial heterogeneity is high, with chemical and physical conditions varying strongly in relation to topography (e.g. hummock-hollow gradient), and influencing vegetation (Okland et al. 2008, Andersen et al. 2011), defining well-defined and easily identified habitats. Thus, *Sphagnum*-dominated peatlands are good model ecosystems to study the community ecology of slow-growing organisms including protists such as testate amoebae.

Testate amoebae are a polyphyletic group of amoeboid protists characterized by a shell (or test), and relatively slow growth rates comparable to those of microscopic Metazoa (Foissner 1999). Among testate amoebae, generation time of *Trinema lineare*, one of the smallest species, was 1.9 days (Schönborn 1975), while only 19 generations per year have been reported for the testate amoeba *Phryganella acropodia* (Schönborn 1992). Members of the genus *Nebela* (*sensu* (Kosakyan et al. 2016) are among the most frequent testate amoebae in *Sphagnum*-dominated peatlands (Gilbert and Mitchell 2006). Species of this genus exhibit few morphological diagnostic characters and are therefore difficult to differentiate from each other. As a result, ecologists and paleoecologists often group members of this genus into morphologically similar "types" of individuals (Charman et al. 2000), even when there is evidence to suggest that these individuals represent independent evolutionary entities. However, as for macroscopic organisms, DNA barcoding studies have revealed the existence of several cryptic and pseudo-cryptic species

(Kosakyan et al. 2012). For example, some of the "types" historically included within *N. tincta* represent phylogenetic species that, in some cases, can be differentiated using morphological characters (Singer et al. 2015). Furthermore, some of these species may have different ecological optima (measured as water table depth) within *Sphagnum*-dominated peatlands (Valiranta et al. 2012). This then raises the question about the degree to which these organisms truly co-exist in syntopy, and if so, what are the likely mechanisms involved that make their co-existence possible? Answering these questions experimentally is impossible, because most testate amoebae (and slow-growing protists in general) are difficult or impossible to culture. Nevertheless, inferring the mechanisms driving community assembly in these highly specialized protists can be achieved indirectly, basing conclusions on correlations calculated from in situ data.

In this purpose, we investigated species composition of members of genus *Nebela* in the whole range of habitats (hummock, lawn, *Pinus* forest and fen) existing in peatlands of the Swiss Jura Mountains. To avoid identification mistakes that may occur with these similar looking species, we used a metabarcoding approach based on the specific sequencing of *Nebela* mitochondrial COI gene from environmental DNA. For each sampling point, we also measured relevant environmental variables (water table depth, nitrogen and phosphorus content, pH, tree cover and topography). We determined whether community composition was correlated to the environmental variables characterizing the different habitats, and if strong niche overlap occurred between species. A significant niche overlap would suggest interspecific competition rather than facilitation, given the close genetic relationship between species and trait similarity (Beltran et al. 2012). Finally, we determined taxon relatedness within the different habitats in order to detect if adaptations to a given environment were phylogenetically inherited or not. Phylogenetic inheritance would result in communities of closely related taxa (phylogenetic clustering) and that are organized by environmental filters, whereas lack of phylogenetic inheritance would result in communities of distantly related taxa (phylogenetic over-dispersion) and that are organized by intraspecific competition (Webb et al. 2002, Emerson and Gillespie 2008, Vamosi 2014).

Accordingly, we expect the distribution of *Nebela* species to reflect these ecological gradients. In addition, given the low nutrient levels (and especially nitrogen) in *Sphagnum* peatlands, we expect that nitrogen deficiency could represent a strong environmental filter that will influence the composition of communities. Secondly, if *Nebela* species are associated to topography as we suppose, niche overlap should be minimal, suggesting low competition. In addition, as it has been observed in other protist groups, we expect that ecological preferences should be phylogenetically inherited by a common ancestor in closely related *Nebela* lineages.

Methods

Sampling and environmental variables

The study was carried out in two similar Swiss peatlands from the Jura Mountains: Le Cachot (47°00'N 6°39'E) and Praz-Rodet (46°33'N 06°10'E). The peatlands were selected according to their similar characteristics of altitude (approximately 1000 m a.s.l.) and geomorphological setting (Lotze and Hölder 1964). Between September 2013 and April 2014, we collected 21 samples containing each 30g of fresh *Sphagnum* sp. from the top 5 cm in four habitats present in the studied peatlands. These habitats were defined on the base of criteria such as vegetation type, topography, and nutrient availability. Two of these habitats were in the center of the peatland, treeless and ombrotrophic, as they mostly received all of their water and nutrients from precipitation. The first of them was defined as hummock type (four samples); it exhibited a mound topography, low pH, low water table depth, and was dominated by the moss *Sphagnum fuscum* and the vascular plant *Vaccinium oxycoccos.* The second ombrotrophic habitat was defined as lawn type (five samples); it exhibited a flat topography, and was dominated by the moss *Sphagnum magellanicum* and vascular plant *Eriophorum vaginatum.* Then we selected two habitat types in the periphery of the peatland: 3) bog forest type (seven samples), dominated by *Pinus mugo* and *Vaccinium* spp.; 4) poor fen type (four samples), a transition zone between the peatland and the surrounding environment (mostly

meadows), characterized by the presence of *Betula pubescens* and *Sphagnum fallax* (Appendix S1: Table S1). Nutrient content and pH are typically lower in the center of the peat bog than in the surrounding environment (Batzer and Baldwin 2012). We measured water table depth (WTD), tree cover (classes: a=75-100%, b=50-75%, c=25-50%, d=0-25% of tree cover), and topography [classes: a=flat, b= slight mound (5-10cm), c= mound (>10cm)] as described in Singer (2016). We then measured total nitrogen (%), total phosphorus (%) and pH_{H2O} following standard protocols (Carter and Gregorich 2007) (Appendix S1: Table S2). The four habitats were differentiated according to their vegetation (Singer et al. 2016) and environmental characteristics (Appendix S1: Fig. S1).

