UNIVERSITÉ DE NEUCHÂTEL

Faculty of Science
Institute of Biology
Laboratory of Soil Biodiversity

Assessing soil micro-eukaryotic diversity using high-throughput amplicons sequencing: spatial patterns from local to global scales and response to ecosystem perturbation.

Christophe Victor William Seppey

For the degree of:

Doctor in Science in Biology

Thesis committee:

Prof. Edward A.D. Mitchell (Supervisor)	University of Neuchâtel (Switzerland)
Dr. Enrique Lara (Co-supervisor)	University of Neuchâtel (Switzerland)
Dr. David Bass (Jury)	Natural History Museum of London (United Kingdom) Centre for environment Fisheries & Aquaculture Science (United Kingdom)
Dr. Micah Dunthorn (Jury)	University of Kaiserslautern (Germany)
Dr. Stefan Geisen (Jury)	Netherlands Institute of Ecology (Netherlands) Wageningen University (Netherlands)

July 4, 2017

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 Christophe SEPPEY

Titre:

"Assessing soil micro-eukaryotic diversity using high-throughput amplicons sequencing: spatial patterns from local to global scales and response to ecosystem perturbation"

sur le rapport des membres du jury composé comme suit:

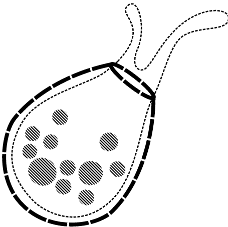
- Prof. Edward Mitchell, directeur de thèse, Université de Neuchâtel, Suisse
- Dr Enrique Lara, co-directeur de thèse, Université de Neuchâtel, Suisse
- Dr Micah Dunthorn, University of Kaiserslautern, Allemagne
- Dr David Bass, Natural History Museum, Londres, UK
- Dr Stefan Geisen, Netherlands Institute of Ecology, Wageningen, NL

Neuchâtel, le 3 juillet 2017

Le Doyen, Prof. R. Bshary

**Choose a job you love,
and you will never have to work a day in your life.**

Confucius



Résumé

Les micro-eucaryotes exhibent une immense diversité qui remplie plusieurs fonctions essentielles dans les écosystèmes colonisés. Ces micro-organismes sont impliqués dans tout les niveaux trophique microbiens, interagissent entre eux ainsi qu'avec d'autre groupes d'organismes tel les procaryotes ou les macro-organismes, et influencent les cycles d'éléments comme ceux du carbone ou de l'azote. Les diversité et écologie des micro-eucaryotes sont étudiées à partir de la morphologie de ces organismes et de plus en plus avec des méthodes moléculaire devenant plus abordable que jamais. Le séquençage haut débit de fragments d'ADN donnant une information taxonomique prise directement de l'environnement est maintenant le standard pour établir les communautés microbiennes et pratiquement saturer la diversité microbienne. Cette thèse profite des avancées dans cette technique pour étudier l'écologie des micro-eucaryotes des sols, organismes qui représentent la base de la plupart des écosystèmes terrestre et sont impliqué dans de critique questions écologique comme les changement climatique ou l'approvisionnement alimentaire. Les cinq chapitres suivent des communautés contraintes par différent niveaux de stress ou perturbation et distribuées autant sur de petite surfaces que sur le globe. Des analyses écologique classique et innovante sont utilisées dans ce travail pour couvrir des questions à propos de bioindication, fonctions, niveaux trophique, distribution spatiale et diversité de ce groupe de micro-organismes peu connu à l'immense diversité.

Mots clefs: Micro-eucaryotes . Sols . ADN environnementale . Séquençage haut débit Illumina . Metabarcoding . Petite sous unité du gène de l'ARN ribosomique

Summary

Micro-eukaryotes exhibit a huge diversity which fulfils many essential functions in the colonized ecosystems. These micro-organisms are involved in every level of microbial trophic networks. They interact with each other and with other biota like prokaryotes or macro-organisms, and influence element cycles like the carbon or nitrogen cycle. The diversity and ecology of micro-eukaryotes are studied based on morphological analyses and more and more with molecular methods which are increasingly affordable. High-throughput sequencing of taxonomically informative DNA fragments taken directly from the environment is now the golden standard to assess microbial communities and virtually saturate the microbial diversity. This thesis takes advantage of the advances in this technique to study the ecology of micro-eukaryotes in soils, which represent the basis of most terrestrial ecosystems and are involved in critical ecological issues like climate changes or food supply. The five chapters follow communities constrained by different levels of stress or perturbation and distributed from very limited areas to global ecosystems. Classical and innovative ecological analyses are used in this work to cover questions about the bioindication, functions, trophic networks, spacial distributions and diversity of these hyper-diverse and largely unknown micro-organisms.

Keywords: Micro-eukaryotes . Soils . Environmental DNA . Illumina high-throughput sequencing . Metabarcoding . SSU rRNA gene

Contents

Résumé . . . v
Summary . . . v

Contents vii

Introduction 1
1 Micro-eukaryotic communities . . . 1
2 Soils . . . 4
3 Thesis aims . . . 7

1 Euglyphids response in cadaver impacted soil 9
1.1 Introduction . . . 9
1.2 Materials and methods . . . 10
1.3 Results . . . 12
1.4 Discussion . . . 13

2 Co-occurrence network of micro-eukaryotes in peatlands 19
2.1 Introduction . . . 19
2.2 Materials and methods . . . 20
2.3 Results and discussion . . . 22

3 Spatial modelling of soil protist diversity in the Swiss western Alps 33
3.1 Introduction . . . 34
3.2 Materials and methods . . . 34
3.3 Results . . . 35
3.4 Discussion . . . 37

4 Algivory: an alternative pathway for the soil microbial loop 59
4.1 Introduction . . . 59
4.2 Materials and methods . . . 60
4.3 Results . . . 62
4.4 Discussion . . . 65

5 Comparison between soils and planktonic micro-eukaryotic communities 81
5.1 Introduction . . . 82
5.2 Materials and methods . . . 83
5.3 Results and discussion . . . 84

Discussion and conclusion 97
1 Background . . . 97
2 Achievements . . . 98
3 Perspectives . . . 99

Bibliography 101

Curriculum Vitae 111

Remerciements / Acknowledgements 115

Introduction

Christophe V.W. Seppey[1*]

[1]Laboratory of Soil Biodiversity, University of Neuchâtel, Rue Émile-Argand 11, 2000 Neuchâtel, Switzerland
[*]**Corresponding author:** christophe.seppey@unine.ch

Abstract: Micro-eukaryotes exhibit a huge diversity which fulfils many essential functions in the colonized ecosystems. These micro-organisms are involved in every level of microbial trophic networks. They interact with each other and with other biota like prokaryotes or macro-organisms, and influence element cycles like the carbon or nitrogen cycle. The diversity and ecology of micro-eukaryotes are studied based on morphological analyses and more and more with molecular methods which are increasingly affordable. High-throughput sequencing of taxonomically informative DNA fragments taken directly from the environment is now the golden standard to assess microbial communities and virtually saturate the microbial diversity. This thesis takes advantage of the advances in this technique to study the ecology of micro-eukaryotes in soils, which represent the basis of most terrestrial ecosystems and are involved in critical ecological issues like climate changes or food supply. The five chapters follow communities constrained by different levels of stress or perturbation and distributed from very limited areas to global ecosystems. Classical and innovative ecological analyses are used in this work to cover questions about the bioindication, functions, trophic networks, spacial distributions and diversity of these hyper-diverse and largely unknown micro-organisms.

Keywords: Micro-eukaryotes . Soils . Environmental DNA . Illumina high-throughput sequencing . Metabarcoding . SSU rRNA gene

1 Micro-eukaryotic communities

What is a micro-eukaryote?

Micro-eukaryotes, include any eukaryotic organism small enough to be invisible to the naked eye and are found in all super-groups of eukaryotes (Pawlowski et al., 2012). These micro-organisms encompass mostly protists, a polyphyletic group containing all eukaryotes except Metazoa (animals), Embryophyceae (higher plants) and Fungi (Whittaker, 1969) (Figure 1). This group includes the phagotrophic protozoa (e.g. Ciliata), phototrophic eukaryotic algae (e.g. Chlorophyceae), and osmotrophic eukaryotes living as free organisms (e.g. Labyrinthulea) or involved in mutualistic (e.g. Trebouxiophyceae) or parasitic symbiosis (e.g. Apicomplexa). Contrarily to its macrobial counterpart, micro-eukaryotic diversity is mostly unknown (Pawlowski et al., 2012). This is due to their small size that 1) makes it difficult to find the organism in a sample - and even harder if the sample contains a high proportion of particles as in soils - and 2) makes it more difficult to find distinctive morphological features as this often requires using scanning electron or transmission electron microscopy (Ekelund and Ronn, 1994; Mulot et al., 2017).

Fungi form the second main group of micro-eukaryotes. Indeed, although many Basidiomycota and Ascomycota produce macroscopic carpophores, many species are only found in a microscopic form like yeasts (e.g. *Saccaromyces*) (Masinova et al., 2017; Hittinger, 2013). In addition, most of their biomass - at least in soils - is formed from undifferentiated cells (hyphae) reaching considerable cumulative length but only a few micrometers in diameter (Gobat et al., 2010d). In many soils, they play an important role in the eukaryotic microbiome as they constitute the main part of the micro-eukaryotic biomass (Ekelund et al., 2001). This osmotrophic taxonomic group is well known to form mutual and parasitic symbiosis with higher plants (Cairney and Chambers, 1999) as well as with other organisms like Metazoa (e.g. Entomophthoromycota (Boomsma et al., 2014)) or even Fungi (e.g. Cryptomycota (Lara and Belbahri, 2011)).

In addition to the first two groups, some Metazoa like rotifers, tardigrades or nematodes are also included in micro-eukaryotes. However, these micro-Metazoa will not be taken into account in this thesis as many of them can reach a size up to a millimetre and are constituted of differentiated cells.

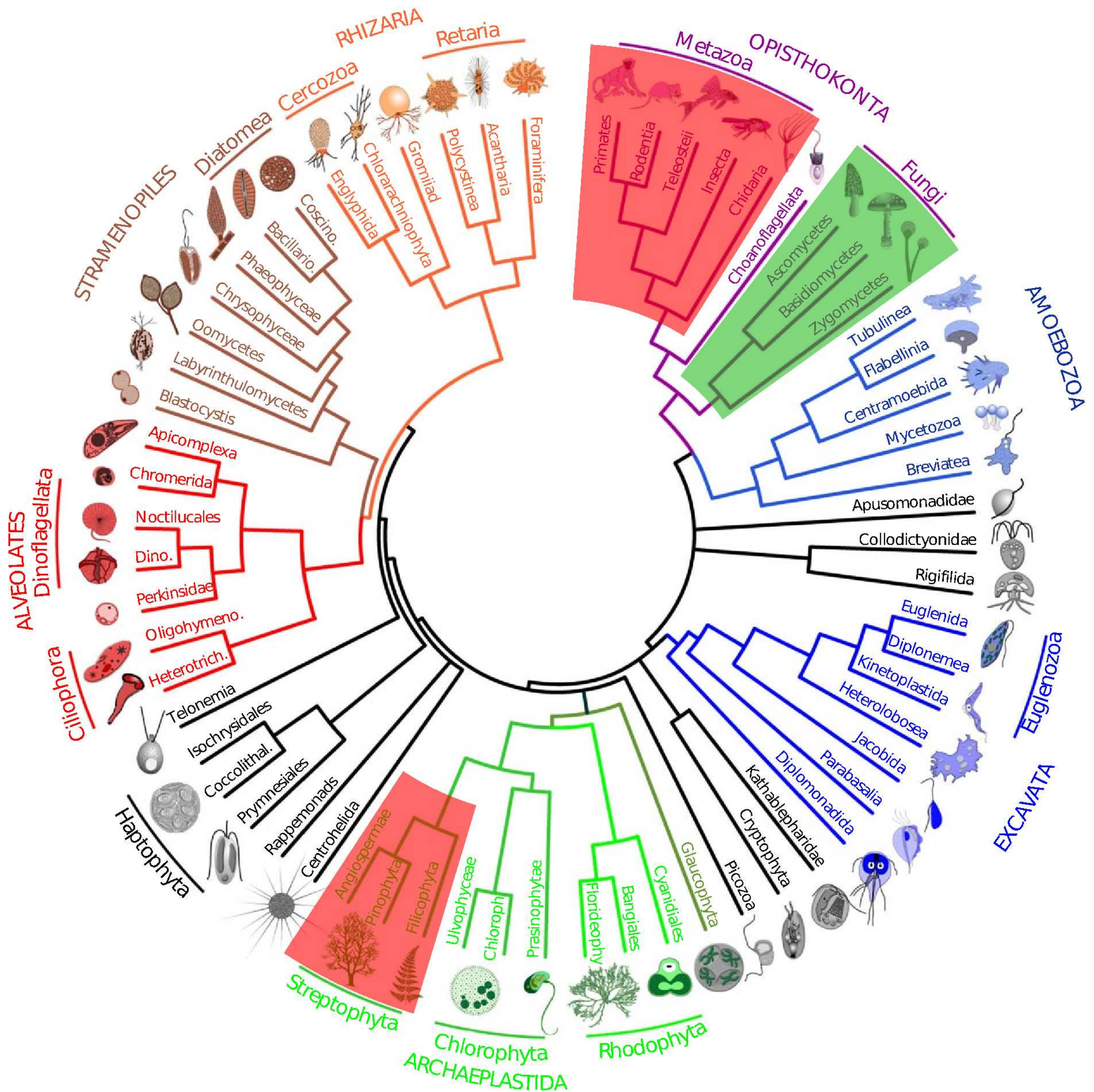


Figure 1: Schematic tree of eukaryotes modified from (Pawlowski, 2014). Micro-eukaryotes constitute almost the entire tree apart from the Metazoa (animals) and macro-Streptophyta (higher plants) (in red). Apart from Fungi (in green) all micro-eukaryotes are protists.

What functions do they fulfil?

Phagotrophs

In every biome on Earth, micro-eukaryotes play fundamental roles that can roughly be divided in five categories. The first, and maybe most acknowledged one, is the role of microbial predators which was conceptualized in the microbial loop (Azam et al., 1983; Clarholm, 1985). According to this theory, phagotrophic protists act as "natural fertilizer producers" as they make nutrients available for other organisms (e.g. plants and algae). These nutrients would otherwise stay sequestrated in bacterial biomass (Bonkowski, 2004). In addition to prokaryotes, other carbon sources are used by phagotrophs like micro-eukaryotic algae (Davidson, 1996), Fungi (Ekelund, 1998; Geisen, 2016; Geisen et al., 2016) or Metazoa (Geisen et al., 2015b; Gilbert et al., 2000). Phagotrophic protists, as a functional group, can thus be considered as predators of virtually the entire microbial world. This predation pressure shapes the prey communities promoting certain characteristics like toxicity, motility or size (Jurgens and Matz, 2002). Nevertheless, certain predators adapt to these defences (Swanson, 2013; Xinyao et al., 2006; Wu et al., 2004) leading to a microbial arms race (Brodie Jr et al., 1991) and constituting a powerful motor of evolution.

Phototrophs

The second function micro-eukaryotes can fulfil in the environment is as primary producers. In aquatic ecosystems, they are an important source of carbon for higher

trophic levels and are the dominant contributors to overall photosynthesis (Vaulot et al., 2008; Jardillier et al., 2010; Falkowski, 2002). In terrestrial ecosystems, where the surface is mainly covered by higher plants, the eukaryotic micro-algae contribution is more discrete but can constitute a significant carbon input in arid regions (Pushkareva et al., 2016; Belnap, 2003). Through their symbiosis with lichens, eukaryotic algae also play - to a certain extent - the role of higher plants by stabilising the soil and improving the water infiltration (Chen et al., 2000; Upreti et al., 2015; Eldridge and Greene, 1994). They can also constitute a significant part of the carbon ingested by certain soil organisms like springtails, earthworms (Schmidt et al., 2016) or phagotrophic protists (Hess and Melkonian, 2013; Seppey et al., 2017).

Mixotrophs

Comparatively to macro-organisms, many micro-eukaryotes are not constrained in strict phagotrophy or phototrophy, and can benefit from the two trophic strategies. The ideal mixotroph (constitutive mixotroph) is an organism capable of transmitting the plastid vertically to the next generation and able to regulate the organelle via nuclear genes (Mitra et al., 2016). In addition to constitutive mixotrophs, more and more phototrophic eukaryotes are observed ingesting preys under light or nutrient limitation (Stoecker, 1998). Phagotrophs can also conserve the chloroplasts from their preys and contribute to primary production (Johnson, 2011a,b). This strategy may be the norm rather than the exception, at least in aquatic ecosystems (Flynn et al., 2013). In addition, this strategy increases the growth rate of the population in comparison to a strictly phototrophic or phagotrophic mode (Burkholder et al., 2008). The possibility to switch from a primary production mode, where carbon is easily available but nutrients and light are mandatory, to a heterotrophic mode, based on preys like prokaryotes but not requiring anything else, allows this functional group to adapt to the availability of nutrients, carbon and light. This adaptability can firstly be seen at the community level when primarily autotrophic mixotrophic species (e.g. cryptophytes) dominate in eutrophic conditions while primarily heterotrophic mixotrophic species (e.g. chrysophytes) dominate in oligotrophic environments (Saad et al., 2016; Lara et al., 2015). The adaptation can also be seen at the species level as it was shown for *Dinobryon sociale* where the bacteria ingestion rate increases at a certain temperature or as nutrient concentrations decrease (Princiotta et al., 2016).

Parasites

Many groups of micro-eukaryotes are known as parasites of various hosts like plants (e.g. *Phytophthora* sp.), animals (e.g. Apicomplexa) or even micro-eukaryotes (e.g. *Rozella* sp.). As all parasites, these micro-eukaryotes can be more or less host specific and may influence the host population in a density-dependent way, that could increase the diversity of the host population (i.e. Janzen-Connel model) (Mahé et al., 2017; Freckleton and Lewis, 2006; Connell, 1972; Janzen, 1970). In addition, like micro-eukaryotic predators, parasites apply a selection pressure on their host and vice versa, which can cause a Red Queen coevolution (Raberg et al., 2014; Rabajante et al., 2016). The ecological models in which parasitic micro-eukaryotes are involved can also be more complex when the symbiont can switch from a parasitic strategy to a mutualistic one depending on the health of the host (Sachs et al., 2011; Fellous and Salvaudon, 2009).

Saprotrophs

Saprotrophic micro-eukaryotes are mainly found within Fungi and some groups of protists (e.g. slime moulds, certain Oomycota, Labyrinthulomycota) (Beakes et al., 2012; Tsui et al., 2009; Mendoza et al., 2002). Their role in terrestrial ecosystems is mainly to recycle dead organic matter into new biomass (Adl and Gupta, 2006). As a result, these organisms constitute the main part of the living biomass in soil (Gobat et al., 2010d).

Where can we find them?

Micro-eukaryotes are everywhere; from marine to freshwater ecosystems and from liquid to solid substrates. Through evolution, a variety of adaptations were selected according to the various possible environments, which led to the diversity of morphology and physiology we can observe today (Katz, 2012). These adaptations allow micro-eukaryotes to colonize from more suitable to harsher and perturbed environments. The best example of adaptation to changing environments is the capacity of many protists to encyst - sometime for years - waiting for more suitable conditions (Foissner, 1987). This is particularly true in soils where the conditions of humidity and nutrients can vary through periods of dryness or proximity of a root exudate for example (Adl and Gupta, 2006). It is even hypothesized that the soil micro-eukaryotic diversity would be mainly constituted by encysted organisms (Foissner, 1987; Clarholm, 2005). This capacity to encyst and wait for optimal conditions partly explains the incomplete picture of the eukaryotic morphological diversity because of our incapacity to establish the conditions needed to reactivate these organisms in the laboratory (Ekelund and Ronn, 1994; Foissner, 1999b).

The various types of symbiosis also help micro-eukaryotes to colonize almost all Earth environments. Lichen, often composed of Fungi and a eukaryotic alga (e.g. Trebouxiophyceae, Chlorophyceae), benefit of their respective adaptation to gather nutrients from bedrock and carbon from the air (Chen et al., 2000; Upreti et al., 2015).

These adaptations more or less fit into the various environmental conditions which influences the spatial distributions of micro-eukaryotes. This phenomenon, added to historical spatial distributions and environmental barriers, helps to explain the present geographic patterns (Foissner,

2006; Bass et al., 2007; Smith et al., 2008; Lara et al., 2016).

How to study them?

Morphological methods

Since the beginning of protistology and mycology, micro-eukaryotes have been cultivated and observed through the microscope (optical and electronic) to characterize their diversity. The combination of these two tools is essential for the tens to hundreds of micro-eukaryotic new species descriptions taxonomists achieve every year (Appeltans et al., 2012). In addition to morphological identifications, these classical tools give precious information about the ecology of the cultivated organisms, like feeding behaviour (Berge et al., 2008; Dumack et al., 2016b; Jassey et al., 2012; Xinyao et al., 2006; Singh, 1946; Verni and Gualtieri, 1997) or tolerance to environmental characteristics like pH (Koch and Ekelund, 2005; Weisse et al., 2013; Germond et al., 2013, 2014), temperature (Pichrtova and Nemcova, 2011), salinity (Scholz and Liebezeit, 2012b; Nielsen et al., 2011) or nutrients (Mitchell, 2004). This information is essential for most interpretations of ecological results despite being unavailable for the large majority of micro-eukaryotes (Ekelund and Ronn, 1994).

Micro-eukaryotic communities can be assessed by microscopy preceded or not by a cultivation step (Singh, 1955; Foissner, 1983). These techniques have the advantage to be inexpensive and simple to perform. The most probable number technique (MPN) allows estimating the number of protists in a sample and can give pure cultures as by-products, which can be used for species description or experiment purposes (Singh, 1955; Darbyshire et al., 1974). This technique mainly makes it possible to assess opportunistic organisms that are not necessarily abundant in the environment but benefit from the cultivation medium which is normally rich in nutrients (Del Campo et al., 2013). Nevertheless, the cultivation medium can be adapted, to a certain extent, to the environmental conditions of the sample of origin in order to better represent the original community (Ekelund and Ronn, 1994). Direct counting can also be used to avoid cultivation bias and thus have a better idea of the actual community (Foissner, 1983; Luftenegger et al., 1988). It was shown that communities of large and well identifiable organisms like testate amoebae (Payne and Mitchell, 2009) or ciliates (Posch et al., 2015) can be reliably assessed through this technique. However, these two techniques suffer of serious drawbacks (Ekelund and Ronn, 1994; Clarholm, 2005; Adl and Gupta, 2006). MPN is constrained by cultivation bias, allowing only a part of the micro-eukaryotic community to grow. On the other hand, community assessment by direct observations may fail to reveal small or inconspicuous organisms, particularly if the samples contain many particles as in soils. In addition, these methods are recognised as time consuming, which reduces the number of samples that can be handled for a single study (Hamsher et al., 2013).

Molecular methods

Methods based on DNA sequences allow overcoming a part of the issues linked to morphological identifications like organism size or cultivability (Moreira and Lopez-Garcia, 2002; Pawlowski et al., 2012). Taxonomically informative DNA fragments (DNA barcodes) like the ribosomal RNA gene were primarily used to improve the classifications of the cultivated and described species through phylogenetic reconstruction (Schlegel, 1994). For this reason, more and more effort is now given to increase and cure databases of these DNA barcodes (Pawlowski et al., 2012; Guillou et al., 2013). As for morphological data, methods emerged to obtain the cellular DNA sequences directly from the environment to overcome the culturing bias and reach the uncultivable micro-eukaryotic diversity, which constitutes the majority of the diversity for many taxonomic groups (de Vargas et al., 2015; Pawlowski et al., 2012). Once the DNA is extracted from the environment, amplicons of the DNA barcodes (metabarcoding) can be grouped according to their similarity to form operational taxonomic units (OTUs) that constitute biological entities like their morphological counterparts observed by microscopy (Taberlet et al., 2012; Valentini et al., 2009). Each OTU can then be taxonomically assigned by comparing its characteristic sequence to a DNA barcode database. The obtained community matrix can then give a good proxy to assess the responses of both dead or alive and active or inactive micro-organisms to environmental constrains (Giner et al., 2016).

Within the last few years, the democratisation of high-throughput sequencing allows assessing micro-eukaryotic OTU communities faster and from much more samples than before (Shendure and Ji, 2008). A constraint of high-throughput sequencing is the use of short fragments, which provide less phylogenetic information but are less prone to break in the environment and/or form chimera (Valentini et al., 2009). Nevertheless, these sequencing technologies develop rapidly resulting in an increase in sequences length and numbers (Guzvic, 2013). Despite giving semi-quantitative results, the number of sequences belonging to each OTU can be used to calculate ecological metrics and give results in line with studies made on morphological data. Metabarcoding is now the golden standard to assess high numbers of micro-eukaryotic communities in a reasonable amount of time (Figure 2).

2 Soils

What is a soil?

Soils are the intersection between the geological, aquatic, aerial and biological worlds. They result from the weathering of the bedrock through physical and chemical processes and the action of pioneer living organisms like lichen or mosses

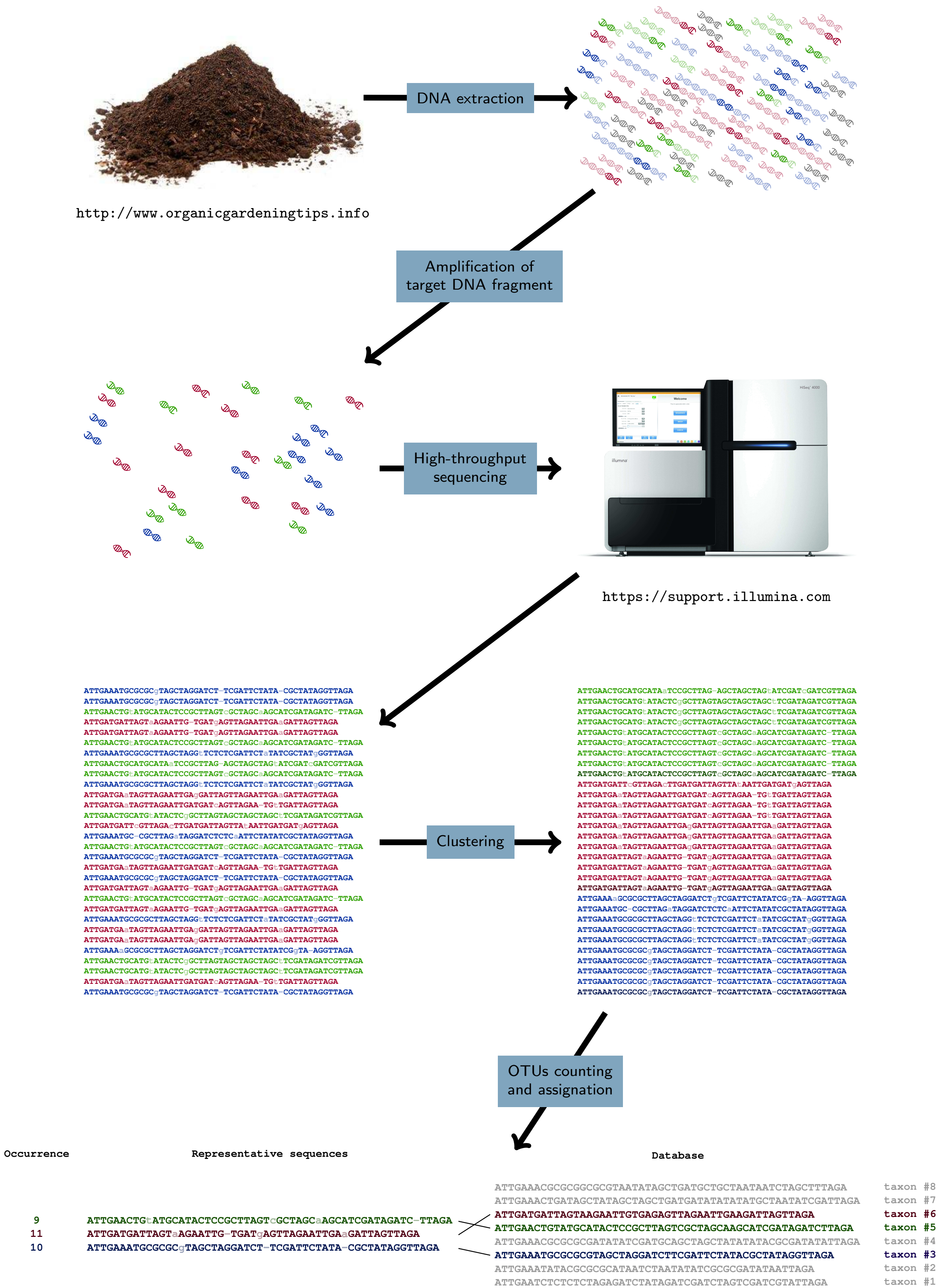


Figure 2: Schematic plan of the metabarcoding analysis of one sample

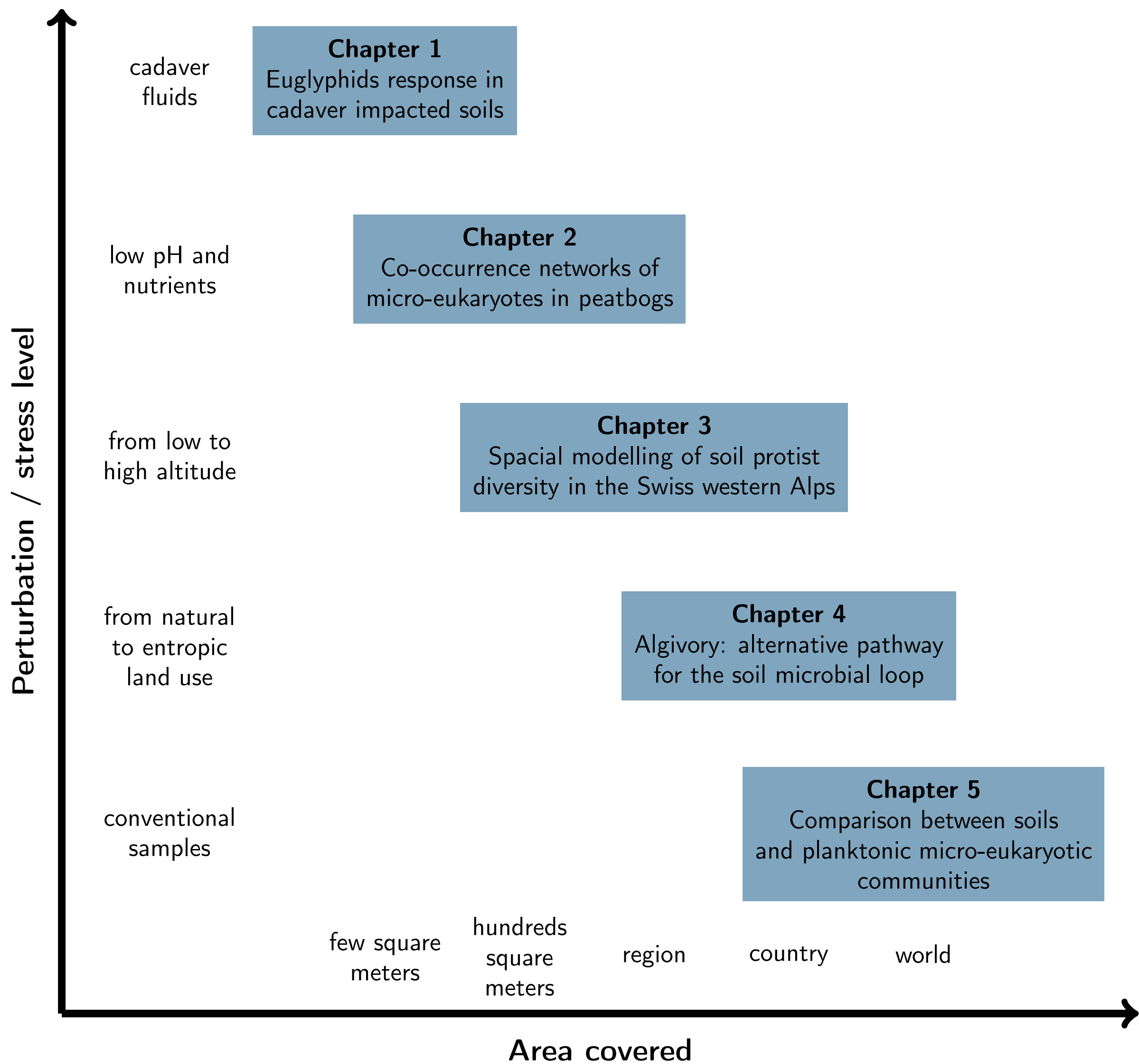


Figure 3: Schematic organisation of the chapters according to the size of area studied and the level of perturbation/stress.

(Coleman et al., 2004a; Bardgett, 2005b; Chen et al., 2000). These organisms contribute to the genesis of soils by taking nutrients from the parental material and the carbon (CO_2) and nitrogen (N_2) from the air (Bardgett, 2005c; Gobat et al., 2010a,b; Coleman et al., 2004b). The dead organic matter produced by the pioneer organisms accumulates and mixes with the fragmented mineral substrate, allowing sequential colonisation waves leading to a deeper soil and a more complex ecosystem (Gobat et al., 2010c).

Soils can firstly be described vertically through their horizons going from the more mineral, closer to the parental material, to the more organic at the surface (Bardgett, 2005a). This vertical gradient obviously influences, and is influenced, by the soil organisms who are more abundant in the upper horizons because of their higher content of nutrients and labile carbon. A soil can also be described by its texture (percentage of clay, silt and sand), as well as by the structure of its aggregates. These characteristics will influence the porosity of the soil and its capacity to retain water and nutrients as well as its resistance to erosion or compaction (Boiffin et al., 1988). As an example, a clay-rich soil retains more water and nutrients with a positive impact on soils organisms (Gobat et al., 2010e).

What are its functions?

Above ground support

Soils are essential components of terrestrial ecosystems. They support plant growth that provides most of the organic carbon that feeds terrestrial life, including humans. This support to plants is firstly influenced by the capacity of the soil to provide nutrients, water and physical support but also catalysed by the below ground diversity (Anderson and Weigel, 2003). Indeed, many soil organisms contribute to plant well-being like mycorhizal fungi which increase the nutrients and water intake surface of a plant (van der Heijden et al., 1998). Plants also benefit from soil micro-organisms from their production of bioavailable nitrogen as symbiotic nitrogen-fixing bacteria

(Pajares and Bohannan, 2016) or by releasing ammonium as metabolic waste from the digestion of their microbial preys (Clarholm, 1985; Bonkowski, 2004). Vertebrates and invertebrates also physically aerate the soil and provide privileged interstices for roots to pass (Gobat et al., 2010f).

The quality of this support to above ground ecosystems is directly affected by what can impact soils physically, chemically and biologically. An example are agricultural soils which are physically compacted, and chemically amended and treated with pesticides. The compaction can create a hardpan which reduces water infiltration to deeper horizons, thus reducing the crop water reserve for dry periods. In addition, water runoff is increased as the liquid is constrained to stay on the surface, causing erosion (Boiffin et al., 1988). Chemically, fertilizers and pesticides have often detrimental effects on the abundance and diversity of soil organisms thus hampering their beneficial effects on soils and plants (Mitchell, 2004; Todorov and Golemansky, 1992; Foissner, 1997).

Heterogeneous habitat

Soils constitute the habitats of a plethora of macro- and micro-organisms. One of the specificity which explains this high diversity is heterogeneity. Indeed, soil characteristics change in the vertical and horizontal spaces according to e.g. topography, parental material, vegetation or human activity (IUSS et al., 1999). At a smaller scale, soils also show micro-heterogeneity through their aggregates, with interstices more or less humid or rich in nutrients. At the microbial scale, this alternation of more or less suitable habitats can allow two species competing for the same niche to live almost in the same place (Adl and Gupta, 2006). Soil conditions are also temporally heterogeneous, for instance when pushing many micro-organisms to encyst under temporally harsh conditions (e.g. dryness) (Adl and Gupta, 2006). This encystment of certain species allows others, that are less competitive under optimal conditions but can support the harsh conditions, to use the available niche (Adl and Gupta, 2006).

An excellent example of how soil spatial and temporal heterogeneity can drive below-ground communities can be illustrated with cadaver decomposition (Barton et al., 2013). Cadaveric fluids drastically change soil chemical properties by altering pH (Benninger et al., 2008), carbon or nitrogen concentration (Hopkins et al., 2000; Anderson et al., 2013) and other nutrients such as phosphorus (Towne, 2000) or potassium (Stokes et al., 2013). In addition, the soil can become temporarily anoxic during the most active parts of decomposition due to cadaveric fluids entering the soil (Aitkenhead-Peterson et al., 2015) and mainly because of the intense bacterial respiration (Dent et al., 2004). Over the short term, these changes favour anoxic organisms and r-strategists, and kill most of the established diversity (Szelecz et al., 2014; Dent et al., 2004). Nevertheless, these high nutrient concentrations are favourable for specific communities that then flourish for some time after the nutrients concentration drops below an acceptable level (Towne, 2000; Seppey et al., 2016). These specific conditions are still not fully explored and thus probably hide a high proportion of unknown diversity (Seppey et al., 2016).

Elements cycles (Carbon and Nitrogen)

Soils also play a central role in the cycles of many elements, including carbon, the most well known. Soils firstly stock a significant amount of the terrestrial organic carbon (2160 Pg; Gobat et al., 2010a; Davidson et al., 2000), principally in waterlogged soils where the oxidation of organic matter is slowed down by anaerobic conditions (Belyea and Malmer, 2004). This phenomenon is particularly true in the case of peatlands which cover ca. 3% of the terrestrial surface (Kivinen and Pakarinen, 1981) and contain ca. 30% of its organic carbon (Gorham, 1991; Rydin and Jeglum, 2013; Clymo and Hayward, 2012). Depending on how these ecosystems respond to on-going climate change they could either buffer or enhance global warming. Indeed, climate change tends to lower the water level, thus exposing the peat to oxidation. This oxygen increase causes a shift from anaerobic to aerobic microbial communities which start to release CO_2, triggering a warming positive feedback (Dieleman et al., 2015; Dorrepaal et al., 2009).

Soils are also a essential compartment in the nitrogen cycle, particularly in terrestrial ecosystems. Indeed, it is in soils that the un-assimilable atmospheric N_2 is fixed by nitrifying bacteria as NH_4^+ that can than be taken by plants and higher terrestrial trophic levels (Gobat et al., 2010b). In many ecosystems, nitrogen is the limiting factor for primary producer growth and it is thus the reason why industrial fixation produces up to 100 Tg of assimilable nitrogen per year for crop uses. Most of this input is however leached in aquatic networks and contributes to the growth of algae and the eutrophication of the ecosystem after the consumption and respiration of the algal biomass (Bunting et al., 2016; Howarth and Marino, 2006).

3 Thesis aims

The aim of this thesis is to improve knowledge about soil micro-eukaryotic diversity and community ecology using the metabarcoding approach. The diversity of micro-eukaryotes is studied at different spatial scales and in relation to various disturbance factors (Figure 3).

The first chapter assesses the very localised effect of a cadaver perturbation on the community of euglyphid testate amoebae. The effect of the perturbation is studied over more than two and a half years, allowing to measure the recovery of the euglyphid communities over time. In addition, the molecular data revealed that, although most OTUs responded negatively, some could benefit from the conditions brought by the cadavers.

The second chapter explores the interactions among different functional groups of micro-eukaryotes in four peatland micro-habitats. As these four micro-habitats differ in nutrient content and humidity, the design allows

linking the interactions between functions to the environmental gradients. The interactions between OTUs are assessed by co-occurrence networks, an innovative analysis in the field of protistology. The results show that the less diverse micro-habitat, which is nutrient-poor and comparatively drier, is the one where the communities are the most tightly linked, suggesting a higher resistance to environmental changes. OTUs assigned to low trophic levels, like small predators or mixotrophs, show the highest degree of linkage. This suggests that the network stability is based on low trophic levels functional groups, implying a bottom-up driven trophic network in stressed habitats like peatlands.

The third chapter shows the preliminary results [1] of a study aiming to predict the micro-eukaryotic diversity of meadow soils through the mountainous regional area of the Swiss western Alps. Spatial distribution models are used to measure the predictive power of both local physico-chemical (e.g. pH, texture, soil temperature) and regional topo-climatic (e.g. slope aspect, air temperature) environmental predictors on the diversity of total protists community and of nine broad phylogenetic groups. In general, the predictive power of regional predictors was similar to that of physico-chemical variables. These results open the door to the extrapolation of protist diversity through a large territory as the topo-climatic predictors are easier to extrapolate to an area than local physico-chemical ones.

The fourth chapter explores the functional and taxonomic diversity of micro-eukaryotes across Switzerland in three common land-uses (forests, meadows, crop lands). The relations between the more dominant phagotroph OTUs and the amount of sequences assigned to phototrophic micro-eukaryotes is assessed. Positive linear relations are found for seven of the phagotrophic assigned OTUs, cumulating between the third and the fourth of the phagotroph assigned sequences. Taxonomic evidence was found in the literature to support the hypothesis according to which these OTUs would be related to algivorous organisms. In addition, observations of three protists sharing a similar taxonomy to the one used to assigned the putative algivores OTUs were done. This suggests that a significant amount of carbon is entering the microbial loop through micro-eukaryotic algae.

The fifth and last chapter aims to compare the micro-eukaryotic diversity of world soils and plankton (marine and freshwater). The results show that soils host at least as much diversity as oceans and suggest that the diversity of certain acknowledged aquatic taxa is higher in soils than previously thought.

[1] The discussion of this chapter should be extended with respect to the predictive power of topo-climatic and edaphic variables on protist diversity. Nevertheless, the results about 1) the communities retrieved in that mountain area, 2) the difference in explicative power of two sets of variables (edaphic and topo-climatic) and, 3) the spatial prediction of soil protist diversity were considered innovative enough to be included in the manuscript.

Int J Legal Med; 103, 2 pp.551-562 (2016)
doi:10.1007/s00414-015-1149-7

Chapter 1

Response of forest soil euglyphid testate amoebae (Rhizaria: Cercozoa) to pig cadavers assessed by high-throughput sequencing

Christophe V.W. Seppey[1*], Bertrand Fournier[1,2], Ildikò Szelecz[1,3], David Singer[1], Edward A.D. Mitchell[1,4], Enrique Lara[1]

[1]Laboratory of Soil Biodiversity, University of Neuchâtel, Rue Émile-Argand 11, 2000 Neuchâtel, Switzerland
[2]Evolutionary Community Ecology Group, CNRS, University of Montpellier 2, 34095 Montpellier Cedex 05, France
[3]Institute of Forensic Medicine, Goethe University, Kennedyallee 104, 60596 Frankfurt, Germany
[4]Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland
[*]**Corresponding author:** christophe.seppey@unine.ch

Abstract: Decomposing cadavers modify the soil environment, but the effect on soil organisms and especially on soil protists is still poorly documented. We conducted a 35-month experiment in a deciduous forest where soil samples were taken under pig cadavers, control plots and fake pigs (bags of similar volume as the pigs). We extracted total soil DNA, amplified the SSU ribosomal RNA (rRNA) gene V9 region and sequenced it by Illumina technology and analysed the data for euglyphid testate amoebae (Rhizaria: Euglyphida), a common group of protozoa known to respond to micro-environmental changes. We found 51 euglyphid operational taxonomic units (OTUs), 45 of which did not match any known sequence. Most OTUs decreased in abundance underneath cadavers between days 0 and 309, but some responded positively after a time lag. We sequenced the full-length SSU rRNA gene of two common OTUs that responded positively to cadavers; a phylogenetic analysis showed that they did not belong to any known euglyphid family. This study confirmed the existence of an unknown diversity of euglyphids and that they react to cadavers. Results suggest that metabarcoding of soil euglyphids could be used as a forensic tool to estimate the post-mortem interval (PMI) particularly for long-term (>2 months) PMI, for which no reliable tool exists.

Keywords: Environmental DNA . Euglyphid testate amoebae . Illumina high-throughput sequencing . Metabarcoding . SSU rRNA gene V9 region . Forensic ecology

1.1 Introduction

The estimation of time since death and more generally the ability to detect the presence of cadavers even when the remains are no longer present are the two major objectives in forensic research (Prangnell and McGowan, 2009; Rodriguez and Bass, 1985).

Calculation of the post-mortem interval (PMI), an essential element of legal medicine and criminal investigation to estab- lish the timing of events that led to the death of a person, becomes less precise with the advance of the decomposition process. Until now, two main approaches are used to estimate the PMI. The medical method provides information ranging from a few hours to several days after death (Henssge et al., 2000; Amendt et al., 2004; Wyss and Cherix, 2006). The second method, forensic entomology, is based on the observation of larval stages of necrophagous flies and beetles and can be used to establish a PMI up to some weeks (Amendt et al., 2004; Wyss and Cherix, 2006). Although well established, the accuracy of entomological methods has been questioned after the full development of the first generation of necrophagous insects (Wyss and Cherix, 2006). Therefore, the development of additional indicators for PMI estimates beyond 1 month would constitute a welcome addition to the toolkit of forensic criminal investigators.

According to Payne (1965), the decomposition of ca-

davers can be separated into six stages: fresh, bloated, active decay, advanced decay, dry and remain stages. During the 'bloated' and 'active decay' stages (Payne, 1965; Carter et al., 2007), the release of cadaver liquids into the soil changes the chemical parameters drastically (Vass et al., 1992) This perturbation of the soil environment has been referred to as 'ephemeral resource patches' (Barton et al., 2013) leading to the development of 'cadaver decomposition islands' (CDI) (Carter et al., 2007). Although most of the decomposition takes place in the first few weeks under optimal conditions, cadaver effects on the soil environment can be long lasting. For example, Towne (Towne, 2000) showed that nitrogen and phosphorus concentration and pH were still significantly enhanced in soil samples taken under cadavers 2 years after laying ungulate cadavers on a prairie, while Melis et al. (Melis et al., 2007) reported enhanced soil calcium content and pH as late as 7 years post mortem in a CDI. Such environmental changes were shown to have an effect on the soil fauna (Bornemisssza, 1957), bacteria (Horswell et al., 2002; Howard et al., 2010; Moreno et al., 2011) and fungi (Hawksworth and Wiltshire, 2011; Carter and Tibbett, 2003). However, knowledge about cadaver effects on soil communities remains very limited, and almost nothing is known about the response of soil protists (Szelecz et al., 2014).

In this study, we focused on euglyphid testate amoebae (Rhizaria: Cercozoa), a highly diverse and abundant group of protists that reacts rapidly to environmental changes by shifts in community structure and abundance (Foissner, 1999b). Euglyphids include about one quarter of the ca. 300 testate amoeba morphospecies known to occur in soils (Foissner, 1999b). These amoeboid unicellular protists range mostly between 20 and 150 µm in length, and their densities typically range between ca. 10^6 and 10^8 individuals per square metre (Foissner, 1987). They build a shell (test) reinforced with ornamented self-secreted siliceous plates, and these shells allow species identification even after the death of the organism (Meisterfeld, 2000a,b). Most euglyphids are heterotrophs and feed mainly on bacteria and fungi (Gilbert et al., 2000). The distribution patterns of soil testate amoebae along environmental gradients and their response to environmental changes have been well studied, including soil humidity (Bobrov et al., 1999; Booth, 2008; Swindles and Ruffell, 2009); temperature (Beyens et al., 2009; Tsyganov et al., 2011); pH (Booth, 2001; Charman, 2001; Mitchell et al., 1999, 2008); and pesticide (Petz and Foissner, 1989), nitrogen, phosphorus and sulphate concentration (Gilbert et al., 1998a,b; Mitchell, 2004; Payne et al., 2010). They can thus be expected to also respond to the presence of decomposing cadavers. The generation time of euglyphids, which ranges from ca. 2 days to 1 week under natural conditions (Heal, 1964), is considerably longer than that of bacteria or smaller protozoa such as nanoflagellates, and this represents an advantage regarding their use as bioindicators. It is indeed short enough to allow them to (re)colonise rapidly suitable habitats (Wanner et al., 2008; Wanner and Elmer, 2009) and respond to environmental change over a period of weeks. However, as euglyphids are highly sensitive to environmental conditions, the effects on communities can be expected to be long lasting under continuous environmental stress. So especially for estimating longer PMIs, euglyphids might be a group to consider for forensic applications.

However, a current limitation to the development of euglyphid analysis (or that of other soil protists) as a standard tool for PMI estimates is taxonomy. Sound taxonomy is indeed a prerequisite for the use of a group of organisms as bioindicators. Up to now, all ecological studies on testate amoebae were based on morphology-based species identifications. The morphological identification of testate amoebae requires taxonomic expertise and is time-consuming. Furthermore, recent molecular taxonomy studies on euglyphids have revealed the existence of a substantial higher diversity than estimated based only on morphology (Chatelain et al., 2013; Heger et al., 2011), and this hidden/unknown diversity may prove to have bioindication value. The molecular approach presented in this study overcomes the current limitation of morphology- based taxonomy and is also faster (i.e. weeks instead of months for the number of samples analysed here).

Analysis of environmental samples (e.g. soil, water, faeces) targeting a specific DNA barcode gene and aiming at characterising the entire community is referred to as metabarcoding (Taberlet et al., 2012). The V9 region of the 18S rRNA gene has sufficient variability for obtaining reasonably high taxonomic resolution (Amaral-Zettler et al., 2009; Adl et al., 2013) and two main advantages for biodiversity surveys as follows: (1) it is short and thus likely to be well preserved in environmental DNA samples and (2) it contains highly conserved sites allowing to designing primers for virtually all eukaryotes (Valentini et al., 2009). The advent of high-throughput sequencing (HTS) now allows using the metabarcoding approach in ecological studies with high sample numbers (e.g. spatial and/or temporal sampling).

In this study, we used a DNA metabarcoding approach applied to the V9 region of the 18S rRNA gene to assess the temporal response of soil euglyphid testate amoebae to decomposing cadavers over a period of 35 months (1051 days). Given the sensitivity of the technique and the well-documented response of testate amoebae to ecological gradients, changes and disturbances, we expected to find (1) a higher diversity of soil Euglyphids than generally inferred from microscopic analyses and (2) a strong generally negative response of communities to decomposing cadavers with i) rapid disappearance of the majority of taxa following the massive release of cadaver fluids in the soil and ii) slow recovery after the end of the active decay phase, ca. 1–2 months after the peak of cadaver fluid input in the soil. If such patterns were indeed found, this may lead to the development of new PMI indicators in the future.

1.2 Materials and methods

Sampling site

The experimental site is situated in a beech- (*Fagus sylvatica*) and oak- (*Quercus robur*, *Q. petraea*) dominated forest near Neuchâtel (Switzerland 47° 00' N; 06° 56' E, elevation 478 m). The overall average temperature measured over the course of the experiment was 10.4 °C (SD 6.09 °C). The mean annual precipitation of the nearest meteorological station (Neuchâtel) for 1993 to 2013 was 974 mm per year (MétéoSuisse).