Molecular analyses

Testate amoeba cells were extracted from 20g of fresh *Sphagnum* and concentrated by sieving at 150µm. We extracted DNA from the product of sieving at <150µm using a MoBio Power Soil® DNA Isolation kit according to the manufacturer's instruction. We amplified a fragment of the mitochondrial COI gene by using the general primer LCO (Folmer 1994) and a specific primer TINCOX (CCATTCKATAHCCHGGAAATTTC) following the protocol recommended for the amplification of *Nebela* species (Kosakyan et al. 2013). The expected size of the fragments ranges from 300 bp to 499 bp (Kosakyan et al. 2013). PCR steps consisted of a 5min initial denaturation step in a 40 cycles program of 15s at 95°C, 15s at 43°C, and 1min and 30s at 72°C with the final extension at 72°C for 10min. The amplicons were cloned into pCR2.1 Topo TA cloning vectors and transformed into *E. coli* TOP10' One Shots cells (Invitrogen kit) according to the manufacturer's instructions. We amplified (M13f, M13R primers (Kosakyan et al. 2015)) and sequenced up to 50 inserts per PCR product to reach the total diversity of *Nebela* species. Sequencing was carried out using a BigDye197 Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) and analyzed with an ABI-3130XL DNA sequencer (Applied Biosystems). In order to evaluate if the total diversity was reached, we performed rarefaction curves with visual check of asymptotic behavior independently for each site, for each habitat, and for the overall data set (Appendix S1: Figs. S2,S3,S4) and we computed Chao estimates for the overall data set using the R vegan package (v2.3-5 (Oksanen J et al. 2015).

Phylogenetic reconstructions

Phylogenetic analysis was used to determine the phylogenetic position of the sequences obtained. We built a reference database containing 32 previously known COI sequences (obtained from single cells) of the *Nebela collaris* species complex from GenBank (Heger et al. 2011, Kosakyan et al. 2012, Kosakyan et al. 2013, Singer et al. 2015) plus three sequences of the testate amoeba *Longinebela tubulosa* (GenBank accession number: JN849020, JN849021, JN849061) as outgroup. The dataset of the 35 reference sequences, together with the newly obtained sequences, was aligned using the Clustal W algorithm (Thompson et al. 1994) (Metadata S1). COI gene sequences of the genus *Nebela* potentially contain post-transcriptional mitochondrial editing resulting on the insertion of single nucleotides in the gene sequence (Kosakyan et al. 2013): these single nucleotides were further removed for all subsequent analyses. Phylogenetic trees were reconstructed using Bayesian and maximum likelihood (ML) approaches as in Kosakyan et al. (2013) and Singer et al. (2015). A Bayesian Markov Chain Monte Carlo (MCMC) analysis was performed using MrBayes v3.2 (Ronquist et al. 2012) with a general time reversible model of sequence evolution. A ML tree was built using RAxML v7.7.1 (Stamatakis et al. 2008) with a general time reversible model as well. Model parameters were directly estimated in RAxML.

Phylotype identification

To evaluate the number of independent evolving units within the genus *Nebela*, we used the Automatic Barcode Gap Discovery (ABGD) method (Puillandre et al. 2012). ABGD is an automated procedure that clusters sequences into candidate species based on pairwise distances by detecting differences between intra- and inter-specific variation without any *a priori* species hypothesis (Puillandre et al. 2012). We used the web-server of ABGD http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html with the Jukes-Cantor (JC69) model and the default parameters to run the analysis.

Niche partitioning and competition

To convert clone numbers to relative proportions of individuals from each species, we calculated the "corrected clone number" (C_{x_corr}) of each phylotype as described in (Kosakyan et al. 2015). Kosakyan et al. (2015) demonstrated that there is significant correlation between the bio-volume of a selected species and the number of clones obtained for the species within the community. C_{x_corr} is obtained by dividing the biovolume of the largest species (*Nebela collaris*) by the biovolume of a given species multiplied by the number of clones found in the community. We assigned the bio-volume of the largest testate amoeba species found in our study site (*Nebela collaris*) to phylotypes belonging to undescribed species (and thus with an unknown bio-volume) to avoid their overrepresentation within local communities (Appendix S1: Table S3).

We assessed whether species communities (here defined as the relative proportions of each species for each sample weighted by their biovolume) of the genus *Nebela* changed significantly among habitats using a redundancy analysis (RDA) with all measured quantitative variables (water table depth, pH, and nitrogen content), but did not use phosphorus because it was correlated to nitrogen content (Pearson correlation: r=0.66, p=0.001). The selected variables are known to describe the main ecological gradients in *Sphagnum*-dominated peatlands (Bridgham et al. 1996, Bridgham et al. 2001).

Direct experimental evidence cannot be provided to assess if competition takes place or not given that *Nebela* spp. do not grow in culture; thus, we used only indirect evidence to evaluate if competition may take place or not. Under this rationale, we computed (1) the C-score index (Stone and Roberts 1990) to evaluate whether niche breadth was significantly different between species and (2) the Pianka's index (Pianka 1973) to estimate the amount of niche overlap between species.

Species can either co-exist because species occupy different niches (reduced fundamental – or Grinnellian niche overlap; Grinnell 1917) or replace each other due to competition (high potential niche overlap but distinct realized – Eltoninan– niche; Elton 1927). (Grinnell 1917, Elton 1927). We

tested whether phylogenetically close species of the genus *Nebela* have the same niche breadth and, thus, exhibit strong interspecific competition (i.e. they exhibit a strong niche overlap and cannot co-occur) using the C-score index. This method requires a species x site matrix to calculate the species occurrence at a given site in the absence of another one. If a high number of species do not co-occur, then a checkerboard pattern arises in the matrix. The occurrence of this pattern is interpreted as an indication of competitive exclusion between species (Diamond and Cody 1975, Stone and Roberts 1990, Gotelli and Graves 1996, Novak et al. 2011). The statistical significance of this analysis was evaluated with a row-column null model (50,000 random matrices, $P < 0.05$), a conservative algorithm that minimizes Type I errors (Gotelli 2000). C-score values higher than expected support the null hypothesis, while values lower than expected reject the null hypothesis. C-score and null model calculations were computed using the R package EcoSimR v0.1.0 (Gotelli and Ellison 2013).

A supplementary line of evidence for showing potential competition was inferred calculating Pianka's niche overlap index. This index requires an incidence matrix with species as rows and niche categories as columns to analyze how much overlap exists in the use of a given resource axis among species. Pianka's index ranges from 0 (no resource used in common between two given species) to 1 (complete overlap in resource use). The statistical significance of this analysis was evaluated with a row-column null model (50,000 random matrices, $P < 0.05$), a conservative algorithm that minimizes Type I errors (Ulrich and Gotelli 2010). Values lower than expected (below the 95% confident interval of the null model expectations) support the occurrence of low niche overlap between taxa, while values higher than expected (above the 95% confident interval of the null model expectations) reject this hypothesis. Pianka's index and null model calculations were conducted using EcoSimR.

In order to test whether ecological preferences are lineage-inherited as predicted by the phylogenetic conservatism hypothesis (Losos 2008, Wiens et al. 2010), local communities were tested for non-random phylogenetic structure using the NTI/NRI metrics (Kembel et al. 2010). Both were calculated using the R package picante v1.6-2 (Kembel et al. 2010). Nearest taxon index (NTI) is

derived from the average branch length to the nearest co-occurring taxon, while net relatedness index (NRI) is derived from the sum of the branch lengths that connect all co-occurring taxa (Kembel et al. 2010). Positive values of NRI or NTI indicate phylogenetic clustering while negative values indicate phylogenetic evenness. The analysis was run with sequence abundance information and the parameters used for the phylogeny reconstruction were the same as described above. The phylogeny was pruned in order to have only one sequence for each taxon: we kept the sequence of the most abundant haplotype of each phylotype defined by the ABGD analysis. In order to understand the role of the environmental correlates on NTI (and also on NRI), we build linear mixed effect models LMEMs (Bunnefeld and Phillimore 2012, Beckerman 2014) with NTI (or NRI) values for each site as the response variable and a set of environmental measurements as explanatory variables. Given that samples were obtained from different habitats, habitat identity was included in the mixed effect models in the random error structure. Moreover, in order to assess the relative importance of each predictor in the statistical models, we used a model averaging approach (Burnham and Anderson 2003): the set of sub-models including all possible combinations of the explanatory variables was generated and the Relative Importance (RI) value of each variable was calculated, on a scale from 0 to 1, as the sum of the Akaike weights of the sub-models in which the variable appears; better models have larger Akaike weights, and a variable that contributes more to model fit will thus have a higher relative-importance value, closer to 1. Model averaging was performed with the R package MuMIn v1.10.5 (Barton 2014) (Appendix S1: Table S4).

Results

The rarefaction curves for the overall data set, for each habitat, and for the single sites reached asymptotic values (Appendix S1: Figs. S2, S3, S4). Also, the overall observed species richness (n=9) was within the confidence interval of the Chao estimates (9.9±2.1). These results confirmed that our sampling effort was reliable to describe the diversity of *Nebela* in the study sites.

The ABGD analysis including our 1028 sequences (Metadata S1) and the 32 reference sequences available in GenBank for the genus *Nebela* group identified a distinctive ''barcode gap'' centered around 7% of divergence of the COI sequences and delimited 14 genetic clusters as candidate species (Fig. 1). This result was consistent in all recursive partitions with prior intraspecific genetic divergence values between 0.28% and 0.77%, and we considered it more likely than the two other alternatives (defining 54 candidate species with intraspecific divergence values below 0.2% or considering only two species with an intraspecific divergence values higher than 1.29%).

According to the most conservative estimate in ABGD, 14 phylotypes could be identified. Nine of the phylotypes were present in our samples, whereas four other ones were represented only by sequences from GenBank (Fig. 1). Six of the phylotypes in our samples corresponded to known species (*N. collaris, N. guttata, N. tincta, N. rotunda, N. pechorensis* and *N. gimlii*) and the other three (referred to here as unknown phylotypes UP1-UP3) represented probably undescribed species (Fig. 2). UP3 corresponded to a sequence derived from a single cell documented also in previous studies (Kosakyan et al. 2012, Kosakyan et al. 2013) (Fig. 1). The four most abundant phylotypes could be affiliated to the species *N. collaris, N. guttata, N. tincta* and *N. gimlii*. Other phylotypes were less represented and never exceeded 20% of the total number of sequences (Fig. 2; Appendix S1: Table S1). Obtained sequences were deposited in GenBank under the following accession numbers: MG764316 to MG764387 (Appendix S1: Table S1).

The RDA analysis (Fig. 3) of *Nebela collaris* s.l. communities in the different habitats from the two peatlands showed that species communities did not differ between the two peatlands ($p = 0.093$) and that the effect of sampling date was not significant ($p = 0.08$), but differed significantly among habitats ($p = 0.018$) with only limited overlap between lawns and hummocks in the first two axes of the ordination space (Fig. 3). The first two axes explained 60% of the total variance, and two significant variables, N (total nitrogen content, $p=0.001$) and WTD (water table depth, $p=0.03$) explained respectively 20% and 7% of the variance (adjusted R^2).

The C-score index was significantly lower than expected (C-score: observed = 0.361; expected = 5.537 ± SD 0.078; p = 0.01), suggesting an absence of competitive exclusion between the different species of genus *Nebela* (i.e. they co-occur in the same site more often than expected). Furthermore, the Pianka's index was lower than expected by chance (Pianka's index: observed = 0.369, expected = 0.773 ± SD 0.002, p < 0.0001), suggesting that *Nebela* species exhibit a low niche overlap and that, therefore, they perform an effective partitioning along the resource axes measured (habitat breadth). Taken together, these two analyses suggested that competition may not play an important role in shaping the communities of *Nebela* species in the surveyed sites.

Regarding the phylogenetic structure of species communities expressed as NTI, nitrogen had a strong and significant effect on this metric (Table 1); a similar effect was found also for NRI (Table 1). NTI indices decreased linearly with increasing nitrogen content (p = 0.0003; Fig. 4). NRI indices present similar pattern as NTI (Appendix S1: Fig S5). Species composition of the most nitrogen-deprived habitats (i.e. lawns and hummocks) indicated a shift in species communities towards communities dominated by the closely related *N. tincta* and *N. guttata* (Figs. 1-3).

Discussion

Cryptic diversity and its potential implications for biomonitoring and paleoecology

Environmental molecular surveys as well as detailed DNA barcoding studies have revealed the existence of huge overall protist diversity and unsuspected cryptic diversity within numerous protist taxa, and the testate amoeba genus *Nebela* that we studied here is a typical example (Kosakyan et al. 2012, Kosakyan et al. 2013, Singer et al. 2015). We found several new phylotypes (UP), in addition to the recently described taxa, thus suggesting that the total diversity in the genus *Nebela* has still not been uncovered. This raises the question whether such diversity corresponds to different ecological optima and/or functions, which may potentially be relevant for bioindication, paleoecology or macroecology. Testate amoeba species are generally lumped into "morpho-types"

in ecological and especially in paleoecological studies (Charman et al. 2000, Mitchell et al. 2008). If cryptic species differ in their ecological optima and/or function, then potentially useful information for biomonitoring or paleoecological inference is lost by lumping. Such a case was demonstrated for aquatic insects used as biomonitors of environmental quality (Zhou et al. 2010). However, if the species cannot be identified even by detailed morphological observations then applications would only be limited to the living communities. If they do not differ in ecology or function, then lumping them in morpho-taxa may be a sound approach.