Experimental setup

The experiment included three treatments: control (plots of forest soil left under natural conditions), fake pig (plastic bags filled with a volume of soil similar to that of the pigs placed in a cotton cloth) and pig (*Sus scrofa*). The fake pig treatment was used to differentiate the chemical effect of pig cadaveric liquids from the physical effects (i.e. humidity, soil compaction) due to the presence of a carcass on the soil. The bag volume was kept approximately similar to that of the pigs by removing soil from the bag to mirror the volume loss of the pig cadavers over time. The pigs (20 kg±1 kg) were killed on the farm with captive bolt stunning and the cadavers immediately brought to the experimental site. The cadavers were placed in strong metal wire cages (90×100×50 cm) to protect them from scavengers. The cages also allowed moving the cadavers for sampling. Pigs are commonly used in comparable forensic studies due to the similarities with humans, comparable thoracic cage size and almost naked skin (Stokes et al., 2013). Each treatment was run in triplicate. The sampling plots were organised into three randomised blocks (15–34 m apart). Within each block, the plots were at least 4 m apart.

Sampling and chemical analyses

Eight sets of samples were collected from the onset of the study (August 5th, 2009=D0) until June 21st, 2012 (Table 1.1). At the onset of the experiment (D0, before the pigs and fake pigs were placed), initial control samples were collected from all sampling plots and pooled for each block (i.e. three pooled samples in total). Sampling days were scheduled according to decomposition stages (Table 1.1) (Payne, 1965). On each sampling day, ca. 25 g of soil was taken to a depth of 10 cm in each plot and stored at -80 °C. Soil subsamples (3 g) were dehydrated (40 °C, 12 h), ground to powder and analysed for total organic carbon (Soil_C) and total nitrogen (Soil_N) using combustion infrared spectroscopy (CHNEA1108-Elemental analyser, Carlo Erba Instrument) after decarbonation with HCl (Harris et al., 2001).

Table 1.1: Sampling dates and corresponding decomposition stages of the pig cadavers in a *Fagus-Quercus* forest above Neuchâtel, Switzerland.

Sample code	Sampling date	Decomposition stage
D0	05/08/2009	Fresh
D8	13/08/2009	Active decay
D15	20/08/2009	Dry stage
D22	27/08/2009	Dry stage
D33	07/09/2009	Dry stage
D64	08/10/2009	Dry stage
D309	10/06/2010	Dry stage
D1051	21/06/2012	Dry stage

Molecular analyses

DNA was extracted from soil samples using the MoBio PowerSoil DNA Isolation Kit (Carlsbad, CA, USA) following the manufacturer instructions. The SSU rRNA V9 region was amplified by PCR using the specific eukaryotic primers 1380f/1510r (`CCCTGCCHTTTGTACACAC` / `CCTTCYGCAGGTTCACCTAC`) (Amaral-Zettler et al., 2009). Forward primers were tagged on the 5' end with a 10 nucleotides strand, specific to each sample. PCR reactions were run in triplicate with a PTC-200 Peltier Thermo Cycler (BioConcept, Allswill, Switzerland) with 1 ng of environmental DNA, 6 µL of 10x PCR buffer, 0.6 µL of each primer, 0.6 µL of each dNTP 400 µM (Promega) and 0.2 µL of 0.05 U/µL Go Taq (Promega). The volume was adjusted to 30 µL with ultra-pure water. Amplification was conducted with the following conditions: denaturation at 94 °C for 3 min, 30 cycles at 94 °C for 30 s, 57 °C for 60 s and 72 °C for 90 s and final extension at 72 °C for 10 min (Amaral-Zettler et al., 2009). PCR products were purified through QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and pooled together with a 4 ng DNA of each sample.A DNA library was prepared using the New England Biolabs's kit NEBNext DNA Sample Prep Master Mix Set 1 (`http://www.neb.com/nebecomm/ManualFiles/manualE6040.pdf`) except for the size selection step. Sequencing was done by the Genomics Core Facility at Brown University (Providence, USA) with an Illumina® HiSeq 2000 sequencer to obtain paired-end reads covering the full length of the V9 region.

Sequence treatment

A database was constructed by selecting 44 complete euglyphid V9 sequences from the GenBank database, using sequences derived both from identified organisms and from related environmental sequences retrieved from GenBank. Each environmental V9 read was compared to the database using the `BLASTn` algorithm (Altschul et al., 1997) in order to select euglyphid sequences. Before the `BLASTn`, each nucleotide with a Phred score below 28 was changed to an unknown nucleotide 'N' in order to avoid unreliable nucleotides. The `BLASTn` algorithm was setup with a match/mismatch ratio of 1:-1, gap open and extend penalty, respectively, of 0 and 2 and a word size of 32 nucleotides. We used an empirically determined e-value threshold as the criterion for classifying a read as belonging to the euglyphids. To determine the appropriate e-value threshold, a subset of eukaryotic V9

sequences (sample D309, block 2, pig treatment) was compared by BLASTn to the previously established euglyphid V9 sequence database with a permissive e-value (i.e. 10). The hit results were sorted by increasing e-value and compared to the GenBank database by BLASTn, using the previous setup, until sequences corresponding to taxa other than euglyphids were found. Once the e-value threshold was found (i.e. $8e^{-29}$), each environmental sequence was compared against the V9 euglyphid database using BLASTn. Only sequences over 130 nucleotides long and occurring at least five times in the 66 samples were retained, in order to remove possible false-positive sequences. As our database showed that some closely related but nevertheless morphologically and genetically (e.g. COI gene or full SSU) distinct euglyphid morphospecies shared exactly the same V9 sequence (e.g. *Euglypha penardi* (EF456753) and *Euglypha cf. ciliata* (EF456754); Lara et al., 2007b), we considered each unambiguous difference in the nucleotide sequence as sufficient for discriminating two OTUs. Conversely, when two sequences differed only in ambiguous nucleotides, they were considered as belonging to the same OTU. The resulting OTU sequences were then counted in each sample.

Numerical analyses

We assessed the response of the 51 OTUs found in the 66 samples to the different treatments using partial redundancy analysis (RDA) on Hellinger-transformed data (Legendre and Gallagher, 2001) with the blocks used as conditional variable. Rare OTUs (present less than three times in a minimum of ten samples) were removed to reduce noise in the model and optimise the adjusted R^2 (Borcard et al., 2011). These thresholds were selected after testing several options (presence threshold 1, 3, 5, 7; minimum number of presence 7, 10, 12, 14). The significance of variables (Soil_C, Soil_N, treatment) and ordination axes (first, second and third) were assessed using Monte Carlo tests (999 permutations, p value threshold=0.05). We assessed the effect of the treatments, relative to control, on the OTU responses over time using a principal response curve (PRC) (van den Brink and Ter Braak, 1999). The model was also tested using a Monte Carlo procedure (999 permutation, p value threshold=0.05). All statistical analyses were performed with R-2.13.1 (R_Core_Team, 2013) using package vegan (Oksanen et al., 2013) for the Hellinger transformation, RDA and PRC analyses.

Retrieval of full-length SSU rRNA gene sequences of selected taxa and phylogenetic analysis

Because sequences of the V9 variable region of the SSU rRNA gene are short (i.e. generally less than 200 bp), they are not suited for inferring the position of OTUs in phylogenetic trees. This is especially problematic if the considered sequences are suspected not to cover a large part of the diversity of the group of interest (Dunthorn et al., 2014). In order to place the OTUs of interest (i.e. showing a strong response to cadavers) in a phylogenetic tree, we used the sequence information included in the V9 region to design specific reverse primers and amplified the rest of the SSU rRNA gene. We designed specific primers to amplify specifically two phylotypes that responded positively to the pig treatment: eugly_13R (CACGAACTGAAGGCAAGCCCA) and eugly_666R (TTCACTTCCAATCACAGGAG). The newly designed primers were used in combination with the euglyphid-specific forward primer Eugly1SSUF (GCGTACAGCTCATTATATCAGCA (Chatelain et al., 2013)) located at the beginning of the SSU rRNA gene. DNA extractions, where the OTU was most abundant, were selected for specific amplification of the SSU rRNA gene of interest. Cycling profile was the same as described above (with 40 cycles). PCR products were cloned into pCR2.1 Topo TA cloning vector (Invitrogen) and used to transform E. coli TOP10' OneShot cells (Invitrogen) according to the manufacturer's instructions. Up to five clones per PCR product were chosen for sequencing. Sequencing was performed with an ABI-3130xl DNA Sequencer (Applied Biosystems). The new sequences obtained were placed into an alignment that comprised all euglyphid sequences retrieved from GenBank, which included both environmental clones and sequences derived from identified organisms. The alignment was performed using MUSCLE (Edgar, 2004). A maximum likelihood tree was built using the RAxML v7.2.8 algorithm (Stamatakis, 2006) as proposed on the portal (http://phylobench.vital-it.ch/raxml-bb/) using a general time-reversible model. Rate heterogeneity was estimated using a CAT model. The two sequences can be found on GenBank with the accession numbers KX999711.1 (eugly_666) and KX999712.1 (eugly_13).

1.3 Results

Diversity and structure of euglyphid OTU assemblages

Of the 247,366,905 raw Eukaryote reads, 187,566 were identified as euglyphids and 57,533 of these were found at least five times overall (Table 1.2). These 57,533 reads were divided into 198 OTUs. Of these, 51 OTUs

Table 1.2: Summary of the sequence filtering of euglyphid testate amoebae from the control, fake pig and pig cadaver plots in a *Fagus-Quercus* forest above Neuchâtel, Switzerland.

Analysis step	Total reads	Euglyphid sequences	Euglyphid dereplicated reads	Euglyphid OTUs
Raw fastq	247'366'905	—	—	—
Blast selection	187'566	187'566	—	—
Reads >= 5 times	57'533	57'533	2'621	—
OTUs building	52'860	57'724	2'624	198
OTUs selection	52'860	52'860	—	51
Triplication D0	57'640	57'640	—	51

Table 1.3: Summary of total euglyphid testate amoeba OTU abundance in the control, fake pig and pig cadaver plots in a *Fagus-Quercus* forest above Neuchâtel, Switzerland.

Euglyphid OTUs	Total abundance
eugly_59	4234
eugly_13	4205
eugly_2	4149
eugly_12	3873
eugly_666	3161
Euglypha rotunda AJ418783.1	3056
Euglypha filifera AJ418786.1	2583
eugly_66	2530
Uncultured eukaryote EF025028.1	2279
eugly_81	2048
eugly_151	1949
eugly_156	1933
eugly_5	1916
eugly_136	1630
eugly_183	1451
eugly_322	1400
eugly_38	1227
eugly_307	1199
eugly_54	1186
Assulina muscorum AJ418791.1	1162
eugly_113	1086
eugly_234	946
Cercomonadida env sample EF024983.1	858

Euglyphid OTUs	Total abundance
eugly_79	675
eugly_33	594
eugly_41	514
eugly_60	417
eugly_290	411
eugly_991	402
eugly_98	382
eugly_320	367
eugly_992	349
eugly_16	314
eugly_473	273
eugly_862	253
eugly_82	238
eugly_76	211
eugly_80	203
eugly_973	195
eugly_233	182
eugly_1245	177
eugly_854	176
Tracheleuglypha dentata X77698.1	176
eugly_282	172
eugly_885	170
eugly_371	165
eugly_250	141
eugly_120	134
eugly_1777	105
eugly_1716	96
eugly_1890	87

respected the thresholds and were thus retained for further analyses. Six of these OTUs matched exactly with sequences from our database. Total OTU abundance data are summarised in Table 1.3, and OTU richness and number of euglyphid reads along time for the three treatments are shown in Figure 1.1.

The partial redundancy analysis (RDA, Figure 1.2) with the blocks used as conditional variable revealed a significant correlation between euglyphid communities and Soil_N and Soil_C (Monte Carlo test, 999 permutations, both p=0.01). Axes 1 and 2 were significant (p=0.005 for both). The RDA ordination showed that the pig treatment samples diverged from the control and fake pigs along the soil nitrogen content gradient until day 309 after which they converged again with the samples of the other two treatments. The RDA also showed that most OTUs responded negatively to the pig treatment. However, some OTUs responded positively to the pig treatment (e.g. eugly_13, eugly_666).

The principal response curve (Figure 1.3) summarises the treatment effects on OTUs over time and shows the average responses of individual OTUs. The first PRC axis explained significantly ($p<0.03$) 42 %

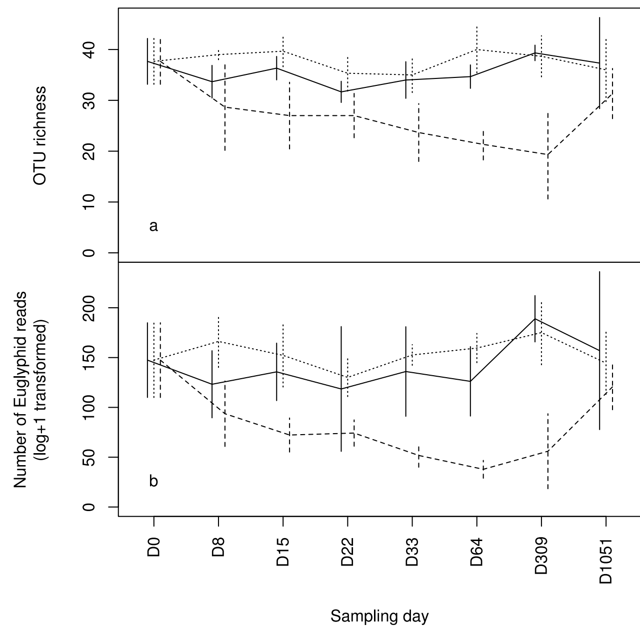


Figure 1.1: Temporal patterns of euglyphid OTU richness (a) and number of reads log +1 transformed (b) in soil samples from control, fake pig and pig cadaver plots in a *Fagus-Quercus* forest above Neuchâtel, Switzerland. Treatments are represented by line type (plain: control, dashed: pig, dotted: fake pig). The vertical lines show the standard deviation of the richness and number of reads for each treatment and sampling date. The lines for the three treatments are slightly offset to improve readability.

of the model variance, while time and treatments explained, respectively, 10 and 27 % of the variance. Qualitatively, the PRC diagram showed an overall negative effect of the pig treatment (D8 to D1051) on the majority of euglyphid OTUs and the positive response of a few OTUs, especially eugly_666 and eugly_13, which were therefore further studied.

Retrieval of full-length SSU rRNA gene sequences of eugly_13 and eugly_666

All clone sequences obtained were identical (n=5 and 8, respectively). Phylogenetic analyses confirmed the position of the two phylotypes within euglyphid testate amoebae (supported with 100 % bootstrap value) and showed that they did not belong to any barcoded family (Figure 1.4) (Heger et al., 2010). They were basal to all known euglyphid families.Eugly_13 branched robustly (80 % bootstrap) with an environmental sequence from freshwater sediments (freshwater 13_2.2 AY620297). By contrast, eugly_666 did not branch robustly with any sequence—be it from environmental samples, cultures, or isolated cells.

1.4 Discussion

Euglyphid community responses to decomposing pigs

This study showed that the presence of decomposing pig cadavers significantly affected the community structure of

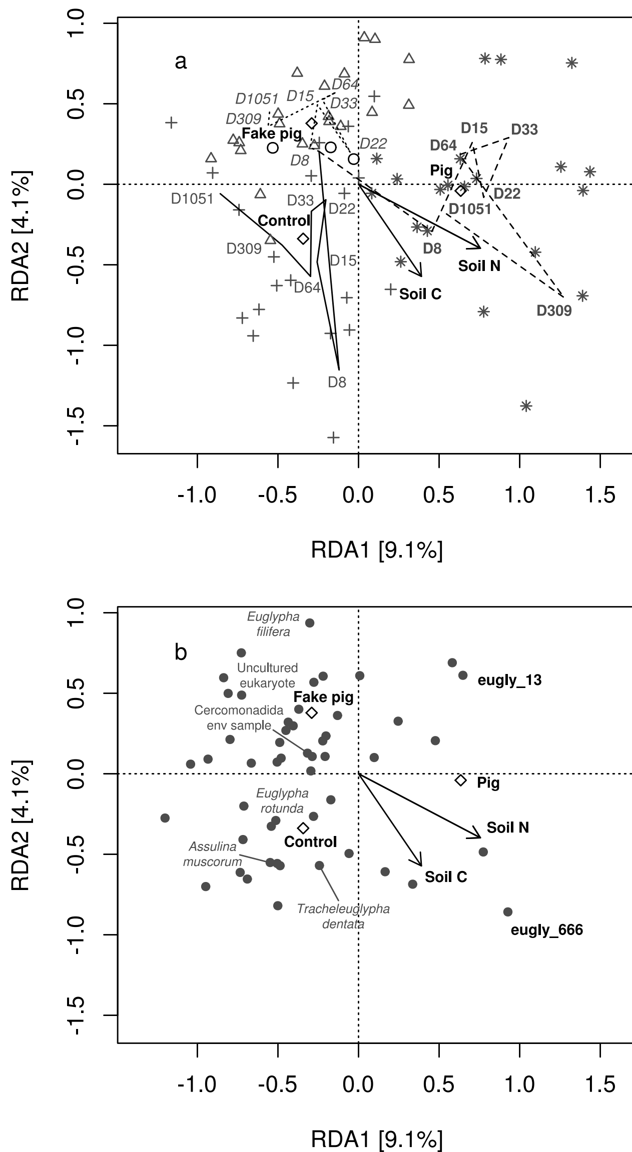


Figure 1.2: Partial redundancy analysis (RDA) ordination diagram showing the temporal patterns of soil euglyphid testate amoeba communities (OTUs) in control, fake pig and pig cadaver plots in a *Fagus-Quercus* forest above Neuchâtel, Switzerland. Diamonds represent treatment centroids and arrows represent weight percentage of total organic carbon (Soil C) and total nitrogen (Soil N). Percentages of variance explained by axes are shown in brackets. In Figure 2a, successive sampling dates for each treatment (Day 0, 8, 15, 22, 33, 64, 309, 1051) are connected by lines. The lines corners correspond to the centroid (average for RDA1 and RDA2 coordinates) of three samples of the same treatment and same sampling day. Treatments are indicated by line type (plain: control, dashed: pig, dotted: fake pig), fonts of the sample labels (plain: control, bold: pig, italic: fake pig) and symbols ("+": control, triangle: fake pig, star: pig). Empty circles represent Day 0 for each block. In Figure 2b, OTUs responses are represented by dots. The two OTUs for which full SSU sequences were obtained are represented in bold. The OTUs with a peffect match with a data base sequence are represented with their names.

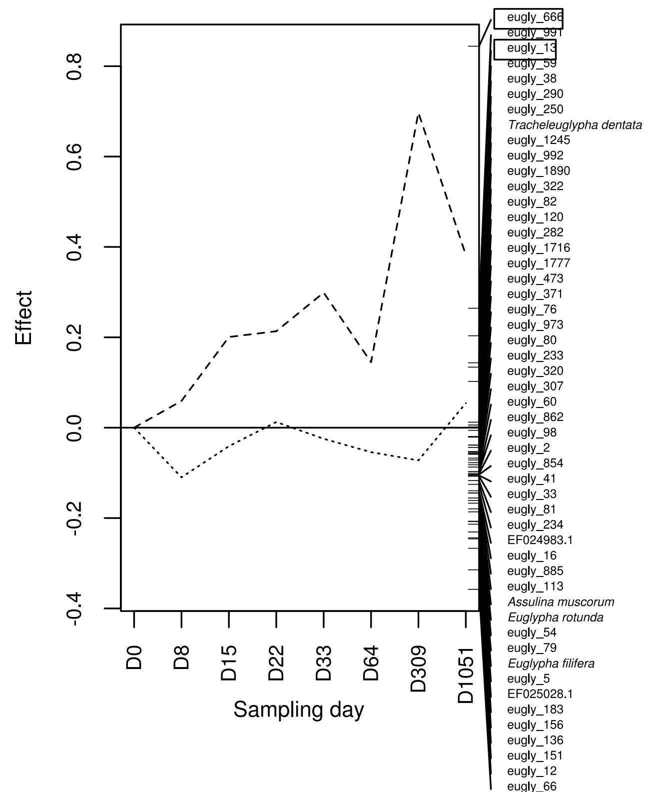


Figure 1.3: Principal Response Curve (PRC) diagram showing the effects of pig (dashed line) and fake pig (dotted line) treatments relative to control treatment over time on soil euglyphid testate amoeba communities in a *Fagus-Quercus* forest above Neuchâtel, Switzerland. The average response of individual OTUs is shown on the right axis. The two OTUs for which full SSU sequences were obtained are framed. GenBank accession numbers represent sequences that matched perfectly with the database.

euglyphid testate amoebae, showing a drastic decrease in sequence abundance and in OTU richness (Figure 1.1). This result is in agreement with our general working hypothesis. The negative effect of a cadaver on euglyphid communities was correlated to the large input of nitrogen and organic carbon in the soil. This result was consistent with previous studies, which show that inputs of nitrogen strongly and negatively influenced testate amoeba communities (Gilbert et al., 1998a,b; Mitchell, 2004). It is probable that most euglyphids died because of anoxic conditions, but a direct or indirect effect of high nitrogen content is also possible.

However, two well-represented OTUs, namely eugly_13 and eugly_666 (eugly_991 also responded positively but was less abundant), responded positively to the presence of cadavers, but only in the late decomposition stage (i.e. after 1 month to 1 year). These OTUs were present but rare at the beginning of the decomposition process as well as in the control and fake pig treatment, but their abundance peaked, respectively, at D33 and D309 in the pig treatments only and in the three replicates simultaneously (Figure 1.5). This suggests that they did not benefit from the initial perturbation brought by the release of cadaveric fluids but rather found optimal conditions (i.e. abiotic, e.g. soil water chemistry, and biotic, e.g. prey and/or low level of competition or predation) for their growth in later stages. These organisms probably benefited from changes in the bacterial communities, as these are supposed to change deeply and progressively underneath a cadaver (Allison and Martiny, 2008; Fierer et al., 2012; Ramirez et al., 2010). Indeed, previous studies have shown that decomposing carcasses cause an increase in soil bacterial biomass (Barton et al., 2013) but also drastically change bacterial community structure (Howard et al., 2010; Moreno et al., 2011). As

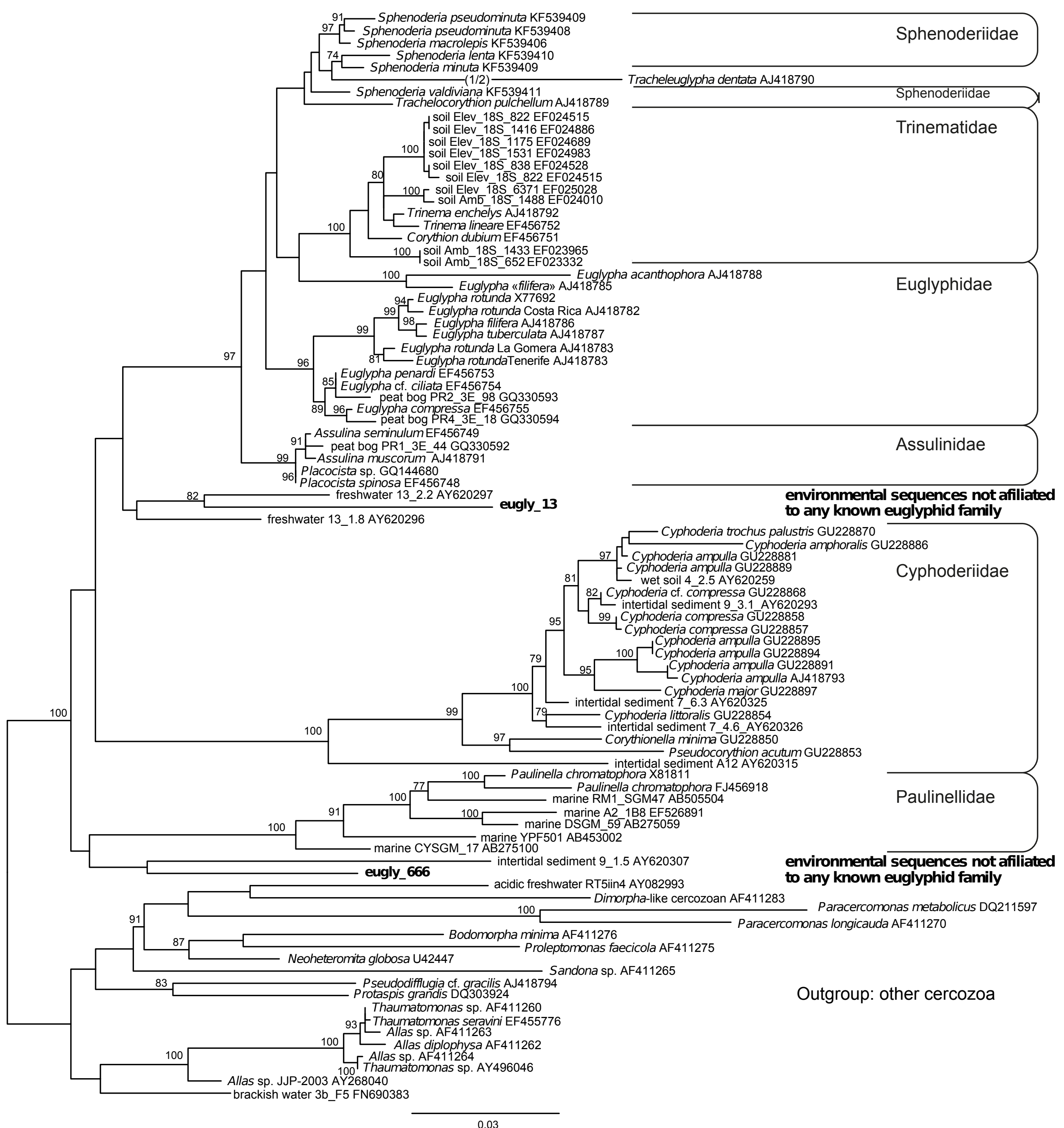


Figure 1.4: Maximum likelihood tree built on full SSU rRNA gene sequences of Euglyphida showing the phylogenetic position of eugly_13 and eugly_666 full length sequences (in bold) obtained from pig cadaver plots in a *Fagus-Quercus* forest above Neuchâtel, Switzerland. The tree was built using RaxML on 1440 positions and rooted with several cercozoan sequences retrieved from GenBank. Robustness of nodes was evaluated by bootstrapping (1000 replicates).

bacteria constitute a large part of euglyphid food regime (Meisterfeld, 2000b), any change in the abundance or community structure of bacteria is likely to also influence the abundance and community composition of euglyphids. It may also be that these taxa represent nutrient-tolerant organisms that benefit well from high abundance of prey organisms, but with low competitive ability in the normally more oligotrophic conditions. The precise mechanism for this response however remains to be elucidated.