Most sequences from unknown *Nebela* species were recovered from *Sphagnum* of forested (tall pine trees) bog sites, which in natural conditions in the Jura Mountains represent a transitional environment between the open, wetter center of peatlands and the surrounding upland fir, beech and spruce forest on predominantly calcareous soil. Further investigations may therefore show that these potential new species have their ecological optima outside of peatlands, possibly in the acidic litter beneath spruce trees. As testate amoebae, including those belonging to genus *Nebela*, have been studied more in peatlands than in other habitats, the true diversity of this genus is likely to be higher and additional species may occur in other habitats and/or regions of the world.

Dispersal limitation

Dispersal limitation is unlikely to have played a role in shaping *Nebela* communities in our sampling sites. Indeed, the communities of both investigated peatlands were not significantly different in spite of the fact that sites were separated by about 100 km. Dispersal is also considered relatively easy in other comparable peatland soil organisms such as mites and rotifers (Fontaneto et al. 2008, Minor et al. 2016).

Environmental gradients, vegetation and driving factors of microbial diversity

Community patterns of soil micro-organisms including testate amoebae may be directly (e.g. through facilitation or allelopathy) or indirectly (e.g. through plant-induced modification of the environment (Philippot et al. 2013, Imparato et al. 2016)) influenced by plants, but in most cases the

possible existence of such mechanisms remains to be determined. The different *Nebela* species were strongly correlated to habitat-type also corresponding to contrasting vascular plant and bryophyte communities. The vegetation partly drives these gradients due to microclimatic effects (e.g. shading, rain interception), or litter quality. Especially, *Sphagnum* mosses strongly influence key factors such as pH, nutrient and cation concentrations and water content (Clymo 1973, Van Breemen 1995, Lindo and Gonzalez 2010). Testate amoebae are considered microbial top-predators. They therefore are likely to contribute to shaping microbial community structure. Inversely, the absence of certain prey organisms may influence their community composition (Gilbert et al. 2003).

Nitrogen as a primary driving factor for Nebela *species distribution among habitats*

Irrespective of any other biotic interaction, the distribution of *Nebela* spp. was highly correlated with nitrogen content (Figs. 3, 4, Appendix S1: Fig S5). Nutrient deficiency is common in *Sphagnum*-dominated peatlands, especially in hummocks (Bridgham et al. 1996), which can be considered as an extreme environment (Dedysh et al. 2006), and is known to shape prokaryotic communities. It is, thus, likely that the especially low availability of nitrogen in hummocks acts as an environmental filter influencing species composition within the *Nebela* species complex. Lower metabolic rates and slow growth could be the key to adaptation to these strongly oligotrophic habitats, as it has been observed in the green algal genus *Dicranochaete* (Caisova 2016).

Other abiotic factors commonly measured in peatlands, such as water table depth, pH, and phosphorus content, which also vary to a similar degree were not correlated with species composition within the *Nebela* species complex. For the broader community of testate amoebae, however water table depth and pH are usually the two main gradients explaining community composition leading to the development of transfer functions used in paleoecology (Mitchell et al. 2008). Altogether, these findings suggest that environmental filters do influence *Nebela* species distribution in different ways, supporting the hypothesis that species differ in their respective ecological preferences. Previous studies have already shown that environmental filters (and thus,

species preferences) may structure soil testate amoeba communities at biogeographical spatial scales (Fernandez et al. 2016, Heger et al. 2016, Lara et al. 2016). Our study extends these findings to ecological gradients at a finer spatial scale.

Phylogenetic niche conservatism

The competitive exclusion principle (Hardin 1960) predicts that two closely related species sharing the same niche should exclude each other. In our case, C-score values suggested a low likelihood that the presence or absence of one species in a given site was directly affected by the presence of another species. Likewise, Pianka's index suggested that, overall, *Nebela* species exhibit low niche overlap, and that species perform an effective partitioning along the resource axis measured (habitat breadth). However, these assumptions are solely based on co-occurrence of species on the same sites and thus cannot be considered as strong evidence against species competition. Observations based on species traits can nevertheless corroborate this hypothesis. Indeed, *Nebela* species can be differentiated, amongst other features, based on the width of the shell aperture and shell size/bio-volume (Kosakyan et al. 2013). For example, there is a positive correlation between the size of the opening of testate amoeba shells and the size of the prey on which testate amoeba species can predate (Jassey et al. 2014). In this study, *N. tincta* and *N. guttata* were often found co-occurring in the same sites (Fig. 2). Given that *N. tincta* has a larger aperture (19% wider) and bio-volume (44% larger) than *N. guttata*, it can be reasonably expected that these species do not compete against each other because they consume prey items of different size (Kosakyan et al. 2013, Kosakyan et al. 2015). Species coexistence by specialization on different prey items has been historically excellently illustrated by Darwin's finches (De León et al. 2014). An alternative hypothesis would be that both species feed on similar but very abundant prey items and hence do not compete for food. *Nebela* spp. have been shown to be broad generalists that feed on items that are abundant in peat bogs such as diatoms and fungi (Gilbert et al. 2000, Gilbert et al. 2003).

Furthermore, we also demonstrate that phylogenetic clustering occurs in the most nitrogen-depleted areas of the peatland. This is illustrated by strongly positive NTI values in habitats where nitrogen concentrations were lowest (hummocks and lawns), meaning that organisms encountered in these habitats are phylogenetically more closely related than expected by chance (Webb et al. 2002). Typical species from nitrogen-poor environments are the sister species *N. guttata* and *N. tincta* (Fig. 1). Therefore, it can be hypothesized that the most recent common ancestor of *N. tincta* and *N. guttata* underwent adaptations to exceptionally nitrogen-poor environments, notwithstanding their morphological differences in mouth aperture and biovolume. Phylogenetically constrained adaptations to nitrogen-poor environments are common in plants (Verboom et al. 2017) and in animals like deep water fishes (Ingram 2011, Parzanini et al. 2017). In *Nebela*, like in the latter example, it may result for instance in long generation times and low metabolic rates. In contrast, in habitats where nitrogen is less limiting (i.e. forest and fen), NTI values were close to 0, which means that genetic distances observed within each community were not significantly different than those expected by chance. We interpret this situation as a decrease in selective pressure in richer (or more mesophilic) environments (Fig. 4).

Wrap-up

Altogether, we show that the different species, as defined by COI (Fig. 2), have contrasting distributions among habitats within *Sphagnum* peatlands (Fig. 1-2), and that these patterns were reproducible (Fig. 3). Therefore, species distribution was not random across habitats and community patterns were indeed correlated to the main ecological gradients characterizing *Sphagnum* peatland (Fig. 3). These species also had different niche breadths and, where they co-occurred, were more likely to be affected by resource partitioning than competition. Finally, our results suggested that niche conservatism takes place in the genus *Nebela*, where only a few species can adapt to excessively low nitrogen contents. Habitat selection and resource partitioning shape *Nebela* communities, and might possibly drive the diversification in the genus and in other slow-growing

unicellular eukaryotes in general, showing that protists are subjected to the same ecological and evolutionary processes commonly reported in larger organisms.

Acknowledgements

This work was funded by Swiss National Science Foundation projects no. 310003A 143960 and 310003A 163254. L.D.F. was supported by projects FONDECYT (No 11170927) and UBO/VRIP (No 170201). We would also like to thank Thierry Heger for the help with the ABGD analysis and Quentin Blandenier for fruitful discussion on the manuscript. We also thank Sébastien Tschanz for giving us the official permission for sampling in the protected le Cachot peatland. We thank also the two anonymous reviewers for the complete and detailed comments on the manuscript.

References

Altermatt, F., E. A. Fronhofer, A. Garnier, A. Giometto, F. Hammes, J. Klecka, D. Legrand, E. Machler, T. M. Massie, F. Pennekamp, M. Plebani, M. Pontarp, N. Schtickzelle, V. Thuillier, and O. L. Petchey. 2015. Big answers from small worlds: a user's guide for protist microcosms as a model system in ecology and evolution. Methods in Ecology and Evolution **6**:218-231.

Andersen, R., M. Poulin, D. Borcard, R. Laiho, J. Laine, H. Vasander, and E. T. Tuittila. 2011. Environmental control and spatial structures in peatland vegetation. Journal of Vegetation Science **22**:878-890.

Barton, K. 2014. MuMIn: multi-model inference. R package ver. 1.10. 5.

Batzer, D. P. and A. H. Baldwin. 2012. Wetland habitats of North America: ecology and conservation concerns. Univ of California Press.

Beckerman, A. P. 2014. What can modern statistical tools do for limnology? Journal of Limnology **73**:161-170.

Beltran, E., A. Valiente-Banuet, and M. Verdu. 2012. Trait divergence and indirect interactions allow facilitation of congeneric species. Annals of Botany **110**:1369-1376.

Bridgham, S. D., J. Pastor, J. A. Janssens, C. Chapin, and T. J. Malterer. 1996. Multiple limiting gradients in peatlands: A call for a new paradigm. Wetlands **16**:45-65.

Bridgham, S. D., K. Updegraff, and J. Pastor. 2001. A comparison of nutrient availability indices along an ombrotrophic-minerotrophic gradient in Minnesota wetlands. Soil Science Society of America Journal **65**:259-269.

Bunnefeld, N. and A. B. Phillimore. 2012. Island, archipelago and taxon effects: mixed models as a means of dealing with the imperfect design of nature's experiments. Ecography **35**:15-22.

Burnham, K. P. and D. R. Anderson. 2003. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer New York.

Caisova, L. 2016. Dicranochaete - an enigmatic green alga with surprising adaptive capabilities. Phycologia **55**:219-229.

Carlson, M. L., L. A. Flagstad, F. Gillet, and E. A. D. Mitchell. 2010. Community development along a proglacial chronosequence: are above-ground and below-ground community structure controlled more by biotic than abiotic factors? Journal of Ecology **98**:1084-1095.

Carter, R. and E. G. Gregorich. 2007. Soil Sampling and Methods of Analysis, Second Edition. CRC Press.

Charman, D. J., D. Hendon, and W. A. Woodland. 2000. The identification of testate amoebae (Protozoa: Rhizopoda) in peats. Quaternary Research Association, London.

Clymo, R. S. 1973. The growth of *Sphagnum*: some effects of the environment. Journal of Ecology **61**:849-869.

Colwell, R. K. and J. A. Coddington. 1994. Estimating Terrestrial Biodiversity through Extrapolation. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **345**:101-118.

De León, L., J. Podos, T. Gardezi, A. Herrel, and A. Hendry. 2014. Darwin's finches and their diet niches: the sympatric coexistence of imperfect generalists. Journal of Evolutionary Biology **27**:1093-1104.

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

Dedysh, S. N., T. A. Pankratov, S. E. Belova, I. S. Kulichevskaya, and W. Liesack. 2006. Phylogenetic analysis and in situ identification of Bacteria community composition in an acidic Sphagnum peat bog. Applied and environmental microbiology **72**:2110-2117.

del Campo, J., L. Guillou, E. Hehenberger, R. Logares, P. Lopez-Garcia, and R. Massana. 2016. Ecological and evolutionary significance of novel protist lineages. European Journal of Protistology **55**:4-11.

Diamond, J. M. and M. L. Cody. 1975. Ecology and evolution of communities / Martin L. Cody and Jared M. Diamond, editors. Belknap Press of Harvard University Press, Cambridge, Mass.

Elton, C. S. 1927. Animal ecology, by Charles Elton; with an introduction by Julian S. Huxley. Macmillan Co., New York.

Emerson, B. C. and R. G. Gillespie. 2008. Phylogenetic analysis of community assembly and structure over space and time. Trends in Ecology & Evolution **23**:619-630.

Esteves, S. D. and A. Vicentini. 2013. Cryptic species in Pagamea coriacea sensu lato (Rubiaceae): evidence from morphology, ecology and reproductive behavior in a sympatric context. Acta Amazonica **43**:415-428.

Fernandez, L. D., B. Fournier, R. Rivera, E. Lara, E. A. D. Mitchell, and C. E. Hernandez. 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**:1216-1227.

Foissner, W. 1999. Soil protozoa as bioindicators: pros and cons, methods, diversity, representative examples. Agriculture Ecosystems & Environment **74**:95-112.

Folmer, O. 1994. DNA Primers for Amplification of Mitochondrial Cytochrome C Oxidase Subunit I from Diverse Metasoan Invertebrates. Rutgers University.

Fontaneto, D., T. G. Barraclough, K. Chen, C. Ricci, and E. A. Herniou. 2008. Molecular evidence for broad-scale distributions in bdelloid rotifers: everything is not everywhere but most things are very widespread. Molecular Ecology **17**:3136-3146.