A possible bias could have been due to the import of euglyphids with the cadavers, either from the farm or during transport. However, at D8, cadaver samples were less different from the control than samples from fake cadavers (which could not have been contaminated by the new plastic bags and cotton cloth), and we therefore conclude that such contamination was negligible.

The effect of cadavers on euglyphids peaked at D309 (Figure 1.3). This time interval seemed quite long in comparison with the results obtained by Szelecz et al. (2014) from the same field experiment (i.e. complete die-out of testate amoebae 22 days post mortem). However, Szelecz et al. (2014) studied litter and not the underlying mineral soil horizon as done here, and they used a direct observation (microscopy) approach,

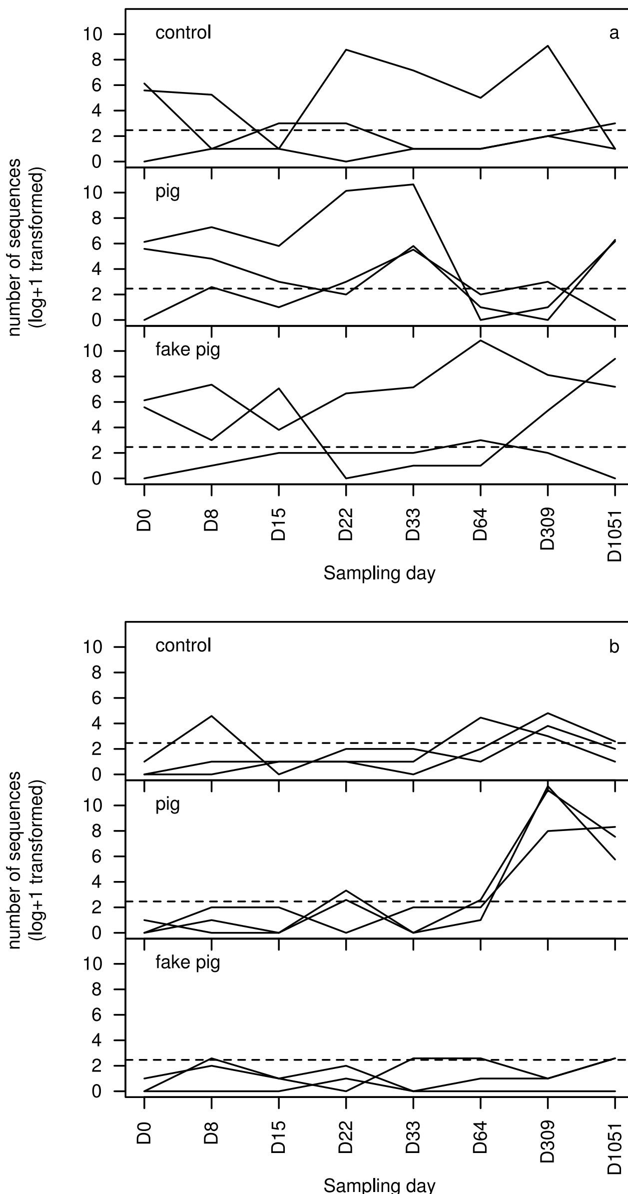


Figure 1.5: Temporal pattern of number of sequences (log+1 transformed) over time in control, fake pig and pig cadaver plots in a *Fagus-Quercus* forest above Neuchâtel, Switzerland for euglyphid testate amoeba OTUs eugly_13 (a) and eugly_666 (b). Full lines represent the number of sequence in each block. The dashed lines represent the average sequence number in a sample independently of the treatment, block or sampling.

which most likely underestimated diversity. Indeed, as OTUs eugly_13 and eugly_666 did not belong to any known euglyphid family (Figure 1.4), their morphology may differ significantly from known forms, they may be very small and thus overlooked or lost in classical sample preparation protocols using 10–20-µm filters and/or they may represent naked forms (i.e. without test), as documented in foraminiferans (Pawlowski et al., 1999).

The RDA and PRC both showed that euglyphid communities had not completely recovered from the influence of cadavers by the end of the experiment (i.e. D1051). This pattern is in line with the observations of Szelecz et al. (2014) who did not observe a full recovery by D309 (end of their experiment). This long resilience time suggests that euglyphid communities (and probably testate amoebae in general) could be used as indicators of cadaver presence over very long periods. The fact that euglyphid communities still indicated an effect of cadavers either shows a lag in return to pre-disturbance community structure or that they still responded to other differences (e.g. soil chemistry, abundance and composition of prey).

In addition—and this is in itself an unexpected result—the positive response of certain euglyphid OTUs at certain time points (eugly_13 at D33, eugly_666 at D309) (Figure 1.5) suggests that individual taxa may respond specifically and positively to some decomposition stages. Such patterns suggest the possibility to use soil testate amoebae as bioindicators for estimating the time elapsed since death (post-mortal interval, PMI), a parameter of considerable importance in forensic sciences.

Unknown diversity of soil euglyphid testate amoebae

after removing rare OTUs, we still found 51 OTUs, 45 of which did not match any sequence in the database. The V9 region does not allow discrimination between close-related species, and it is unclear to how many morphologically and genetically different taxa these 51 OTUs correspond. Regardless of the short length of the barcode, these results reveal the existence of a very high overall diversity of euglyphids in forest soils. This technique yields large amounts of data from small sample volumes, requires much less taxonomic expertise than classical morphological analyses and does not depend on the existence of a reliable taxonomy (which is often lacking for protists).

Perspectives and potential future application

Focusing our study on a specific taxonomic group allowed us to define OTUs at high resolution, using a threshold adapted to already barcoded morphospecies. This approach allowed us to use metabarcoding at a taxonomic resolution close to morphological analysis—much more than what is generally achieved in studies using general eukaryotic marker. Indeed in most studies, more sequences are pooled into OTUs, each of which corresponds to broader taxonomic units than what we achieved in this study. The approach we used to study the response of euglyphid testate amoebae to the impact of decomposing cadavers can also be used to study the responses of any other group of soil eukaryotes. It is indeed very likely that many other taxonomic groups will also show comparable responses to those documented here for euglyphids. Our study shows that some of this unknown diversity could be of potential use for applied purposes such as forensic science. If such patterns can be explored in details, we believe that it will be possible to develop accurate and reliable new molecular bioindicator tools for PMI estimations and other applications.

Acknowledgements

This study was funded by the Stiftung Forensisches Forum Frankfurt/Main, Germany; Vereinigung von Freunden und Förderern der Goethe-Universität, Frankfurt/Main, Germany; the University of Neuchâtel, Switzerland; and the Swiss National Science Foundation (project. No 31003A_141188 to E.M.). We thank Jan Boni (forest engineer of Neuchâtel) for providing the permission to use the site, Dr. Sonia Estelle Tarnawski and Luc Dolivo for the help with DNA extraction, Dr. Roxane Kohler-Milleret for infrared spectroscopy and Emanuela Samaritani for the assistance with the Illumina sequencing. We also thank Balazs Laurenczy for the constructive discussion on the informatics part.

Chapter 2

Micro-eukaryotic diversity and co-occurrence network structure change drastically across micro-environmental gradients in Sphagnum peatlands

David Singer[1*+], Christophe V.W. Seppey[1*+], Battle Karimi[2], Edward A.D. Mitchell[1,3], Enrique Lara[1]

[1]Laboratory of Soil Biodiversity, University of Neuchâtel, Rue Émile-Argand 11, 2000 Neuchâtel, Switzerland
[2]Agroecology, French National Institute for Agricultural Research (INRA), Rue Sully 17, 21065 Dijon, France
[3]Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland
[*]**Corresponding author:** david.singer@unine.ch, christophe.seppey@unine.ch
[+]**Co-first authors**

Abstract: Micro-eukaryote communities are hyper-diverse, support various functional groups and maintain a high number of biotic interactions. In *Sphagnum* peatlands, these interactions can be studied along several environmental gradients such as nutrients, pH and humidity. We hypothesize that the harsh and stable conditions of this habitat pushed the *Sphagnum* inhabitant micro-eukaryotes to co-evolve and develop tight biological linkages. From environmental amplicons sequencing of micro-eukaryotic communities taken from four peatlands across the Jura mountains, we assessed the OTU co-occurrence networks from four characteristic micro-habitats distributed along a gradient of nutrients and humidity. Ordination analyses showed that micro-habitats had an effect on communities and that their diversities increased according to nitrogen concentration and humidity. Harsher micro-habitats also showed a higher connectance and number of first order links that mainly pass by low trophic levels heterotrophs. These results suggest that low diversity micro-habitats can support tightly linked communities and a bottom-up trophic network.

Keywords: Micro-eukaryotes . peatlands . environmental amplicon sequencing . co-occurrence network . functional ecology . environmental gradient

2.1 Introduction

Latest technological developments like high throughput sequencing (HTS) are currently revealing an immense genetic diversity of microbial eukaryotes in the environment (Mahé et al., 2017; Pawlowski et al., 2011; de Vargas et al., 2015). These organisms, tagged protists, algae and fungi, are functionally diverse and may act as primary producers, saprotrophs, symbionts (mutualists and parasites) or predators. These broad functions group many subcategories; e.g. some predators are highly specialised on certain preys (Hess and Melkonian, 2013) while others are more generalist (Johnke et al., 2014). Mutualistic symbiosis, often implying a phototrophic and a heterotrophic partner, is very common in marine systems (de Vargas et al., 2015). The photosynthetic partner is often preadapted to form symbioses and will be found within larger eukaryotes, which often cannot live without them (Gomaa et al., 2014; Anderson, 2014). Parasitism is extremely widespread in the environment (Geisen et al., 2015a; Mahé et al., 2017), and most often implies specialized partners. These broad functional categories can also be cumulated, the most frequent and documented example being mixotrophy, where an organism combines heterotrophy and phototrophy. This strategy has been shown to be particularly efficient in oligotrophic environments like oceans (Unrein et al., 2014) or peaty lakes (Lara et al., 2015). Therefore, in line with its immense genetic diversity, functional diversity in microbial eukaryotes fulfils an immense part of the roles micro-

organisms can have in an ecosystem.

In natural systems, these organisms interact very intensively with each other. These interactions constitute complex networks called interactomes (Lima-Mendez et al., 2015), which can be measured with a series of metrics which evaluate their structure and robustness (Karimi et al., 2017). In environmental microbiology, these networks are computed on complex communities as provided by environmental DNA surveys, mostly generated by HTS (Barberan et al., 2012). There, each co-occurrence between two OTUs is used as a proxy for interaction. The exact nature of the interaction between the organisms is in most cases unknown; thus, it is good practice to follow the few known interactions to validate the others (Lima-Mendez et al., 2015).

Patterns of diversity (including species richness and diversity indices) together with co-occurrence network analysis can reveal ecological thresholds along environmental gradients and tipping points beyond which perturbations such as pollution (Karimi et al., 2016; Payne et al., 2013), climate change (Delarue et al., 2015), or land-use (Lentendu et al., 2014) can cause shifts in the functioning of ecosystems. While the function of most prokaryotic OTUs encountered in the environment is still unknown, much more can be told about eukaryotes. Functional assignation to eukaryotic environmental sequences, although not always straightforward, can be inferred at least to a certain extent. Indeed, entire groups can be labelled as primary producers, like diatoms and chlorophytes for instance. Within ciliates, entire clades can be labelled as first order (like Scuticociliatia) or second order predators (Haptoria and Suctoria) (Lara and Acosta-Mercado, 2012). Other groups are entirely mixotrophic like Synurales (Stramenopiles). Based on these functional assignations, it is possible to interpret co-occurrence networks beyond usual metrics, in order to evaluate which organisms are keystone in an ecosystem and why.

In this work, we surveyed the diversity of micro-eukaryotic communities associated to the different micro-habitats of *Sphagnum*-dominated peatlands, and followed the structure of their interaction networks in order to understand their structure and characterize which organisms are essential elements for the maintenance of the whole ecosystems. In other terms, which organisms are keystone species, i.e. their removal would disrupt the network with probably dramatic consequences on the ecosystem as a whole.

Sphagnum-dominated peatlands are of global significance as they are responsible for one third of the whole amount of C sequestered in soils at the global level (Gorham, 1991; Clymo and Hayward, 2012). This function may be threatened by the on-going global warming. Indeed, under a global change scenario, peatlands tend to become a C-source which emits more greenhouse gases than it may stock C. This phenomenon is principally caused by an increase of microbial respiration (Dieleman et al., 2015; Gorham, 1991) and the loss of mixotrophic organisms which are responsible for significant amounts of C mineralization (Jassey et al., 2015). Hence, it is of utmost importance to understand where the weak points of interactions networks are to understand the resistance of these microbial communities to environmental changes.

In a peatbog, several micro-habitats can be easily characterised based on topology and also on associated plant communities. These micro-habitats are characterized by different values of water table depth and hydrochemistry (pH, amounts of mineral nutrient and water table depth) (Batzer and Baldwin, 2012; Rydin and Jeglum, 2013; Hajek et al., 2006; Gilbert and Mitchell, 2006). They can be classified in function of gradients; for instance the dryness gradient ranges micro-habitats from hummocks (i.e. small mounds ca. 30cm above the average water table) to pools. Additionally, peat bogs are notably deprived of nutrients like nitrogen (Lamers et al., 2000), especially in areas far from the water table like hummocks (Singer et al., 2016). Peat bogs forests however are an exception to the rule as the trees provide nutrients from their litter regardless of a low water table because of the trees evapotranspiration (Bragazza et al., 2007).

In this study, we aimed at characterizing both the micro-eukaryotic diversity and interactome through different micro-habitats within *Sphagnum* peatlands, in order to infer keystone organisms. Microbial communities were assessed using high throughput environmental amplicons sequencing (Illumina, V9 SSU rRNA) and the Operational Taxonomic Units (OTUs) were identified taxonomically and functionally. We hypothesized that the most restrictive micro-habitats (here: hummocks) should be less diverse than other more nutrient-rich and wet microsites. On the other hand, we assume that the number of interactions increase with harshness of the environment, in line with observations on numerous symbiotic associations in nutrient-deprived environments (Lara and Gomaa, 2017). And, finally, as nutrients are deficient in peatlands, bacteria and yeasts should be limiting, and therefore, systems should be bottom-up regulated, which implies that first order predators (i.e. small bacterivorous and mixotrophic protists) should be the keystone species in such environments.

2.2 Materials and methods

Sampling and environmental variables

Sphagnum spp. samples were taken from five *Sphagnum*-dominated peatlands (Le Cachot, Frasne, Praz-Rodet, Étang de la Gruère and Ponts-de-Martel) in the Jura Mountains in Switzerland and France. The sites were located at similar elevation (900-1000m a.s.l.) and have the same geomorphological settings. Sampling was conducted over a 4-days period in summer (24-28.06.2013). In each site, four contrasting micro-habitats were sampled: hummock, lawn, pool and forest (Singer et al., 2016). These four micro-habitats were chosen because they capture the main environmental gradients in *Sphagnum* peatlands (Rydin and Jeglum, 2013). In order to assess the intra-peatland variability five samples per micro-

habitats were taken in le Cachot bog while two replicates per micro-habitat were taken in the four other sites. At each sampling site, we collected ca. 2g of fresh *Sphagnum* in sterile conditions and samples were kept in DNA preservation buffer (Lifeguard, MoBio Carlsbad CA, USA).

The water table depth (WTd) was measured during the sampling in a 2cm diameter auger hole after ca. 30 minutes to allow the water level to stabilise (Singer et al., 2016). Total carbon (%) and nitrogen (%) were measured by combustion infrared spectroscopy (CHNEA1108-Elemental analyser, Carlo Erba Instrument) and pH_{H20} was measured according to the method proposed in Carter and Gregorich (2007).

Molecular analysis

The environmental DNA was extracted from each *Sphagnum* sample with the MoBio PowerSoil DNA Isolation Kit (Carlsbad, CA, USA) following the manufacturer instructions. The V9 region of the rDNA SSU was then amplified using the general eukaryotic primers 1380f/1510r (CCCTGCCHTTTGTACACAC / CCTTCYGCAGGTTCACCTAC) according to (Amaral-Zettler et al., 2009). The PCR reactions were run with a PTC-200 Peltier Thermo Cycler (BioConcept, Allswill, Switzerland) with 3 µL of environmental DNA, 6 µL of 10x PCR buffer, 0.6 µL of each primer, 0.6 µL of each dNTP 400 µM (Promega) and 0.2 µL of 0.05 U/µL Go Taq (Promega) in a volume adjusted to 30 µL. The PCR reactions were conducted following a denaturation at 94 °C for 3 min, followed by 40 cycles at 94°C for 30 sec, 57 °C for 60 sec and 72 °C for 90 sec, and a final extension at 72 °C for 10 min (Amaral-Zettler et al., 2009). The amplification products were then purified through a QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). Amplicons were sequenced paired-end using Illumina HiSeq HTS technology (Fasteris, Geneva, Switzerland).

Bioinformatic analysis

The environmental sequences were merged with the program `flash` (Magoc and Salzberg, 2011). Sample tags and primers were then trimmed and sequences were demultiplexed using a custom script (available on request). Low quality sequences were removed from the analysis if the average of the probabilities of miss identification (based on the phred score) on a window of 50 nucleotides was higher than 0.01. Chimeras were removed from each sample by, using the `vsearch` program (Rognes et al., 2016). In order to avoid artefacts, we retained only sequences occurring at least three times in at least two samples according to the method used in de Vargas et al. (2015). We then clustered sequences into OTUs with the program `swarm` v. 1.2.5 (Mahé et al., 2014). OTUs were finally taxonomically assigned by aligning the dominant sequence of each OTU to the trimmed PR^2 database (Guillou et al., 2013) using the `ggsearch36` program (`fasta` package; Pearson, 2000).

Numerical analysis

Community matrix preparation

Every OTU assigned with a percentage of identity (pid) below 80% were regarded as undetermined, following the method proposed in de Vargas et al. (2015). We also removed sequences affiliated to Embryophyceae and Metazoa or to prokaryotes (pid < 65; threshold determined manually by aligning low pid environmental sequences (60-80%) to the GenBank database with `blast` algorithm). The number of sequences per sample was then normalised to 10'000 using the function `rrarefy` (`vegan` package; Oksanen et al., 2013).

Diversity analysis

We first determined whether community composition was more influenced by the micro-environment or by the peatland. For this purpose, we performed a non-metric multidimensional scaling analysis (NMDS) on all micro-habitats. The effect of the environmental variables (C, N, pH, WTd), peatland and micro-habitat were then tested on the NMDS ordination using a permutation test (10'000 permutations, function `envfit` package `vegan` v. 2.0-10; Oksanen et al., 2013). Supplementary NMDSs were calculated for each pair of micro-habitats in order to assess which communities differ from each other. Then, we calculated common metrics used to characterize diversity, i.e. species richness (R) and Shannon index of diversity (H) for each sample. These calculations were performed with the function `diversity`, in the package `vegan` v. 2.0-10 (Oksanen et al., 2013). Based on these metrics, we characterized the diversity of the different micro-habitats, and determined which ones were significantly different from each other, respectively. For this purpose, we performed a multiple comparisons rank sums test (Nemenyi test; Hollander and Wolfe, 1999) using the function `posthoc.kruskal.nemenyi.test` in the R package `PMCMR` v. 4.1; Pohlert, 2014). In order to avoid the noise caused by rare OTUs, we only kept the ones present in more than 10% of the 51 samples for analysis (in this case, in 6 samples or more) (Schiaffino et al., 2016).

Co-occurrence network

For network analyses we considered only OTUs that were present in at least half of the samples of each micro-habitat, respectively. In order to estimate the variability of the microbial networks and to compare the networks structure between micro-habitats, we bootstrapped six replicates of network using randomly 10 from the 13 (12 for the lawns) samples. A correlation matrix was then computed (`Hmisc` package v. 4.0-2; Harrell Jr and Dupont, 2016) to identify non-random co-occurrences between pairs of OTUs for each micro-habitat or bootstrap replicate. In order to test each correlation between OTU pairs, p-values were corrected and adjusted by the False Discovery Rate method for multiple comparisons (Benjamini and Hochberg, 1995). The significant

co-occurrences were selected for adjusted p-value below the 0.05 threshold and were recorded in an adjacency matrix. Correlations values and directions (i.e. positive or negative) were used to estimate the interaction weights. Non-significant correlations were interpreted as absence of links between OTUs and were replaced by 0 in the adjacency matrix. Co-occurrence network was inferred using this adjacency matrix and structure network indices were computed. For each bootstrap sample, we assessed the connectance (package `statnet` v. 2016.9; Handcock et al., 2016), the betweenness centrality for each OTU and the number of first degree positive links (package `igraph` v. 1.0.1; Csardi and Nepusz, 2006). Betweenness centralities and number of first degree links were summed according to each functional group in order to assess which one maintains more links. The taxonomy of each OTU involved in the networks was individually verified by aligning the sequence on the GenBank database. The effect of micro-habitats and functional groups on the connectance was tested by Nemenyi test (Hollander and Wolfe, 1999) (function `posthoc.kruskal.nemenyi.test`, package `PMCMR` v. 4.1; Pohlert, 2014). All statistical analyses were conducted with the R language v. 3.1.0 (R_Core_Team, 2014).

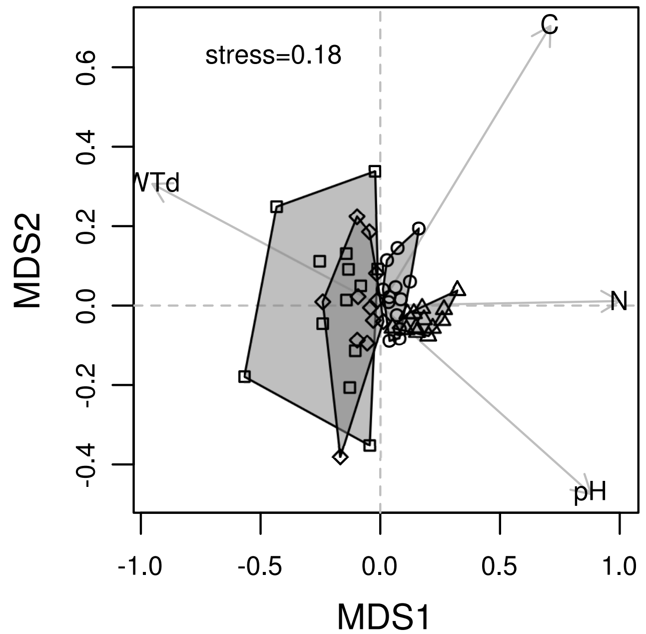


Figure 2.1: Non-metric multidimentional scaling (NMDS) calculated on eukaryotic OTU communities from *Sphagnum* taken from four micro-habitats (forests: rounds, hummocks: squares, lawns: diamonds, pools: triangles) in five peat bogs in Switzerland and France. Arrows show the projection of environmental variables on the communities.

2.3 Results and discussion

Overall diversity

We obtained a total of 4'625 OTUs, mostly rare organisms, from which we kept 1'321 OTUs for the analyses on diversity. From these, communities were dominated by Fungi, mostly Ascomycota (29% of sequences: 151 OTUs), Mucoromycota (14% or sequences: 17 OTUs) and Basidiomycota (10% of sequences: 115 OTUs). Chytridiomycota, Cryptomycota and Glomeromycota were also abundant, especially in pools (Figure S2.1). Stramenopiles, the second most abundant group in all micro-habitats, were represented by the phagotrophic/mixotrophic Chrysophyceae/Synurophyceae (6% of sequences: 184 OTUs), and the often parasitic Oomycota (6% of sequences: 18 OTU). These two taxonomic groups had already been identified as common peatland inhabitants (Singer et al., 2016; Lara et al., 2010). In addition, four other supergroups were found in all the four micro-habitats, namely Alveolata, Chlorophyta, Cercozoa and Cryptophyta. The percentage of OTUs related to Fungi (excepted Chytridiomycota and the insect parasites Entomophthtromycota) decreases with WTd. In contrast, photosynthesising organisms (Dinophyceae, Chlorophyceae, Cryptophyceae and phototrophic Chrysophyceae) follow the opposite trend. Forests stand out as an exception, possibly due to the higher amount of nutrients brought by pine leaves (Rydin and Jeglum, 2013).

Table 2.1: P-values obtained from the permutation tests (10000 permutations) calculated on the NMDSs. The left part shows the significance of the two factors (peat bog and micro-habitats) and four variables (N, C, pH, WTd) on the micro-eukaryotic communities from *Sphagnum* taken from four micro-habitats in five peat bogs in Switzerland and France. The right part shows the significance of the difference between each pair of community according to the micro-habitats (forest: F, hummocks: H, lawn: L, pools: P).