Fukami, T. and P. J. Morin. 2003. Productivity-biodiversity relationships depend on the history of community assembly. Nature **424**:423-426.

Futuyma, D. J. and G. Moreno. 1988. The Evolution of Ecological Specialization. Annual Review of Ecology and Systematics **19**:207-233.

Geisen, S., A. T. Tveit, I. M. Clark, A. Richter, M. M. Svenning, M. Bonkowski, and T. Urich. 2015. Metatranscriptomic census of active protists in soils. Isme Journal **9**:2178-2190.

Gilbert, D., C. Amblard, G. Bourdier, F. André-Jean, and E. A. Mitchell. 2000. Le régime alimentaire des thécamoebiens (Protista, Sarcodina). L'année Biologique **39**:57-68.

Gilbert, D. and E. A. D. Mitchell. 2006. Microbial diversity in *Sphagnum* peatlands. Pages 287-317 *in* I. P. Martini, A. Matinez Cortizas, and W. Chesworth, editors. Peatlands: basin evolution and depository of records on global environmental and climatic changes. Elsevier, Amsterdam.

Gilbert, D., E. A. D. Mitchell, C. Amblard, G. Bourdier, and A. J. Francez. 2003. Population dynamics and food preferences of the testate amoeba *Nebela tincta major-bohemica-collaris* complex (Protozoa) in a *Sphagnum* Peatland. Acta Protozoologica **42**:99-104.

Gotelli, N. J. 2000. Null model analysis of species co-occurrence patterns. Ecology **81**:2606-2621.

Gotelli, N. J. and A. M. E. Ellison. 2013. EcoSimR

Gotelli, N. J. and G. R. Graves. 1996. Null models in ecology. Smithsonian Institution Press.

Grinnell, J. 1917. The niche-relationships of the California Thrasher. The Auk **34**:427-433.

Grossmann, L., M. Jensen, D. Heider, S. Jost, E. Glücksman, H. Hartikainen, S. S. Mahamdallie, M. Gardner, D. Hoffmann, and D. Bass. 2016. Protistan community analysis: key findings of a large-scale molecular sampling. The ISME journal.

Hanson, C. A., J. A. Fuhrman, M. C. Horner-Devine, and J. B. Martiny. 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. Nature Reviews Microbiology **10**:497-506.

Hardin, G. 1960. The competitive exclusion principle. Science **131**:1292-1297.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes

fulgerator. Proceedings of the National Academy of Sciences of the United States of America **101**:14812-14817.

Heger, T. J., N. Derungs, J. P. Theurillat, and E. A. D. Mitchell. 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.

Heger, T. J., E. Lara, and E. A. D. Mitchell. 2011. Arcellinida testate amoebae (Arcellinida: Amoebozoa): model of organisms for assessing microbial biogeography.*in* D. Fontaneto, editor. The importance of being small: does size matter in biogeography? Cambridge University Press.

Hutchinson, G. E. 1961. The Paradox of the Plankton. American Naturalist **95**:137-145.

Imparato, V., S. S. Santos, A. Johansen, S. Geisen, and A. Winding. 2016. Stimulation of bacteria and protists in rhizosphere of glyphosate-treated barley. Applied Soil Ecology **98**:47-55.

Ingram, T. 2011. Speciation along a depth gradient in a marine adaptive radiation. Proceedings of the Royal Society B-Biological Sciences **278**:613-618.

Janzen, D. H., J. M. Burns, Q. Cong, W. Hallwachs, T. Dapkey, R. Manjunath, M. Hajibabaei, P. D. Hebert, and N. V. Grishin. 2017. Nuclear genomes distinguish cryptic species suggested by their DNA barcodes and ecology. Proceedings of the National Academy of Sciences **114**:8313-8318.

Jassey, V. E. J., L. Lamentowicz, B. J. M. Robroek, M. Gazbka, A. Rusinska, and M. Lamentowicz. 2014. Plant functional diversity drives niche-size-structure of dominant microbial consumers along a poor to extremely rich fen gradient. Journal of Ecology **102**:1150-1162.

Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, and C. O. Webb. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics **26**:1463-1464.

Kooistra, W. H. C. F., P. Hargraves, R. A. Andersen, S. Balzano, A. Zingone, and D. Sarno. 2005. Pseudo-Cryptic Diversity and Phylogeography in the Centric Diatom Skeletonema. Phycologia **44**:56-57.

Kosakyan, A., F. Gomaa, E. A. D. Mitchell, T. J. Heger, and E. Lara. 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.

Kosakyan, A., T. J. Heger, B. S. Leander, M. Todorov, E. A. D. Mitchell, and E. Lara. 2012. COI Barcoding of Nebelid Testate Amoebae (Amoebozoa: Arcellinida): Extensive Cryptic Diversity and Redefinition of the Hyalospheniidae Schultze. Protist **163**:415-434.

Kosakyan, A., D. J. G. Lahr, M. Mulot, R. Meisterfeld, E. A. D. Mitchell, and E. Lara. 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**:606-623.

Kosakyan, A., M. Mulot, E. A. D. Mitchell, and E. Lara. 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**:311-320.

Lara, E., L. Roussel-Delif, B. Fournier, D. M. Wilkinson, and E. A. D. Mitchell. 2016. Soil microorganisms behave like macroscopic organisms: patterns in the global distribution of soil euglyphid testate amoebae. Journal of Biogeography **43**:520-532.

Lentendu, G., T. Wubet, A. Chatzinotas, C. Wilhelm, F. Buscot, and M. Schlegel. 2014. Effects of long-term differential fertilization on eukaryotic microbial communities in an arable soil: a multiple barcoding approach. Molecular Ecology **23**:3341-3355.

Lindo, Z. and A. Gonzalez. 2010. The Bryosphere: An Integral and Influential Component of the Earth's Biosphere. Ecosystems **13**:612-627.

López-García, P., F. Rodriguez-Valera, C. Pedrós-Alió, and D. Moreira. 2001. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature **409**:603-607.

Losos, J. B. 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecology Letters **11**:995-1003.

Lotze, F. and H. Hölder. 1964. Handbuch der stratigraphischen Geologie: Bd. IV Jura. Ferdinand Enke.

Mahé, F., C. de Vargas, D. Bass, L. Czech, A. Stamatakis, E. Lara, D. Singer, J. Mayor, J. Bunge, and S. Sernaker. 2017. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. Nature Ecology & Evolution **1**:0091.

Minor, M., S. Ermilov, D. Philippov, and A. Prokin. 2016. Relative importance of local habitat complexity and regional factors for assemblages of oribatid mites (Acari: Oribatida) in Sphagnum peat bogs. Experimental and Applied Acarology **70**:275-286.

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

Novak, M., J. W. Moore, and R. A. Leidy. 2011. Nestedness patterns and the dual nature of community reassembly in California streams: a multivariate permutation-based approach. Global Change Biology **17**:3714-3723.

Ofiteru, I. D., M. Lunn, T. P. Curtis, G. F. Wells, C. S. Criddle, C. A. Francis, and W. T. Sloan. 2010. Combined niche and neutral effects in a microbial wastewater treatment community. Proceedings of the National Academy of Sciences of the United States of America **107**:15345-15350.

Okland, R. H., K. Rydgren, and T. Okland. 2008. Species richness in boreal swamp forests of SE Norway: The role of surface microtopography. Journal of Vegetation Science **19**:67-74.

Oksanen J, Blanchet F.G, Kindt R, Legendre P, Minchin P.R, O'Hara R.B, Simpson G.L, Solymos P, Stevens M.H.H, and Wagner H. 2015. vegan: Community Ecology Package.

Parzanini, C., C. C. Parrish, J. F. Hamel, and A. Mercier. 2017. Trophic ecology of a deep-sea fish assemblage in the Northwest Atlantic. Marine Biology **164**.

Philippot, L., J. M. Raaijmakers, P. Lemanceau, and W. H. van der Putten. 2013. Going back to the roots: the microbial ecology of the rhizosphere. Nature Reviews Microbiology **11**:789-799.

Pianka, E. R. 1973. The Structure of Lizard Communities. Annual Review of Ecology and Systematics **4**:53-74.

Puillandre, N., A. Lambert, S. Brouillet, and G. Achaz. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular Ecology **21**:1864-1877.

Rissler, L. J. and J. J. Apodaca. 2007. Adding more ecology into species delimitation: Ecological niche models and phylogeography help define cryptic species in the black salamander (Aneides flavipunctatus). Systematic Biology **56**:924-942.

Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Hohna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Systematic Biology **61**:539-542.

Santos, H., J. Juste, C. Ibanez, J. M. Palmeirim, R. Godinho, F. Amorim, P. Alves, H. Costa, O. De Paz, G. Perez-Suarez, S. Martinez-Alos, G. Jones, and H. Rebelo. 2014. Influences of ecology and biogeography on shaping the distributions of cryptic species: three bat tales in Iberia. Biological Journal of the Linnean Society **112**:150-162.

Scheffer, M., S. Rinaldi, J. Huisman, and F. J. Weissing. 2003. Why plankton communities have no equilibrium: solutions to the paradox. Hydrobiologia **491**:9-18.

Schönborn, W. 1975. Estimation of Annual Production of Soil Protozoa .1. Euglyphidae (Rhizopoda, Testacea). Pedobiologia **15**:415-424.

Schönborn, W. 1992. Comparative Studies on the Production Biology of Protozoan Communities in Freshwater and Soil Ecosystems. Archiv für Protistenkunde **141**:187-214.

Singer, D., A. Kosakyan, A. Pillonel, E. A. D. Mitchell, and E. Lara. 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**:79-85.

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

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. Systematic Biology **57**:758-771.

Steinberg, C. E. W., S. Kamara, V. Y. Prokhotskaya, L. Manusadzianas, T. A. Karasyova, M. A. Timofeyev, Z. Jie, A. Paul, T. Meinelt, V. F. Farjalla, A. Y. O. Matsuo, B. K. Burnison, and R. Menzel. 2006. Dissolved humic substances - ecological driving forces from the individual to the ecosystem level? Freshwater Biology **51**:1189-1210.

Stone, L. and A. Roberts. 1990. The Checkerboard Score and Species Distributions. Oecologia **85**:74-79.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. Nucleic Acids Research **22**:4673-4680.

Ulrich, W. and N. J. Gotelli. 2010. Null model analysis of species associations using abundance data. Ecology **91**:3384-3397.

Valiranta, M., A. Blundell, D. J. Charman, E. Karofeld, A. Korhola, U. Sillasoo, and E. S. Tuittila. 2012. Reconstructing peatland water tables using transfer functions for plant macrofossils and testate amoebae: A methodological comparison. Quaternary International **268**:34-43.

Vamosi, S. M. 2014. Phylogenetic community ecology as an approach for studying old ideas on competition in the plankton: opportunities and challenges. Journal of Limnology **73**:186-192.

Van Breemen, N. 1995. How *Sphagnum* bogs down other plants. Trends in Ecology & Evolution **10**:270-275.

Van Hannen, E. J., G. Zwart, M. P. van Agterveld, H. J. Gons, J. Ebert, and H. J. Laanbroek. 1999. Changes in bacterial and eukaryotic community structure after mass lysis of filamentous cyanobacteria associated with viruses. Applied and environmental microbiology **65**:795-801.

Verboom, G. A., W. D. Stock, and M. D. Cramer. 2017. Specialization to Extremely Low-Nutrient Soils Limits the Nutritional Adaptability of Plant Lineages. American Naturalist **189**:684-699.

Wanner, M. and W. E. R. Xylander. 2005. Biodiversity development of terrestrial testate amoebae: is there any succession at all? Biology and Fertility of Soils **41**:428-438.

Webb, C. O., D. D. Ackerly, M. A. McPeek, and M. J. Donoghue. 2002. Phylogenies and community ecology. Annual Review of Ecology and Systematics **33**:475-505.

Weiner, A., R. Aurahs, A. Kurasawa, H. Kitazato, and M. Kucera. 2012. Vertical niche partitioning between cryptic sibling species of a cosmopolitan marine planktonic protist. Molecular Ecology **21**:4063-4073.

Weisse, T., U. Scheffel, P. Stadler, and W. Foissner. 2007. Local adaptation among geographically distant clones of the cosmopolitan freshwater ciliate Meseres corlissi. II. Response to pH. Aquatic Microbial Ecology **47**:289-297.

Wiens, J. J., D. D. Ackerly, A. P. Allen, B. L. Anacker, L. B. Buckley, H. V. Cornell, E. I. Damschen, T. J. Davies, J. A. Grytnes, S. P. Harrison, B. A. Hawkins, R. D. Holt, C. M. McCain, and P. R. Stephens. 2010. Niche conservatism as an emerging principle in ecology and conservation biology. Ecology Letters **13**:1310-1324.

Wilkinson, D. M. 1998. Fragments of an entangled bank: do ecologists study most of ecology? Oikos **82**:393-394.