Variables	p-value
peat bog	0.9
micro-habitat	< 0.01
N	< 0.01
C	0.7
pH	0.2
WTd	< 0.1

	F	H	L
H	< 0.001		
L	< 0.001	> 0.5	
P	< 0.001	< 0.001	< 0.001

OTU communities differ significantly between all micro-habitats, with the exception of samples coming from hummocks and lawns which share similar ordination space on the NMDS (Figure 2.1, S2.3). The permutation test realised on fitted factorial variables onto the NMDS also showed that communities differed significantly among micro-habitats ($P < 0.001$) but not among peatlands ($P > 0.05$). N ($P < 0.01$) and WTd (marginally; $P < 0.1$) were also significantly correlated to community composition. R (richness) was highest in forests and decreased sequentially to pools, lawns and was lowest in hummocks (Figure 2.2). Grossly the same pattern could be observed for Shannon diversity where forests and pools harbour higher values then lawns and hummocks. As predicted, richness was lowest in the most constrained environment (hummocks) because of lack of nutrients and free water (Figure S2.2).

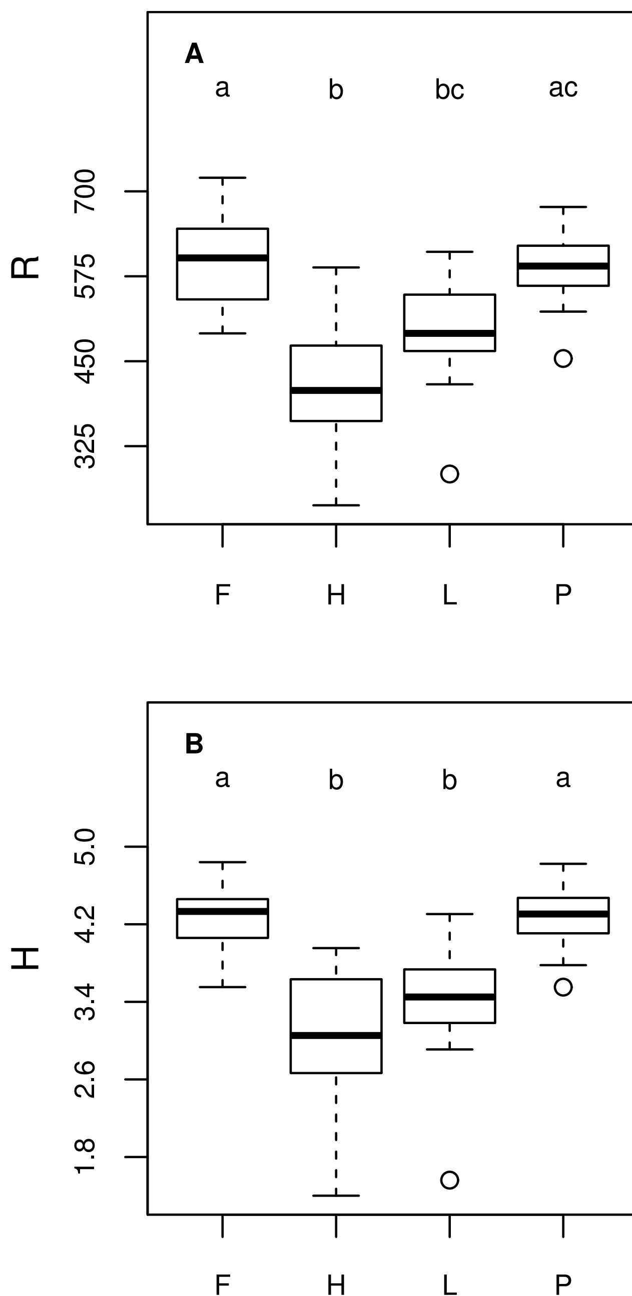


Figure 2.2: Ecological diversity metrics (OTU richness: A, Shannon index: B) calculated on eukaryotic OTU communities from *Sphagnum* taken from four micro-habitats (forests: F, hummocks: H, lawns: L, pools: P) in five peat bogs in Switzerland and France. The letters above the boxplots indicate the significantly different distributions according to a multiple comparisons of mean rank sums test (Nemenyi test) to the threshold of 0.05.

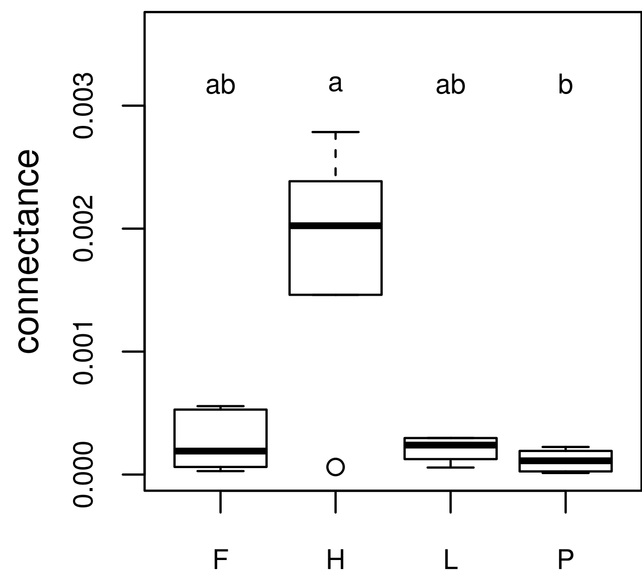


Figure 2.3: Networks connectance calculated on OTU communities retrieved from four micro-habitats in peat bogs of the Jura mountains. The distributions were obtained from 6 bootstraps using 10 samples. The letters above the boxplots indicate the significantly different distributions according to a multiple comparisons of mean rank sums test (Nemenyi test) to the threshold of 0.05.

Co-occurrence network

The co-occurrence networks were calculated on 597, 361, 419 and 559 OTUs appearing in at least half of the samples of forests, hummocks, lawns and pools respectively. Connectance was significantly higher in hummocks than in pools and shows intermediate values in the two other micro-habitats (Figure 2.3). Accordingly, we observed a higher number of first degree links in hummocks (409) than in forests (110) and lawns (131), and very few links in pools (6) (Figure 2.4, S2.4, S2.5). Models show that a system characterized by higher connectance and number of first degree links should be highly resistant to secondary species loss after random species extinction, even if diversity is low (Dunne et al., 2002). This would suggest that hummocks would be the most resistant micro-habitat studied here, despite a lower diversity. In addition, a high connectance is also thought to be linked to a higher resilience of a community to perturbations (Tylianakis et al., 2010). As example, a community of generalist may be thought as more buffered because, if a prey or predator disappear, it will be replaced quickly by an other (**?**). This high connectance can be explained by the harshness of the micro-habitat. The restrictive conditions of low-nutrient and dryness would select only a few number of species which can be associated to a higher connectivity (May, 1972).

Keystone species

The OTUs characterized by highest betweenness centrality values can be considered as keystone species in an environment (Karimi et al., 2017). The most central OTUs which presented the highest numbers of links (keystone OTUs) were, respectively, a mixotrophic chrysophyte related to *Chromulina* sp. (lawns), a small Cercozoan (hummocks) and another mixotrophic chrysophyte (forests) (Figure S2.4). Beyond these examples, most keystone OTUs represented organisms that could be classified either as first order predators feeding on bacteria, or on small mixotrophs equally preying on bacteria. Saprotrophs like Fungi developed also many links although to a lesser extent. Parasites, phototrophs and second order predators were clearly less keystone as they had fewer links and were less central. This can be illustrated when the betweenness centrality is summed according to the functional groups in each micro-habitat (Figure 2.5). Indeed, first order predators, mixotrophs

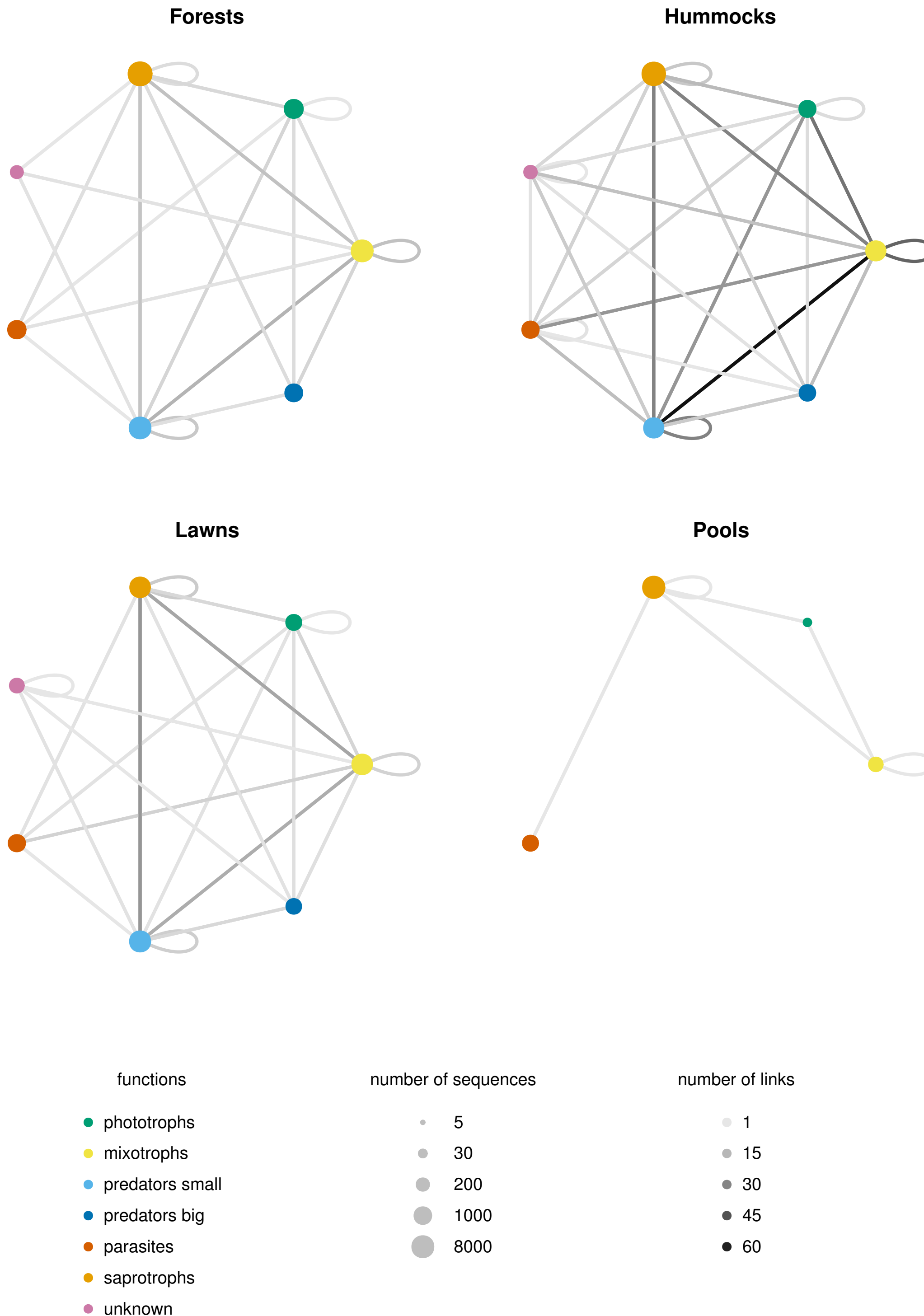


Figure 2.4: Number of positive links between each pair of functional groups from co-occurrence networks of OTUs retrieved from four micro-habitats in peat bogs of the Jura mountains. The size of the points represent the number of sequences found in the micro-habitat for each functional group. The shade of each line represents the number of links between two functional groups for a specific micro-habitat.

and saprotrophs always dominate. These results suggest that the stability of the communities is guaranteed by bacterivorous protists plus some fungi, which in turn points towards a bottom-up control of foodwebs. This appears logical since bioavailable nutrients are the limiting factor in peatlands, which maintain thus relatively low numbers of bacteria compared with more productive systems (Elser et al., 1995). Strictly phototrophic organisms, extremely common in highly productive systems such as eutrophic freshwater plankton (Falkowski and Raven, 2013) probably cannot easily obtain N in a pristine peatland, whereas mixotrophic feeding may be much more advantageous.

Nucleomorph presence

A link between two OTUs (X176 and X324) in pools, assigned respectively to a *Cryptomonas* sp. and a cryptophyte nucleomorph attracted our attention (Figure S2.4)(Richards et al., 2005). Indeed, the two OTUs

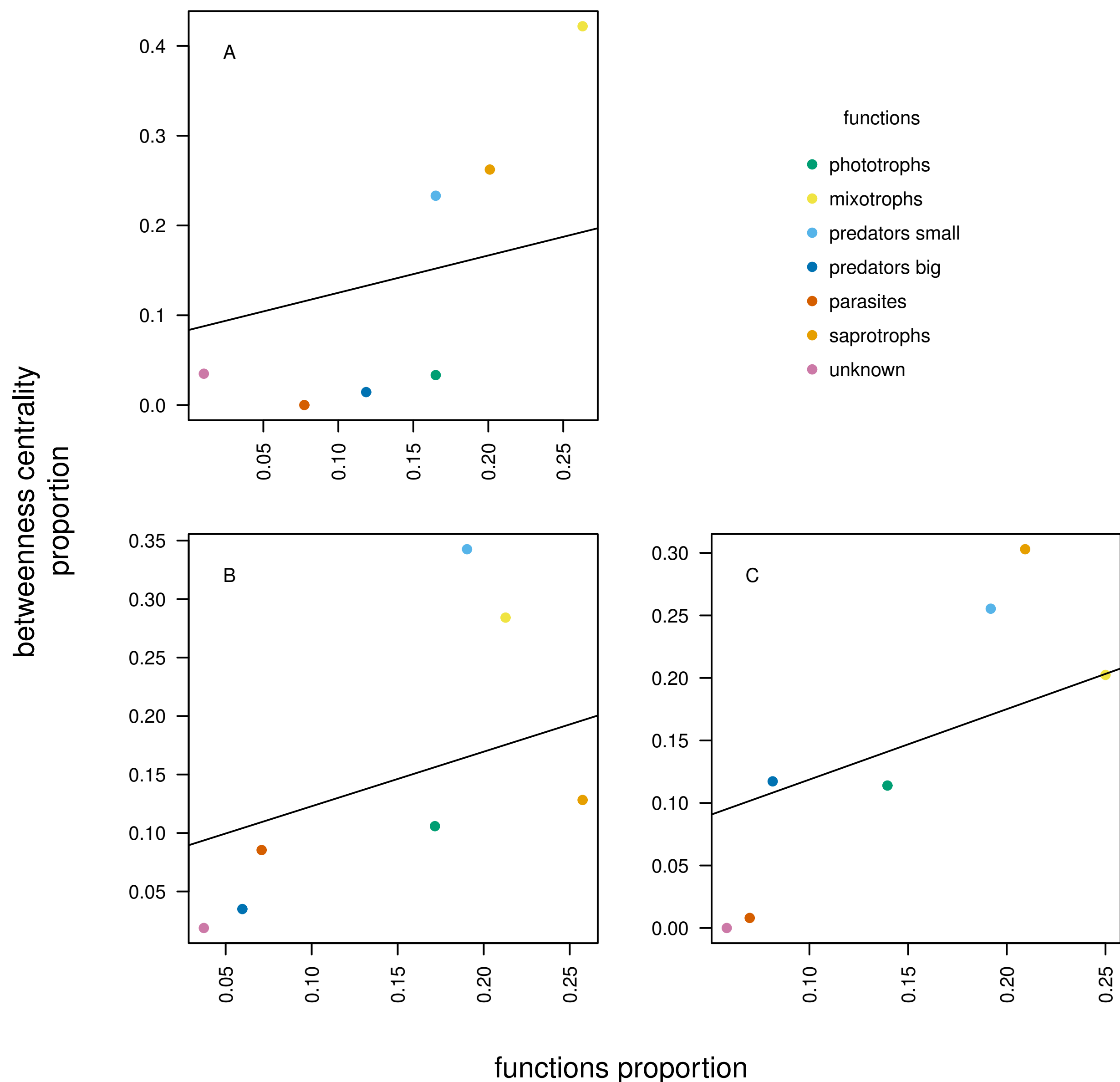


Figure 2.5: Proportion of shortest path between two OTUs (betweenness centralities: BC) calculated on co-occurrence network built from OTU communities retrieved from three micro-habitats (forests: A, hummocks: B, lawns: C) in peat bogs of the Jura mountains. The BC were calculated for each OTUs and summed according to the functional group for each micro-habitat. For each micro-habitat, the BC were plotted against the proportion of each functional group.

possibly came from the same organisms, which would explained their co-occurrence in the micro-habitat. Despite the obvious biological reason to see these two OTUs co-occur, no link respected the significance threshold for the three other micro-habitats despite a significant correlation in each case (Pearson coefficient $P < 0.05$) (Figure S2.6). This raises the question whether the significance threshold to assess a link between two OTUs might not be too stringent and produces a high rate of false negative.

General conclusions

Our results show negative relationship between diversity (highest in more productive environments) and network complexity. Such results are possible only if many organisms do not interact in nutrient rich environments, but create links in poor habitats. These links can be of different nature, from symbiosis to simple facilitation. Highly connected networks become then resistant to secondary species loss (Dunne et al., 2002). If communities are indeed bottom-up regulated, keystone eukaryotes depend directly on nutrients directly (like Fungi) or indirectly (like bacterivorous protists). These results suggest that an addition of nutrients like N or P would reduce the number of interactions, and hence the resistance of the system. Nutrients (or the lack of them) are, therefore, the Achille's heel of peatland ecosystems, and airborne N deposition may well be Pâris' arrow.

Acknowledgements

This study was partly funded by Swiss National Science Foundation projects no. 310003A 143960 to EL.

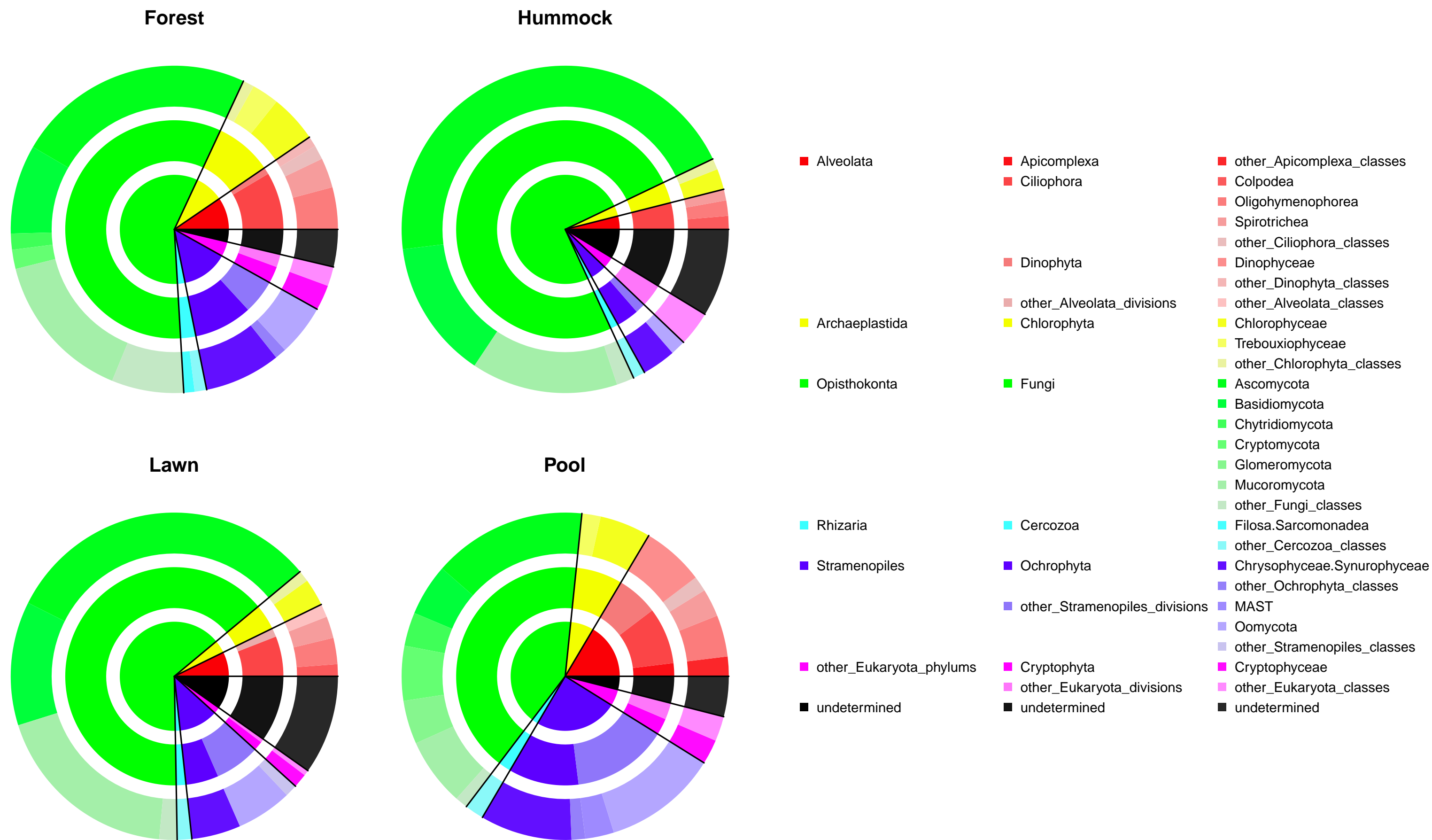


Figure S2.1: Relative abundance of taxon sequence abundances from four micro-habitats of peat bogs located in the Jura Mountains. Only taxa representing at least 1% of the total abundance for a micro-habitat are represented.

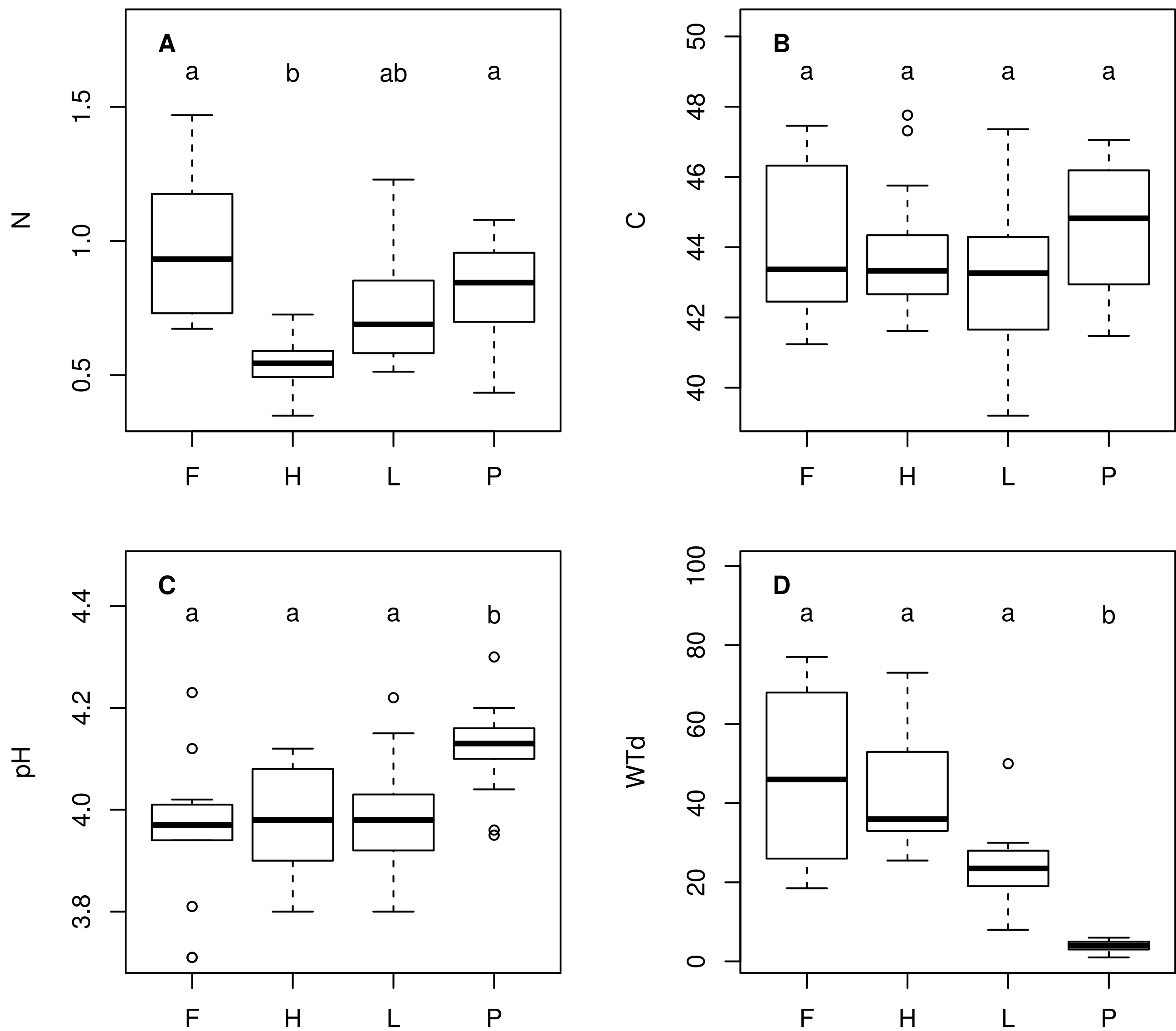


Figure S2.2: Environmental variables (total nitrogen percentage: A, total carbon percentage: B, pH: C, and water table depth: D) retrieved from four micro-habitats (forests: F, hummocks: H, lawns: L, pools:P) in peat bogs of the Jura mountains. The letters above the boxplots indicate the significantly different distributions according to a multiple comparisons of mean rank sums test (Nemenyi test) to the threshold of 0.05.

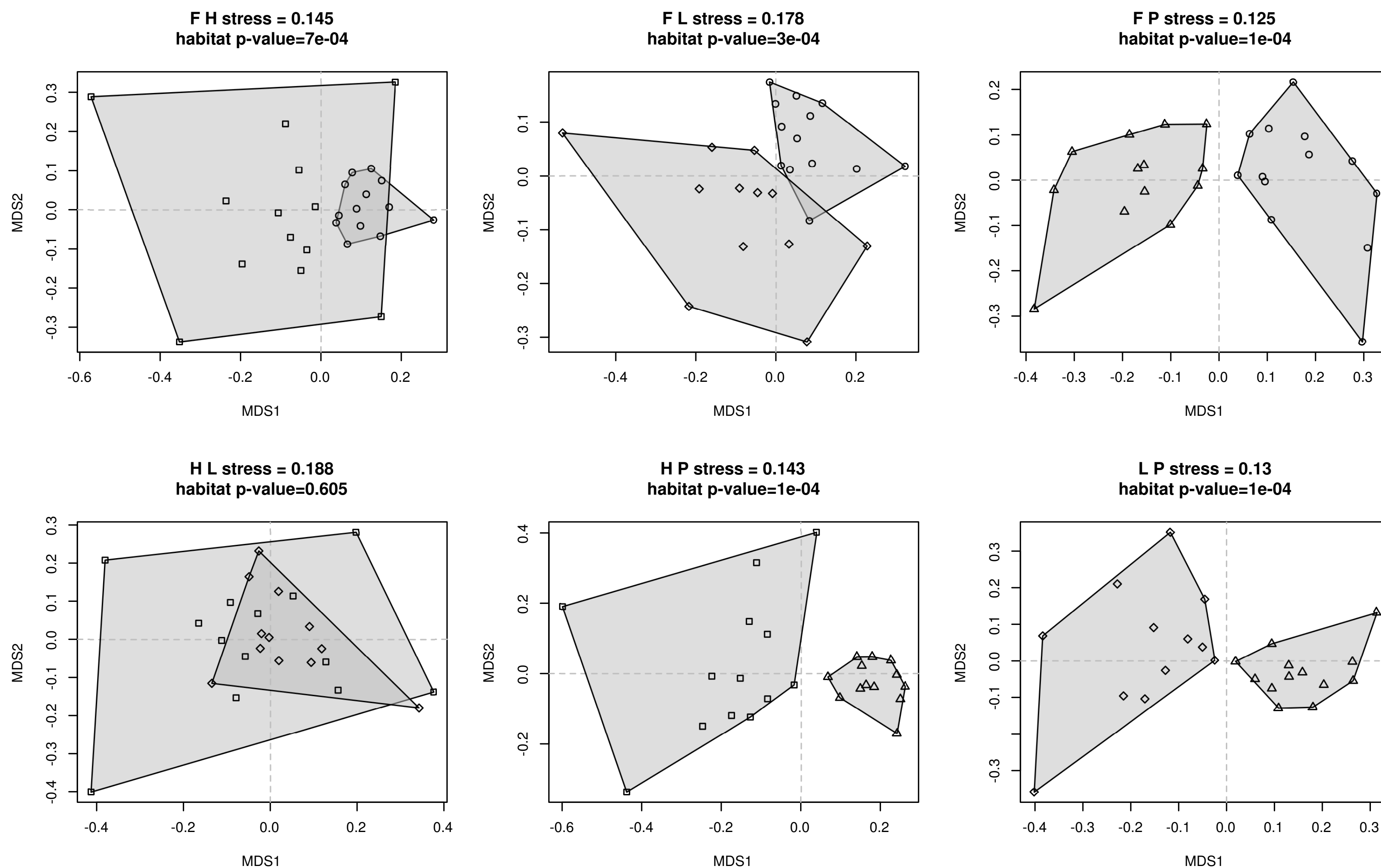


Figure S2.3: Non-metric multidimentional scaling (NMDSs) calculated on eukaryotic OTU communities from *Sphagnum* sp. taken from four micro-habitats (forests: rounds, hummocks: squares, lawns: diamonds, pools: triangles) in five peat bogs in Switzerland and France. Each sub-figure represents a pair of micro-habitats which are labelled on the top of each graphic (forests: F, hummocks: H, lawns: L, pools: P). The significance of the habitat effect on community is also shown on the top of each graphic. The peat bog effect on the communities is not significant for any pairs of micro-habitat.

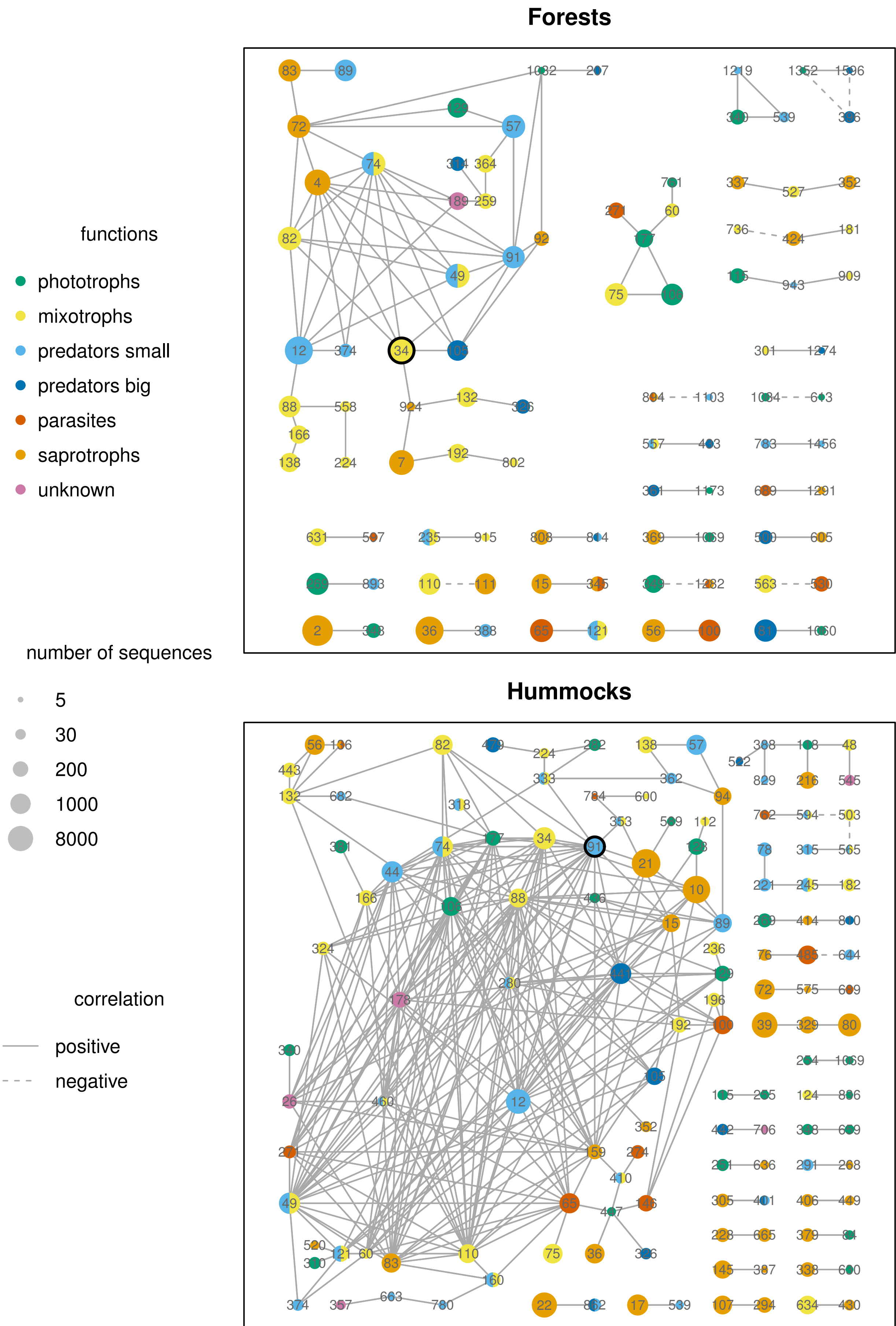


Figure S2.4: Co-occurrence networks of OTUs retrieved from four micro-habitats in peat bogs of the Jura mountains. Each point represents an OTU, the different colors represent the functional group associated with it and the size the number of sequences found in the micro-habitat. Each link represents a significant positive (plain line) or negative (dashed lines) correlation between two OTUs after adjustment according to the False Discovery Rate method. The OTUs surrounded in black are the central OTUs for each micro-habitat.

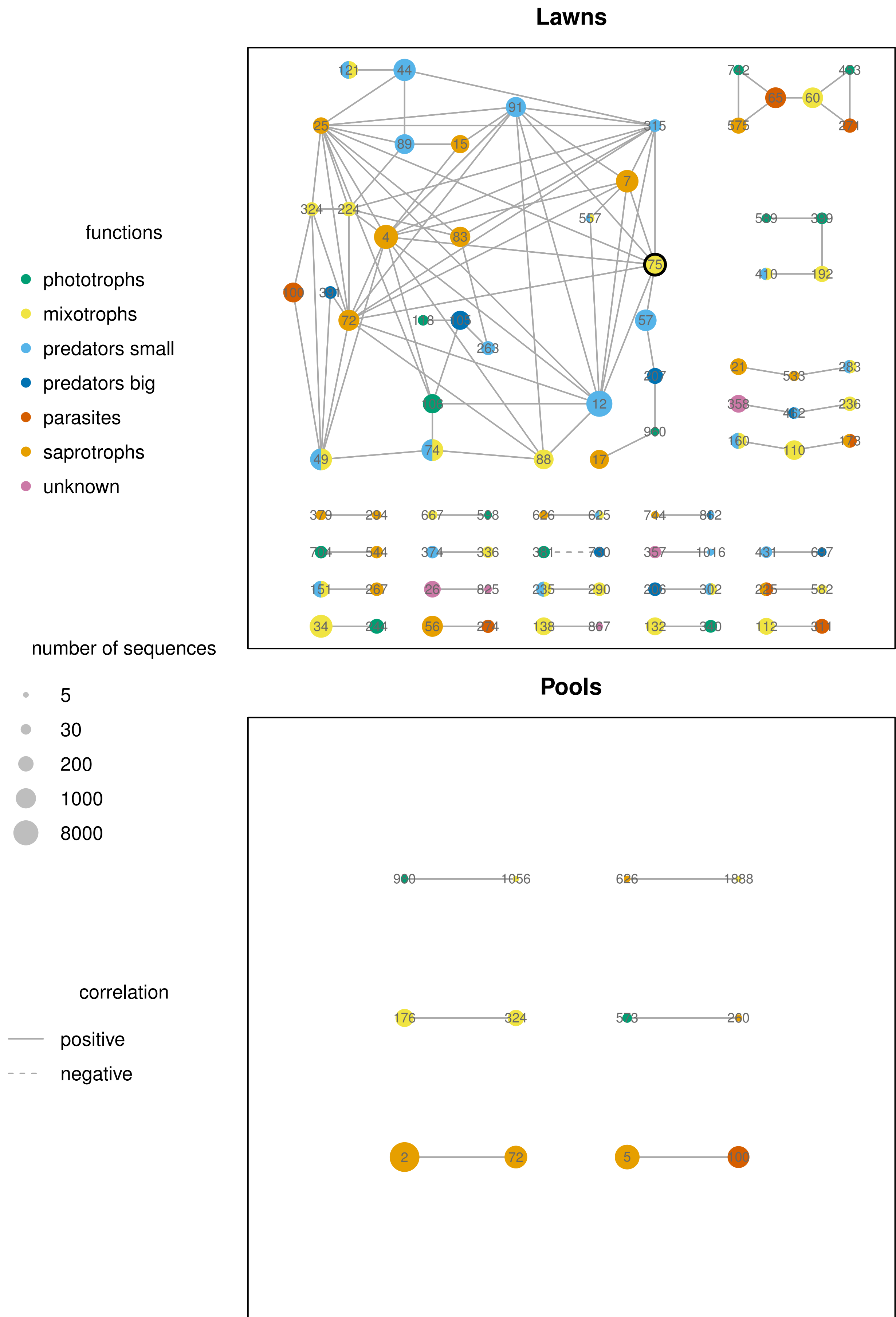


Figure S2.4: continuation

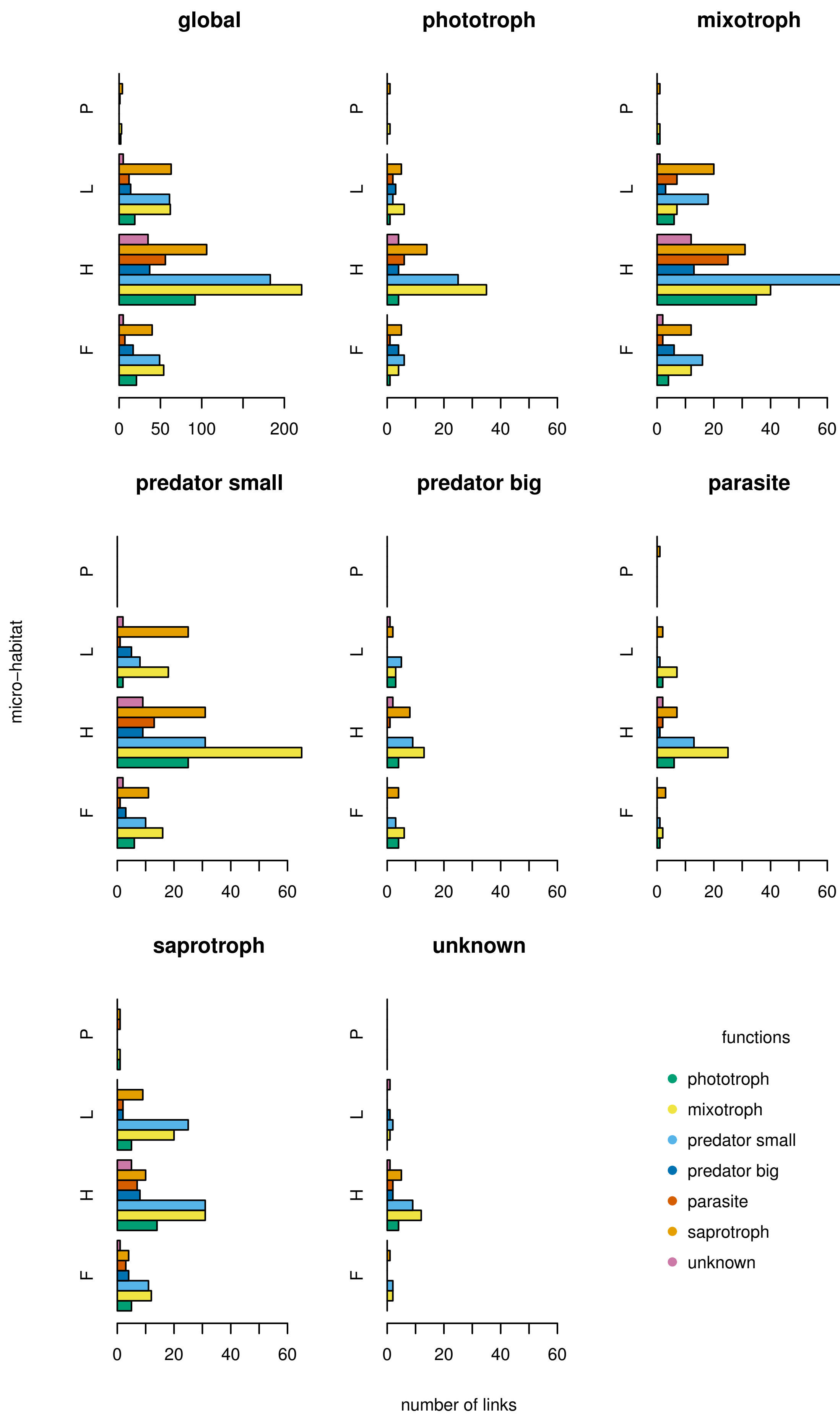


Figure S2.5: Number of first degree links (FDL) formed in co-occurrence networks calculated from micro-eukaryotic OTU communities in peat bogs located in the Jura mountains (global). The number of FDL is also summed according to the functional group and micro-habitat for the overall links and for each functional groups individually.

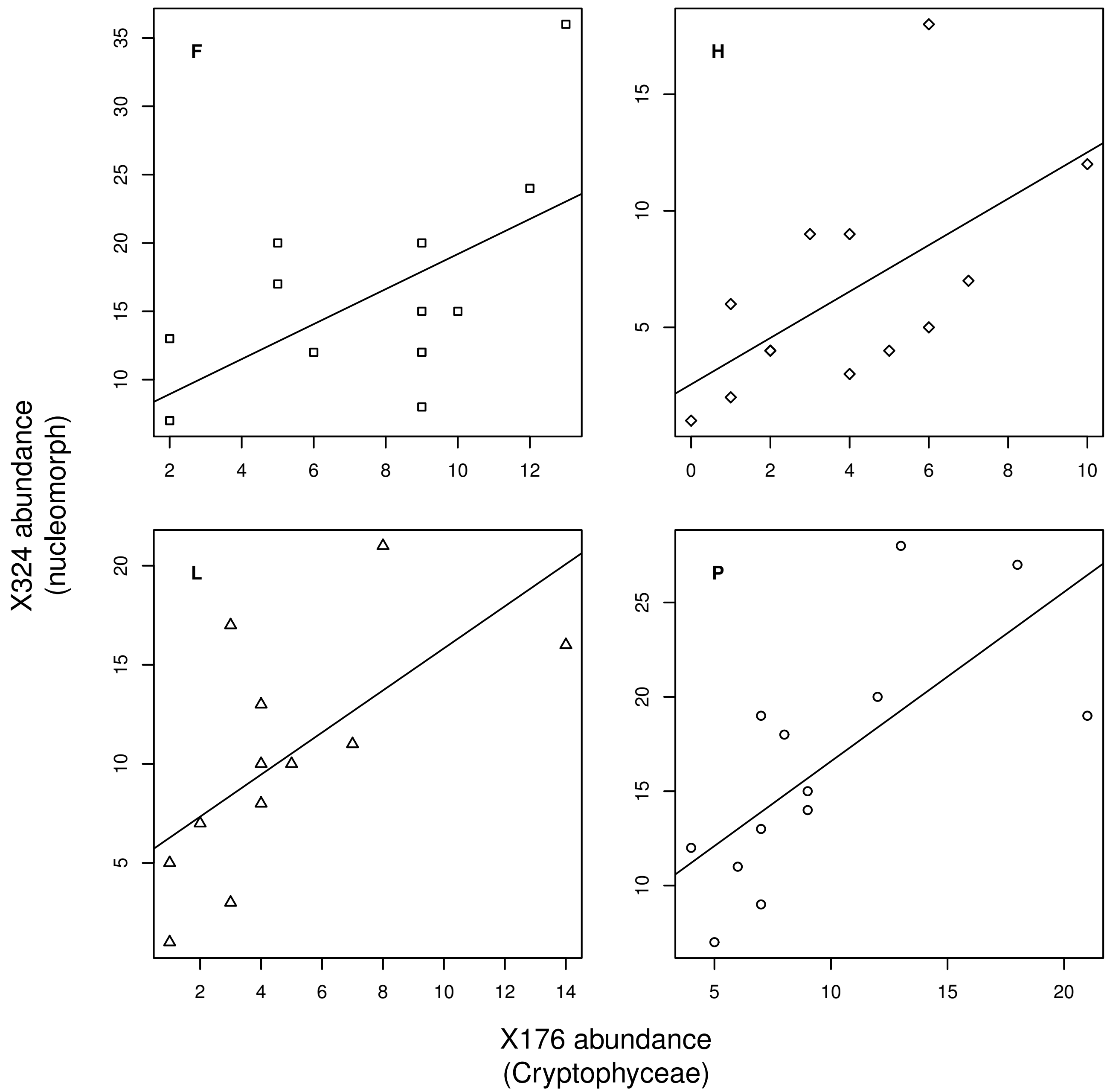


Figure S2.6: Abundances of two OTUs assigned to a Cryptophyceae (X176) and probably its nucleomorph (X324). In each micro-habitat the two OTUs correlate significantly (Pearson $P < 0.05$).

Chapter 3

Spatial patterns of soil protist diversity are better predicted from topo-climatic than from edaphic variables in the Swiss western Alps

Christophe V.W. Seppey[1*], Olivier Broennimann[2,3], Aline Buri[3], Erika Yashiro[2,4], Eric Pinto[2,4], David Singer[1], Quentin Blandenier[1], Edward A.D. Mitchell[1,5], Hélène Niculita Hirzel[6], Enrique Lara[1+], Antoine Guisan[2,3,+]

[1]Laboratory of Soil Biodiversity, University of Neuchâtel, Rue Émile-Argand 11, 2000 Neuchâtel, Switzerland
[2]Department of Ecology and Evolution, University of Lausanne, Biophore, 1015 Lausanne, Switzerland
[3]Institute of Earth Surface Dynamics, University of Lausanne, Géopolis, 1015 Lausanne, Switzerland
[4]Department of Fundamental Microbiology, University of Lausanne, Biophore, 1015 Lausanne, Switzerland
[5]Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland
[6]Institute for Work and Heath, Universities of Lausanne and Geneva, Route de la Corniche 2, 1066 Epalinge-Lausanne, Switzerland
[*]**Corresponding author:** christophe.seppey@unine.ch
[+]**Co-last authors**

Abstract: Spatial patterns of plant, animal and soil microbial diversity are generally well predicted from topo-climatic variables (e.g. temperature, slope aspect), and, for plants and soil microbes, by edaphic factors (e.g. pH, C/N). However, knowledge about patterns and drivers of soil protist diversity is more limited. Given the huge diversity of soil protists, spatial patterns and their drivers may vary among taxonomic and functional groups (e.g. primary producers, saprotrophs, parasites). Metabarcoding of soil DNA now allows efficiently assessing the diversity of soil protists, its spatial patterns and drivers in comparison with other groups of organisms.

We compared the explicative and predictive power of topo-climatic vs. edaphic variables for soil protist diversity using 178 soil samples from an altitude-stratified random sampling in the Swiss western Alps. We defined protist Operational Taxonomic Unit (OTU) communities by metabarcoding of the V4 region of the ribosomal RNA small sub-unit gene. We assessed and modelled the diversity (Shannon index) patterns of the full protist community and of nine clades belonging to three functional groups: parasites (Apicomplexa, Oomycota, Phytomyxea), phagotrophs (Sarcomonadea, Tubulinea, Spirotrichea) and phototrophs (Chlorophyta, Trebouxiophyceae, Bacillariophyta) as a function of topo-climatic and/or edaphic variables using Generalized Additive Models.

OTU diversity was dominated by phagotrophs (Sarcomonadea & Tubulinea) and parasites (Apicomplexa). Edaphic variables explained a higher proportion of variance in diversity than topo-climatic variables for most groups but the largest part of variance was explained jointly by both sets of variables. The significance of topo-climatic variables varied among taxonomic and - to a certain extent - functional groups: while most variables significantly explained the diversity of phototrophs, this was not the case for parasites. Topo-climatic variables had a slightly better predictive power than edaphic variables but predictive power varied most among taxonomic groups.

Our results show that soil protist diversity can be predicted to various degrees from both edaphic and topo-climatic variables. Contrasts among protist taxonomic and functional groups are as large as between plants and animals, in line with the broad range of functional roles and likely environmental drivers within protists. Such spatial models can be used to predict hotspots of diversity or pathogens infections.

Keywords: Spacial modelling . predictive modelling . protist diversity . topo-climatic variables . edaphic variables

3.1 Introduction

Protists are acknowledged as a hyper-diverse group of organisms in soils (Mahé et al., 2017) and known to respond to edaphic conditions such as pH, nutrient and moisture gradients as well as pesticides and other perturbations (Foissner, 1997). These organisms fulfil many essential roles in this ecosystem as primary producers, saprotrophs, predators, or symbionts (both mutualist and parasite) (Adl and Gupta, 2006). Interactions range from specialized, for instance between parasitic Apicomplexa and their animal hosts (Votýpka et al., 2017), to more generalized, as for Phytomyxea clades which can infect hosts from different eukaryotic kingdom (Neuhauser et al., 2014). Predatory protists occupy different levels of the microbial food web, as primary consumers of algae (cyanobacteria or eukaryotic), fungi and bacteria (Bonkowski and Clarholm, 2012; Dumack et al., 2016b; Hess and Melkonian, 2014), to higher levels like other phagotrophic protists or even micro-Metazoa (e.g. nematodes) (Geisen et al., 2015b; Gilbert et al., 2000). The role of protists as primary producers (e.g. Bacillariophyta, Chlorophyceae) is well recognised in aquatic ecosystems where they are the main contributors to photosyhthesis and carbon sequestration (Falkowski, 2002; Nelson et al., 1995). However photosynthetic protists also play essential roles in terrestrial ecosystems as essential components of cryptogamic crusts (e.g. Trebouxiophyceae) (Elbert et al., 2012; Pushkareva et al., 2016) and a still poorly documented, but probably significant source of organic carbon for soil organisms (Schmidt et al., 2016; Seppey et al., 2017). This diversity has never been as close to be fully revealed than with high-throughput sequencing (de Vargas et al., 2015; Mahé et al., 2017). These data give a valuable proxy about microbial communities and their functions, and can be used to calculate ecological metrics that allow to compare microbes to other better-known groups of organisms.