Worden, A. Z., M. J. Follows, S. J. Giovannoni, S. Wilken, A. E. Zimmerman, and P. J. Keeling. 2015. Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. Science **347**.

Zhou, J. Z., Y. Deng, P. Zhang, K. Xue, Y. T. Liang, J. D. Van Nostrand, Y. F. Yang, Z. L. He, L. Y. Wu, D. A. Stahl, T. C. Hazen, J. M. Tiedje, and A. P. Arkin. 2014. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. Proceedings of the National Academy of Sciences of the United States of America **111**:E836-E845.

Zhou, X., L. M. Jacobus, R. E. DeWalt, S. J. Adamowicz, and P. D. N. Hebert. 2010. Ephemeroptera, Plecoptera, and Trichoptera fauna of Churchill (Manitoba, Canada): insights into biodiversity patterns from DNA barcoding. Journal of the North American Benthological Society **29**:814-837.

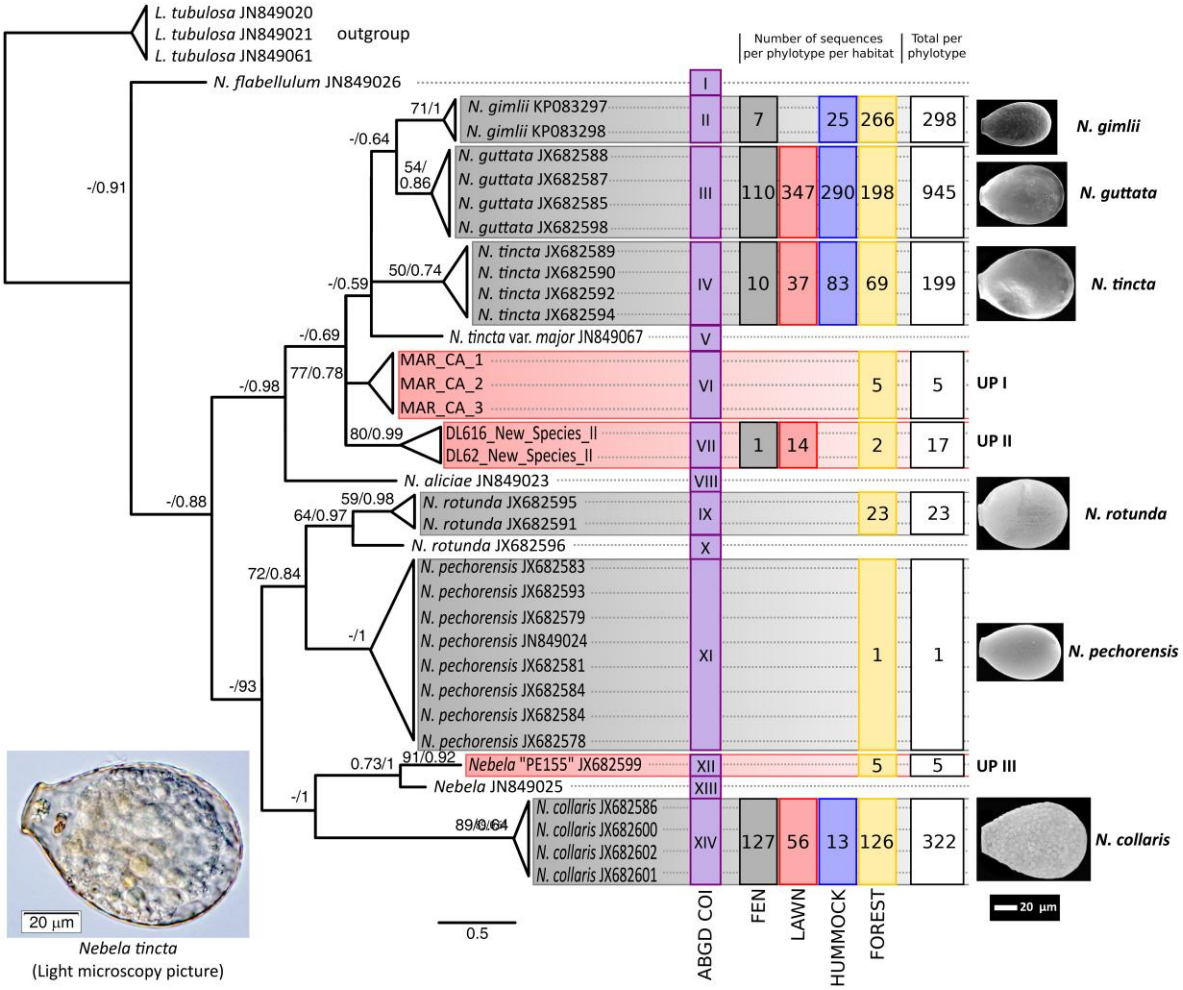
Figure legends

Fig 1) Bayesian Markov Chain Monte Carlo consensus tree of 32 single cells of *Nebela collaris* s.I. mitochondrial cytochrome oxidase, subunit I (COI) sequences from two peatlands in the Swiss Jura Mountains and 1028 environmental sequences, plus three sequences of *Longinebela tubulosa* as outgroup. The numbers along the branches represent, respectively, bootstraps obtained by maximum likelihood method and posterior probabilities as calculated with Bayesian analyses. Only values above 50/0.50 are shown. Sample colors indicate the five different habitats within peatlands (red = lawn, yellow = forest, blue = hummock, black = fen).

Fig 2) Community profiles of testate amoebae from the *Nebela collaris* complex s.l. extracted from *Sphagnum* samples from two peatlands in the Swiss Jura Mountains. Each barplot represents the relative ratio of the clone sequences corrected by the cell biovolume (Kosakyan, 2015). The phylotypes that are not related to any known species are represented by "UP" (Unknown Phylotype).

Fig 3) Redundancy analysis diagram of *Nebela collaris* s.l. communities extracted from *Sphagnum* sample collected from two peatlands in the Swiss Jura Mountains, Le Cachot bog (empty symbols) and Praz-Rodet bog (filled symbols). Significant environmental variables ($p<0.05$) are represented by arrows.

Fig 4) Nearest Taxon Index (NTI) values from *Nebela collaris* s.l. communities extracted from *Sphagnum* sample collected from two peatlands in the Swiss Jura Mountains in function of *Sphagnum* Nitrogen content (%). The line shows the linear regression ($p=0.004$) and the dotted lines correspond to the 95% confidence interval.



Nebela tincta
(Light microscopy picture)