The effect edaphic characteristics (e.g. pH, texture, humidity) have on protist taxonomic and functional diversities is largely recognized (Dupont et al., 2016; Mills and Adl, 2006; Foissner, 1999b). If these variables give essential information on the conditions and mechanisms shaping the protists communities, they are difficult to integrate in spatial modelling at broad scales in order to predict the diversity, or any other metrics calculated from the community matrix. For a spatial prediction of the diversity, all environmental parameters used to predict the ecological metric are needed on every point of the map, what is hardly feasible if the measurements need to be taken directly in the soil. On the other hand, topo-climatic variables like slope steepness or air temperature can be easily measured on large scale areas using remote sensing methods like satellite data. These last variables already did their proofs in spatial modelling of macro-organisms, mostly plants (Guisan and Zimmermann, 2000). However, much less examples of spatial modelling of micro-organisms are found, and when found it is mostly in aquatic environments (Mitchell et al., 2000; Langer et al., 2013; Bulit, 2014; Zaric et al., 2006; Fraile et al., 2008). The development of such tools would be of primary importance to predict a zone at risk of pathogenic micro-eukaryotes infection or biodiversity hotspots.

We aimed to model protist diversity, in general as well as specifically for nine broad taxa chosen within three functional groups, in an elevation gradients in the Swiss western Alps. Protist communities were assessed from 178 meadow soil samples by metabarcoding of the V4 regions of the small sub-unit rRNA. This study aimed to assess the extent of protist diversity in mountainous meadows and determine to what extent two sets of environmental variables (edaphic and topo-climatic) could explain and predict this diversity. Based on this, we extrapolated this diversity over the Swiss western Alps.

3.2 Materials and methods

Sampling

Meadow soils were sampled from 178 plots distributed across the Swiss western Alps. Sampling was performed from July 4th to September 1st 2015 according to a random sampling stratified on the altitude. The samples consisted from five soil cores (5 cm diameters X 5 cm deep) taken in the four corners and the centre of a 2 m^2 plot. The five cores, were then pooled in a sterile plastic bag and kept in an icebox until the DNA extraction and soil analyses. For more details, see Yashiro et al. (2016).

Edaphic variables

We selected eight edaphic variables that were measured directly in the field or on the soil samples. Soil temperature was measured on the field. The relative humidity (rh) was assessed by weighing the mass of the soil sample before and after drying at 105 °C during 2 days. Carbon organic matter was determined by loss on 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 was determined by colorimetric analysis after a mineralisation at 550 °C with $Mg(NO_3)_2$. C/N ratio was calculated with 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.

Topo-climatic variables

For each sample location, seven topo-climatic variables were retrieved from maps of 25 square meter resolution. We used the growing degree day (gdd: threshold 0 °C), potential evapotranspiration (etp), topography (topo), slope southness (asp) and slope steepness (slp) (Zimmermann and Kienast, 1999; Zimmermann et al., 2007). In addition, we also calculated the summer

average temperature (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.

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 fragments were then amplified using the general eukaryotic primers TAReuk454FWD1 and TAReukREV3 (`CCAGCASCYGCGGTAATTCC` / `TYRATCAAGAACGAAAGT`; Stoeck et al., 2010). Both primers were tagged with an eight nucleotides fragment to allow the sorting of sequences according to the samples. PCR mix were constituted of 3 µL of 10 × diluted DNA extract, 0.4 µL of 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 triplicates with a PTC-200 Peltier Thermo Cycler (BioConcept, Allswill, Switzerland). DNA was then quantified to 20 ng for each replicate before to be pooled according to the replicate. A DNA library was prepared for the three replicate pools using the TruSeq Nano PCR-free Library Preparation kit and the sequencing made with an Illumina MiSeq by the University of Geneva (Molecular Systematics & Environmental Genomics Laboratory).

Bioinformatics pipeline

Good quality sequences were then selected on the basis of 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` v. 1.11.1, (Rognes et al., 2016) by comparing the environmental sequences 1) with each other for each replicate and 2) against the PR^2 database trimmed according the V4 primers (downloaded the 12 September 2016; Guillou et al., 2013). To reduce the noise caused by very rare sequences, we then removed every singleton. Replicates were then pooled according to their respective samples and OTUs were built with the program `swarm` v. 2.1.8 (Mahé et al., 2015). The dominant sequence of each OTU was then taxonomically assigned by aligning it to the trimmed PR^2 database using the global pairwise alignment program `ggsearch` v. 36.3.6 (Pearson, 2000).

We then removed every OTU not assigned to protists, namely Metazoan, Embryophyceae and Fungi. We also discarded OTUs with an obvious misidentification (Syndiniales, Phaeophyceae, Radiolaria, Rhodophyta and Opalinata) or a percentage of identity (PID) with the database below 65% as they were putative prokaryotes (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 coming from the same plot. In addition of the total community matrix, we selected nine broad taxonomic resolution taxa in three functional groups (parasites: Apicomplexa, Oomycota, Phytomyxea; phagotroph: Sarcomonadea, Tubulinea, Spirotrichea; phototroph: Chlorophyceae, Trebouxiophyceae, Bacillariophyta). For each of the taxonomic groups, we established a PID threshold verified manually on GenBank to discarded potential misidentification (Apicomplexa: 80%, Oomycota: 80%, Phytomyxea: 75%, Sarcomonadea: 80%, Tubulinea: 75%, Spirotrichea: 90%, Chlorophyceae: 90%, Trebouxiophyceae: 85%, Bacillariophyta: 77%).

Numerical analyses

We firstly discarded certain variables (gdd, etp, psum) to keep the correlation between two predictors below 0.7 to avoid collinearity issues (Zimmermann et al., 2007) (Figure S3.2). For each of the ten communities (total community plus nine broad taxa), OTU richness and Shannon diversity (H) were measured, and the differences between the distributions tested by a multiple comparisons of mean rank sums test (Nemenyi test; Hollander and Wolfe, 1999, `posthoc.kruskal.nemenyi.test` function, `PMCMR` package v. 4.1; Pohlert, 2014). For each of the ten communities, H was modelled in function of the uncorrelated variables using Generalized Additive Model (GAM). Variation partitioning and predictive power were calculated on 100 bootstraps composed of 80% of the 178 samples. The adjusted R^2 was used to estimate the variance explained and was calculated on three models; one for each set of variables (topo-climatic and edaphic) and a third on the overall variables. Variance explained by each set of variables was calculated as the difference between the variance explained by the overall model and the variance explained by the other set of variables. For each H and each set of variables, the predictive power was estimated as the Root Mean Square Error (RMSE) of the samples not included to build the model (20% of the samples). The effect of taxa and predictor nature on the RMSE was tested by a Nemenyi test. Finally, the diversities of the nine broad taxa and total community were extrapolated on the full area of the Swiss western Alps with a GAM using the topo-climatic variables.

3.3 Results

Diversity

We retrieved a total of 24'322'487 good quality sequences of which 97% were not chimaeric and 71% were not singletons. The 17'234'091 remaining sequences were

Table 3.1: Number of sequences and OTUs through each step of the analysis for the total community and for the nine broad taxa. The numbers between parenthesis represent the percentage of the total community.

	nb of sequences		nb of OTUs	
quality check	24'322'487			
chimera removal	23'724'876			
singleton removal	17'234'091			
clustering	17'110'114		41'048	
unwanted taxa removal	3'230'736		18'287	
samples pooling	2'692'209		17'969	
Apicomplexa	507'504	(13)	1'560	(6)
Oomycota	150'500	(5)	372	(2)
Phytomyxea	19'631	(<1)	128	(<1)
Sarcomonadea	486'693	(18)	3'520	(18)
Tubulinea	208'943	(5)	1'591	(5)
Spirotrichea	130'016	(5)	377	(2)
Chlorophyceae	130'550	(5)	249	(1)
Trebouxiophyceae	27'542	(<1)	180	(<1)
Bacillariophyceae	19'135	(<1)	300	(1)

clustered into 41'048 OTUs of which 18'287 were assigned to protists and 17'696 remained after pooling the samples (Table 3.1). The protists community was dominated (proportion of sequences) by Cercozoa, (principally Sarcomonadea and Thecofilosea), and Alveolata of which more than half were assigned to Apicomplexa (dominated by 94% of Gregarine sequences) and ca. 40% to Ciliophora (mostly from classes Spirotrichea, Oligohymenophorea, Litostomatea and Colpodea) (Figure S3.1). The three other dominant groups were the Stramenopiles (i.e. including Oomycota and Bacillaryophyta), Amoebozoa (including Tubulinea) and Archaeplastida (with Chlorophyceae and Trebouxiophyceae) (Figure S3.1).

The nine chosen taxa jointly contributed to over half (54%) of all retained sequences and represented over 35% of the total OTU richness (Table 3.1). The average richness per sample of these clades varies from 7 (Phytomyxea) to 249 (Sarcomonadea). Richness (100 bootstraps of the samples) was lowest for phototrophs and highest for phagotrophs (Figure 3.1). Shannon diversity indices followed the same trend, varying from an average value of 1.1 (Phytomyxea) to 4.3 (Sarcomonadea).

Explicative models

The percentage of variance of the total and broad taxa communities explained by the 12 environmental variables ranged from 6% (Oomycota) to 36% (Chlorophyceae) (Table 3.2). The environmental variables significantly ($P < 0.05$) correlated to protist diversity in the model with the two sets of variables were: slope steepness (5 taxa), pH (4 taxa), mean summer temperature (3 taxa), loss on ignition (2 taxa), shale percentage (1 taxon) and electric conductivity (1 taxon) (Table 3.2). None of the two sets of variables seemed to explain the diversity of protists more significantly. The two sets of variables seemed to explain the diversity of phototrophic protists best and parasites least.

The variation partitioning showed that, in general, topo-climatic variables explained less variance than the edaphic ones (except for Bacillariophyta) (Figure 3.2). Furthermore, for most taxa

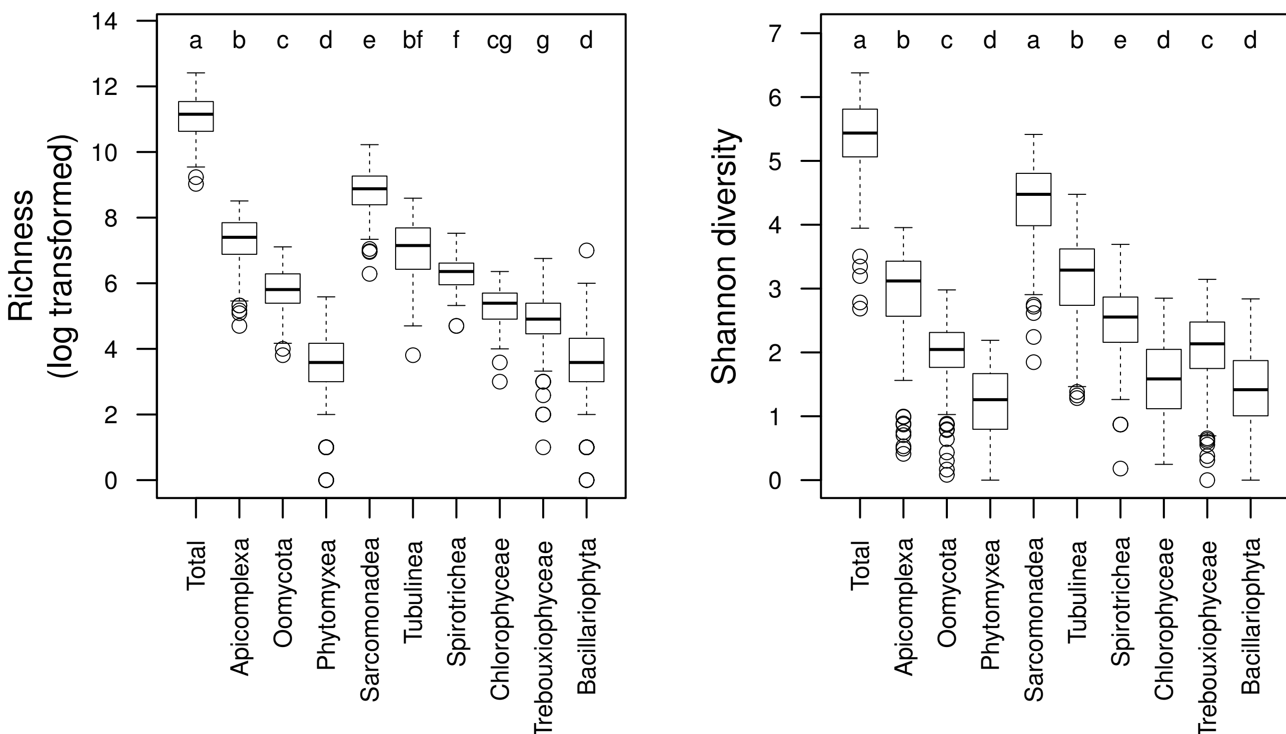


Figure 3.1: Shannon diversity and richness (log transformed) distributions of protist OTU communities retrieved from 187 plots in the Swiss western Alps. The distributions are shown for the total community as well as for nine coarse taxa. The letters above the boxplots represent groups according to a multiple comparison mean rank sums test (Nemenyi test $P < 0.05$).

Table 3.2: Significance of physico-chemical and topo-climatic predictors on the diversity modelled (Generalized Additive Model) from total micro-eukaryotic community and nine coarse taxonomic groups from OTUs gathered from Swiss western Alps meadow soils. 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 < 0.1, one sign: P < 0.05, two signs: P < 0.01, three signs: P < 0.001). 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 Figure S3.3.

	Edaphic								Topo-climatic				R^2
	Soil temp	rh	pH	EC	P	C/N	LOI	Shale	topos	asp	slp	tmean678	
Total		(-)	(+)								(-)		0.14
Apicomplexa								-+					0.21
Oomycota							-					(+)	0.06
Phytomyxea											-		0.12
Sarcomonadea			++							(+)	-		0.16
Tubulinea							+				--		0.14
Spirotrichea			+			(+)						+++	0.20
Chlorophyceae			-	--++	(-)							(+)	0.36
Trebouxiophyceae			++	(-+)							-	+	0.31
Bacillariophyta											--++	-	0.18

(except Apicomplexa, and Bacillariophyta), the variance explained jointly by topo-climatic and edaphic variables was in average higher than the variance explained by either group of variables alone.

Predictive models

Topo-climatic variables showed on average a better predictive power (RMSE) than edaphic variables for all taxa except the Chlorophyceae. However, this was only significant for the Spirotrichea (Figure 3.3). The RMSE varied in function of the taxonomic group and the diversity of certain taxa were significantly better predicted (e.g. Oomycota) than others (e.g. Apicomplexa).

3.4 Discussion

General patterns of micro-eukaryotic diversity in soils

The high proportion of Cercozoa sequences was in line with findings from other soil eukaryotic DNA surveys (Bates et al., 2013; Harder et al., 2016; Seppey et al., 2017). Cercozoa were already known to be major components of soil protist diversity before the advent of molecular methods (Adl and Gupta, 2006). Ciliates were also well-represented in accordance to these previous studies, with Spirotrichea being the most abundant, in line with other studies on soils (Lara et al., 2007a). Stramenopiles were the third best represented group, and were dominated by oomycetes that are common and diverse in temperate soil systems (Singer et al., 2016; Seppey et al., 2017) and contain many plant parasites, but also animal pathogens and free-living saprotrophic forms (Lara and Belbahri, 2011; Beakes et al., 2012). In contrast, this group is less abundant and diversified in neotropical forest soil ecosystems, where it comprises mostly animal parasites (Mahé et al., 2017).

Explicative power of topo-climatic and edaphic variables

The two groups that are better explained by the combination of topo-climatic and edaphic variables are the phototrophic Chlorophyceae and the Trebouxiophyceae. Relationships between diversity of Chlorophyceae and pH in soils (Lenarczyk, 2015) and other soil parameters (Gonzalez Garraza et al., 2011) have been found in vegetation studies i.e. based on morphological data. Accordingly, Trebouxiophyceae are often associated to lichens (Upreti et al., 2015), which are well-known as bioindicators and may have a narrow ecological tolerance with respect to abiotic parameters (van Herk, 2001; Cobanoglu, 2015).

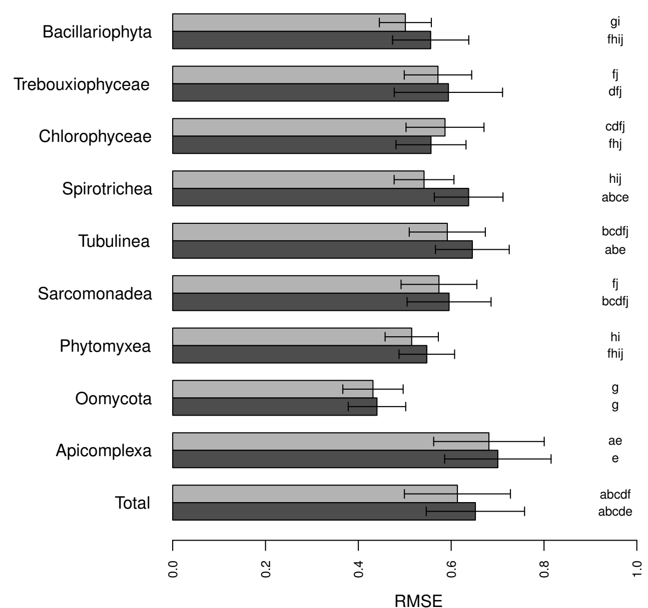


Figure 3.3: Average and standard deviation of Root Mean Square Error (RMSE) according to the predictor nature (micro / macro -environmental variables) for the total protists community and nine broad taxa. For each of the 20 situations the RMSE distributions were calculated from 100 cross validations of Generalized Additive Models performed with 20% of the samples as test dataset. The letters on the right of the boxplots represent groups according to a multiple comparison mean rank sums test (Nemenyi test P < 0.05).

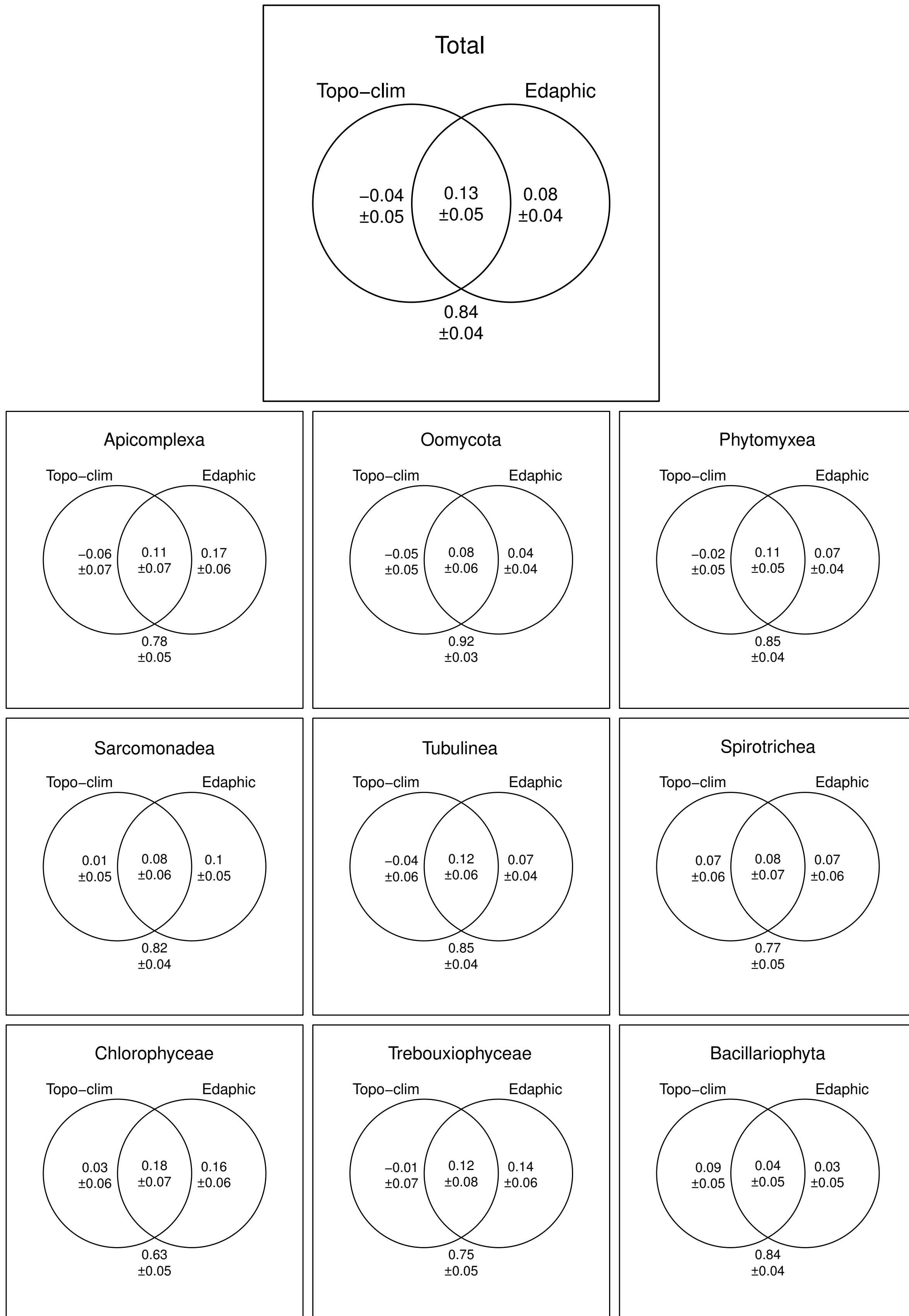


Figure 3.2: Variation partitioning of topo-climatic and edaphic variables on the diversity of protist OTUs from total communities and nine broad taxa retrieved from 178 meadow soils in the Swiss western Alps. The value represent the average and standard deviation (between parenthesis) of the adjusted R^2 retrieved from 100 GAM bootstraps calculated on 80% of the samples.

Slope steepness and pH were the two variables that were found most often to significantly explain protist diversity. Slope steepness affects drainage and leaching of nutrients. Thus, steeper slopes may represent more stressful habitats and this may explain the general trend for lower diversity in steeper slopes. Soil

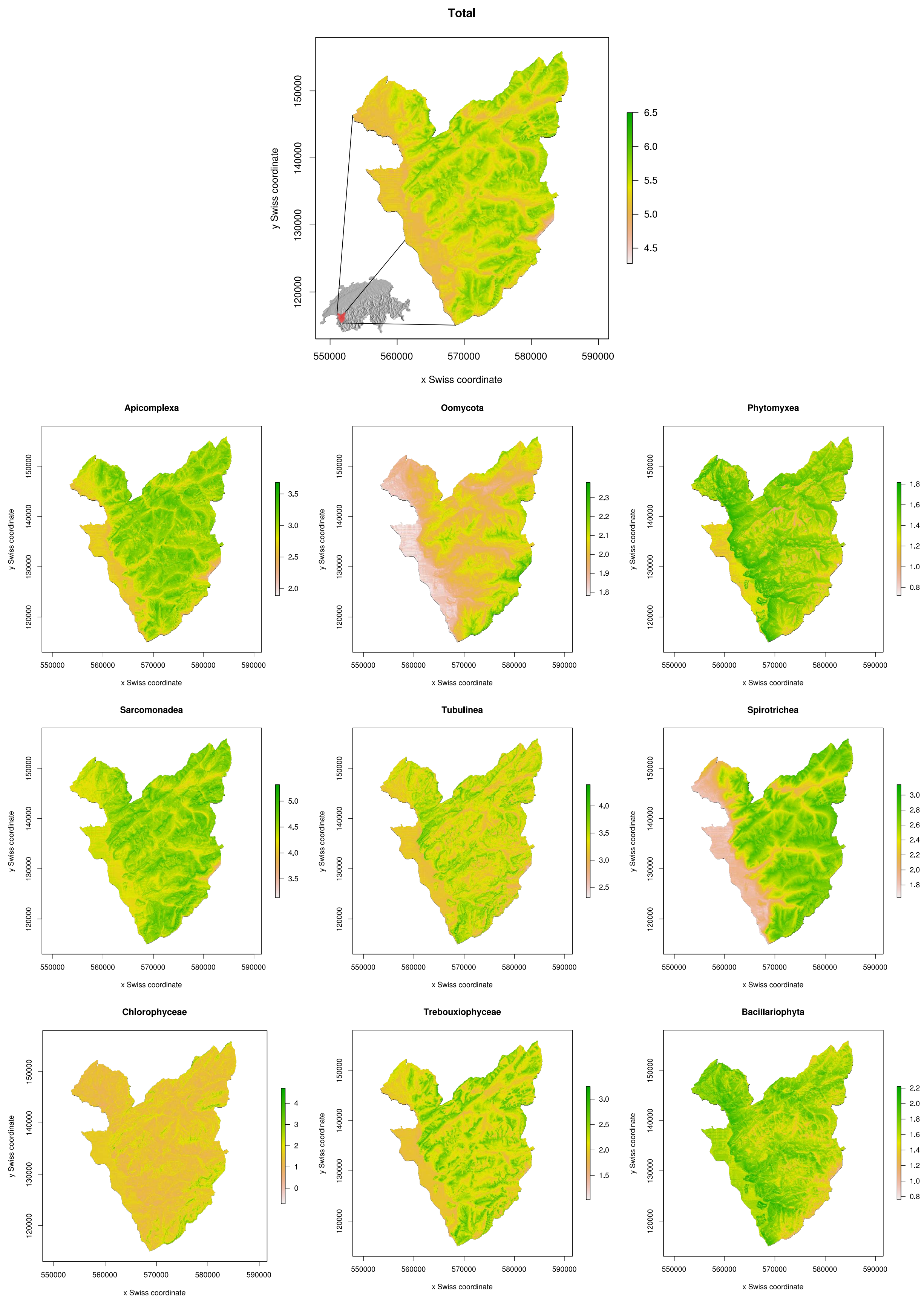


Figure 3.4: Diversity of the total protist community and nine coarse taxa predicted from Generalized Additive Model based on the topography, slope southness, slope steepness and average temperature from June to August.

pH is well known as a major driver of microbial diversity, including bacteria (Santoyo et al., 2017), fungi (Noyce et al., 2016; Zhang et al., 2016) and protists (Dupont et al., 2016). In our case, diversity values were often positively correlated with higher pH values, as most of the soils were acidic (i.e. most values between 4 and

6).

Among the two datasets (i.e. edaphic and topoclimatic) edaphic variables explained more variance than topoclimatic variables, but most of the variance was explained by both sets of variables. This result is in line with what was observed for euglyphid testate amoebae at a global scale, where edaphic variables such as the C/N ratio were more or less equally co-responsible with temperature to explain diversity distribution (Lara et al., 2016). Our results suggest that the use of both sets of explanatory variables may be extended to explain diversity patterns of many other groups of protists, and possibly to most of them.

Interpretation of the diversity spatial patterns modelled with topo-climatic variables

According to the diversity predictions, protists show clear spatial patterns (Figure 3.4), which seems to be positively driven by the summer temperature in most of the cases (Table S3.1), either in a positive (Apicomplexa, Bacillariophyta, Phytomyxea, Trebouxiophyceae and Tubulinea), in a unimodal (Sarcomonadea and Spirotrichea) or in a negative way (Chlorophyceae, Oomycota). Positive relationships correlated with altitudinal gradients are a typical pattern in macroecology, and this corresponds to the species-energy model (Fernández et al., 2016). Unimodal distributions correspond to an optimum in diversity which also as been reported in other studies. They can be interpreted as a compromise between two gradients, like for e.g. temperature and moisture (water energy model: (Fernández et al., 2016) others). Finally, Chlorophyceae and Oomycota are typically sensitive to high temperatures and dessication, the latter often using flagellated life stages for dispersal and do not encyst (Jeger and Pautasso, 2008). In addition, the high diversity of Chlorophyta in the lowest temperature zone (Figure 3.4) could be explained by the fact that micro-eukaryotic algae have a higher grow rate at low temperatures, favouring diversification in cold environments (Rose and Caron, 2007).

Conclusions

Our results suggest that topography and climate contribute in explaining protist diversity and complement the commonly used edaphic parameters very well. We also showed that the diversity of some taxa and - to a certain extent - functional groups, are better explained than others by the topo-climatic and edaphic conditions; particularly with phototrophic taxa when only topo-climatic variables are taken into account.

This approach could be applied at finer taxonomic levels to predict the distribution of individual species, notably in the case of invasive pests of economic importance such as certain oomycetes. The models could be improved by refining the taxonomic groups as taxa responding more homogeneously to the environment may show stronger correlation with abiotic variables than the broad groups we used. For instance, Oomycota contain organisms belonging to other functional groups than parasites (e.g. saprotroph; Beakes et al., 2012; Lara and Belbahri, 2011) or targeting a wide range of hosts (e.g. Phytophothora cinnamomi; Hardham, 2005). These improvements would pave the way to reliable extrapolation of protists diversity and provide useful tools to identify likely biodiversity hotspots or predict spatially the risk of pathogen infection.

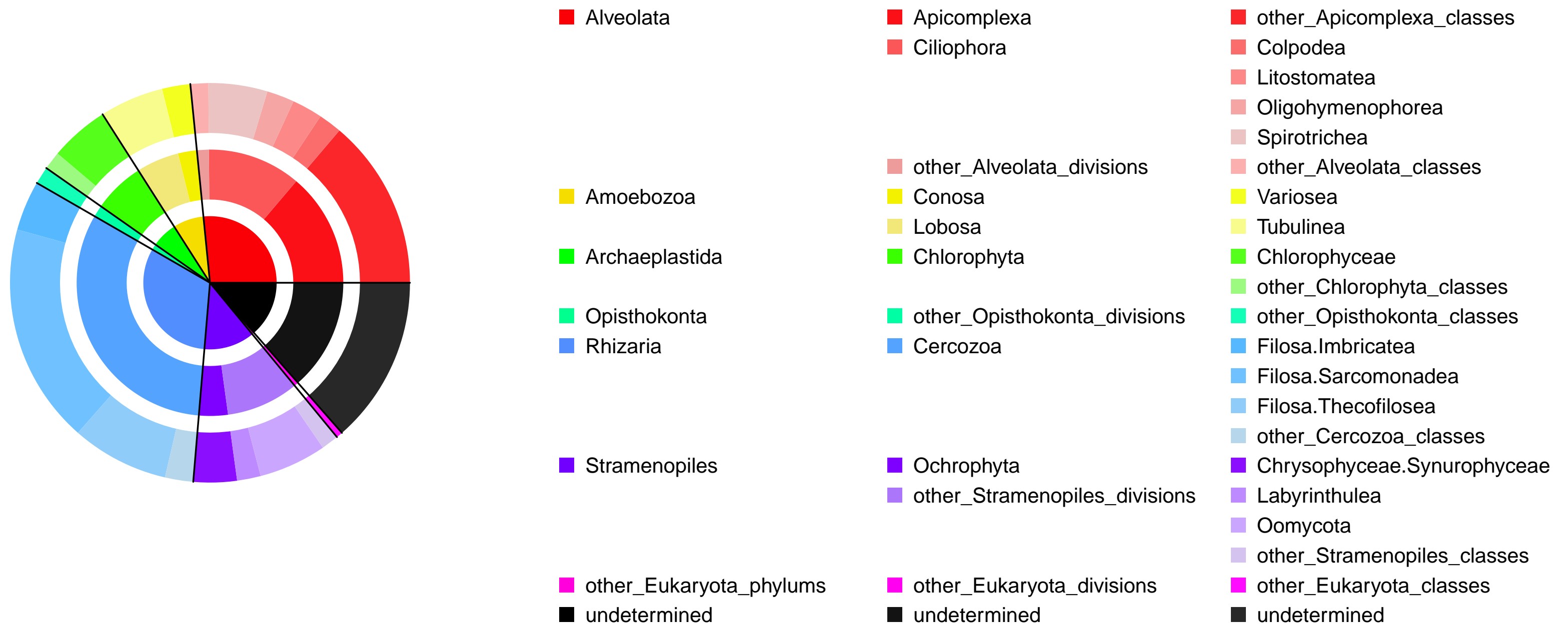


Figure S3.1: Relative abundance of taxon sequence abundances from 178 meadow soils from the Swiss western Alps. Only taxa representing at least 1% of the total abundance are represented.

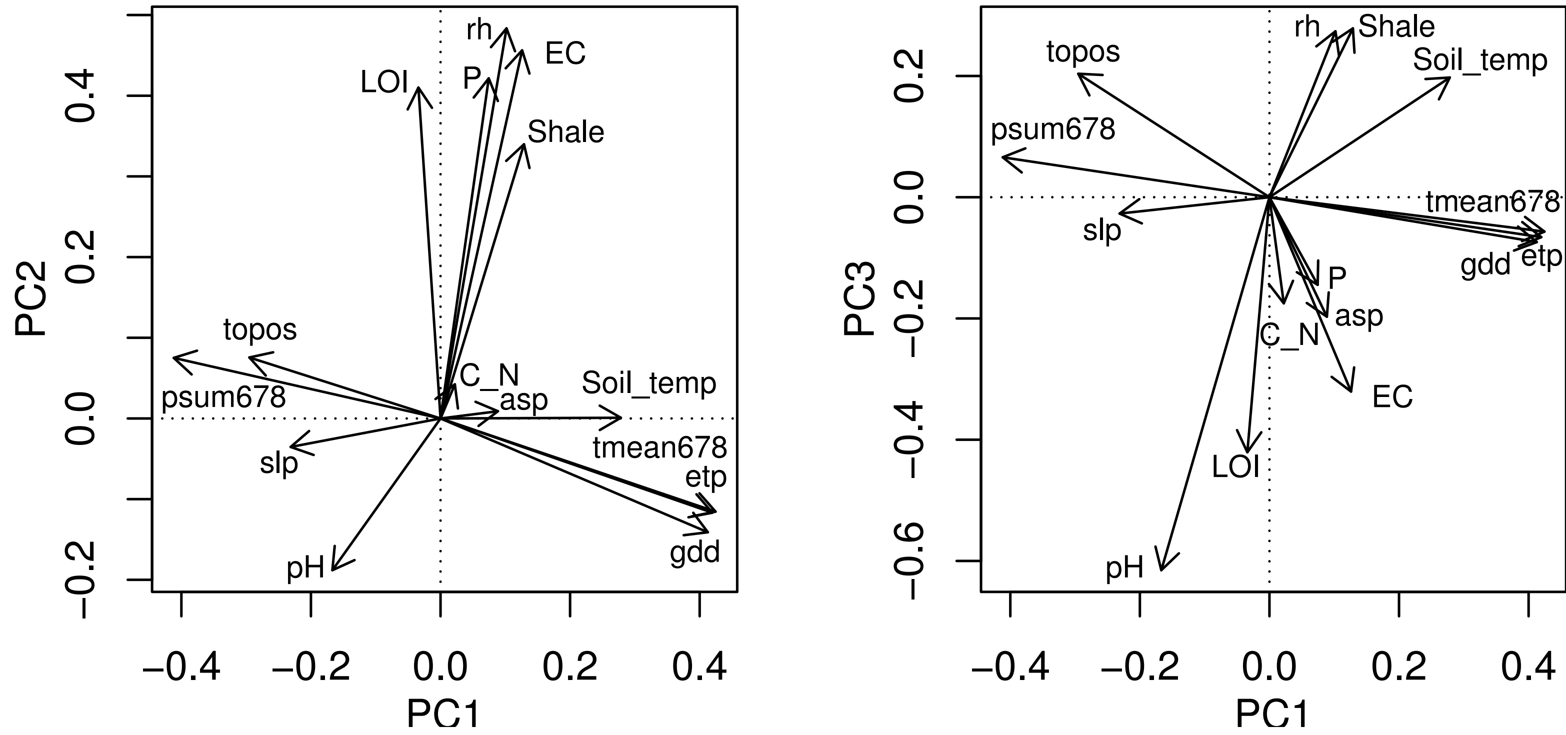


Figure S3.2: Principal component analysis calculated on the environmental variables (soil temperature: Soil_temp, electro-conductivity: EC, relative humidity: rh, loss on ignition: LOI, percentage of shale: Shale, pH, phosphorus percentage: P, C/N ratio: C_N, growing degree day: gdd, potential evapotranspiration: etp, topography: topo, slope southness: asp, slope steepness: slp, summer average temperature: tmean678, summer precipitation sum: psum678) measured from 178 plots through the Swiss western Alps. The first and second, and first and third axis are shown. Note the tight correlation between the tmean678 and the variables removed for the final analyses, namely gdd, etp and psum678.

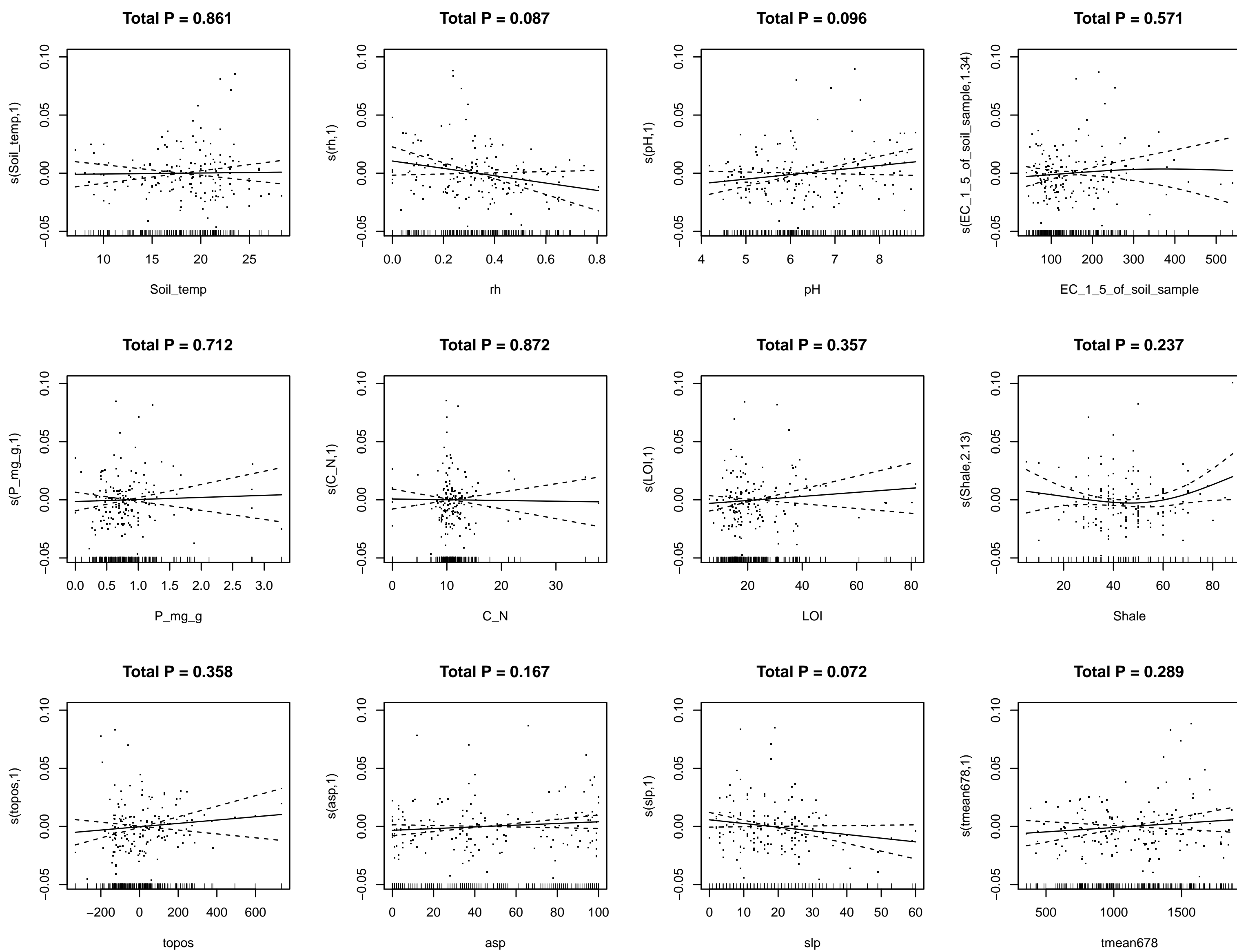


Figure S3.3: Diversity of the total protist community (Total) and all nine broad taxa in function of four topo-climatic predictors (topography, slope southness, slope steepness, average temperature from June to September) through Generalized Additive Models.

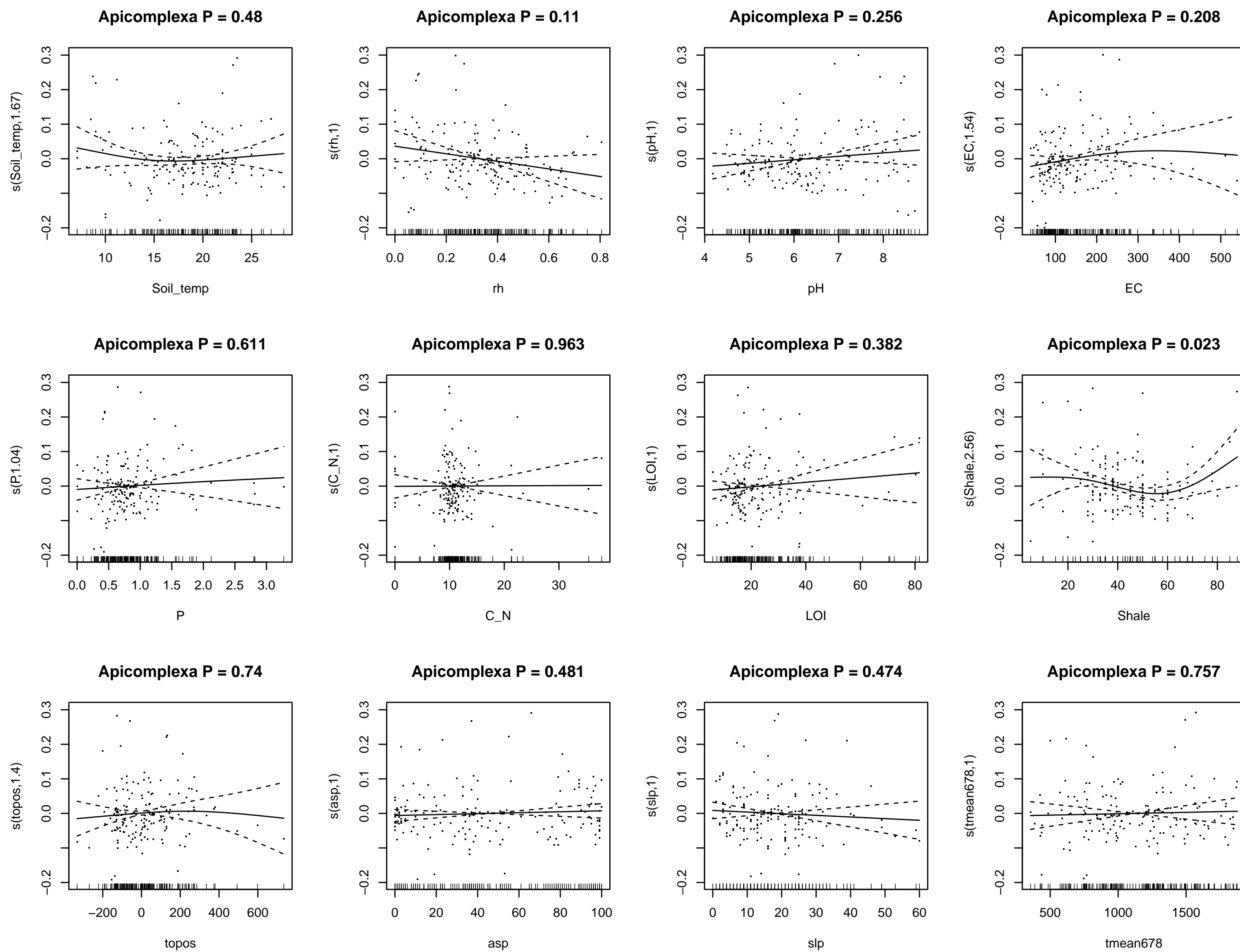


Figure S3.3: continuation

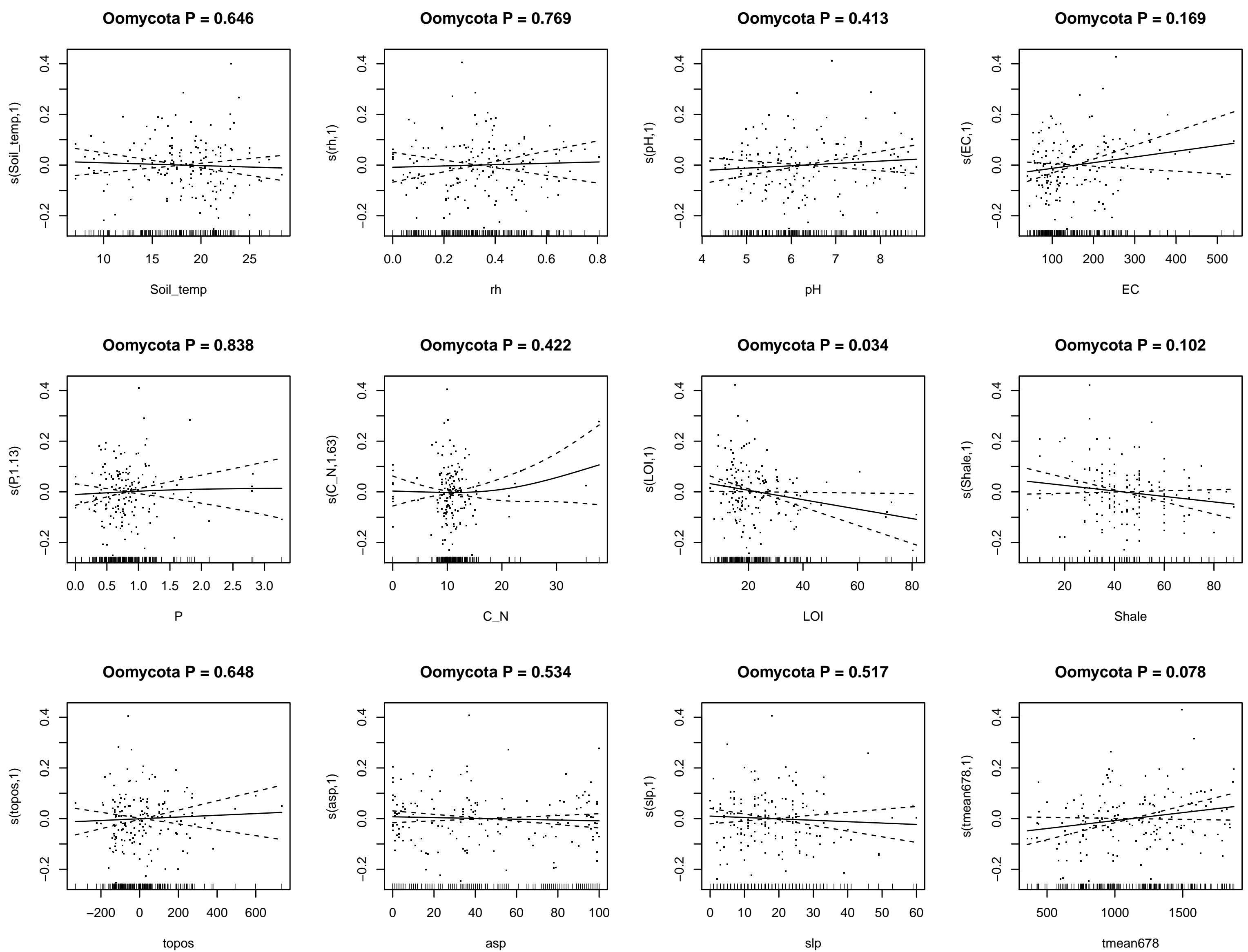


Figure S3.3: continuation

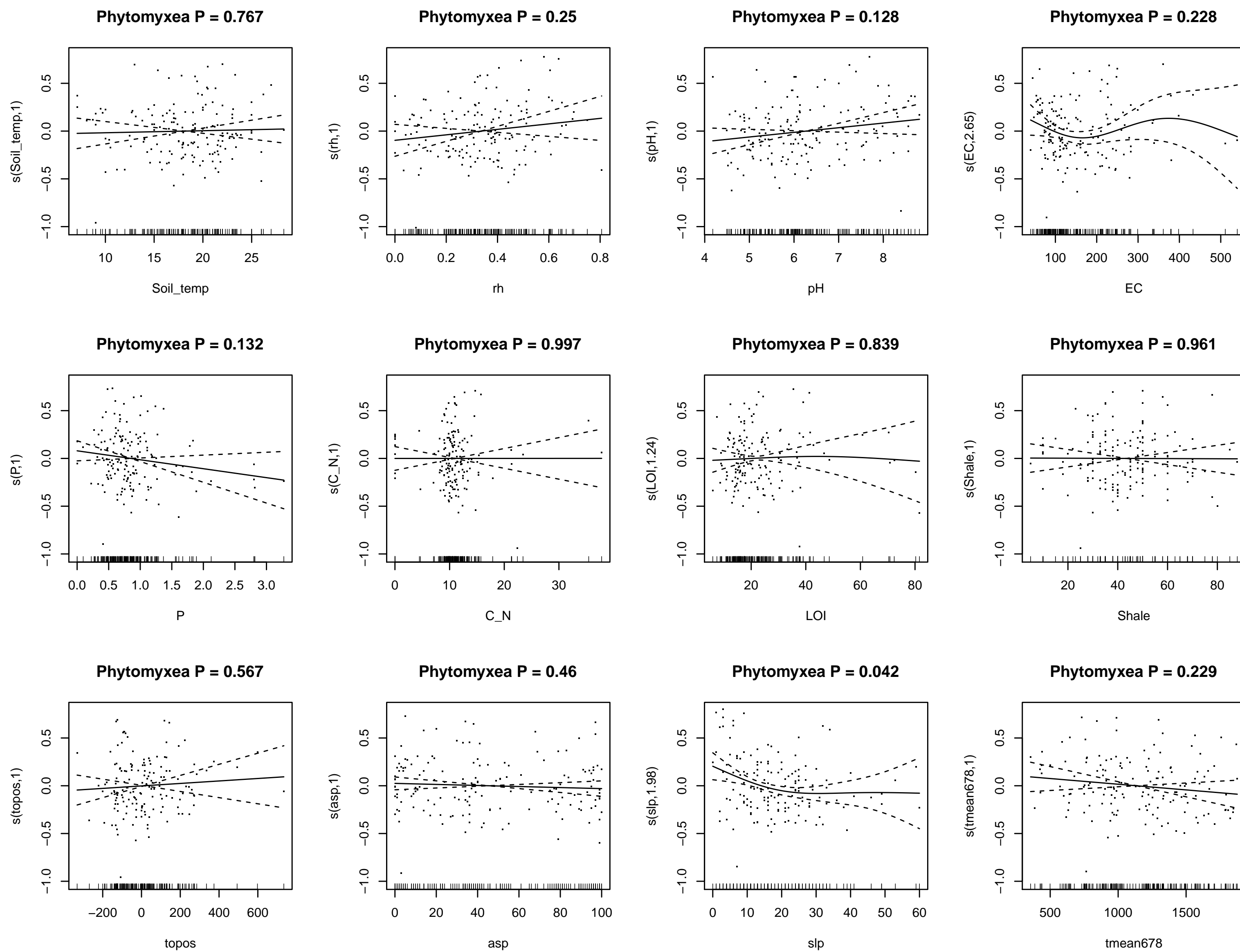


Figure S3.3: continuation

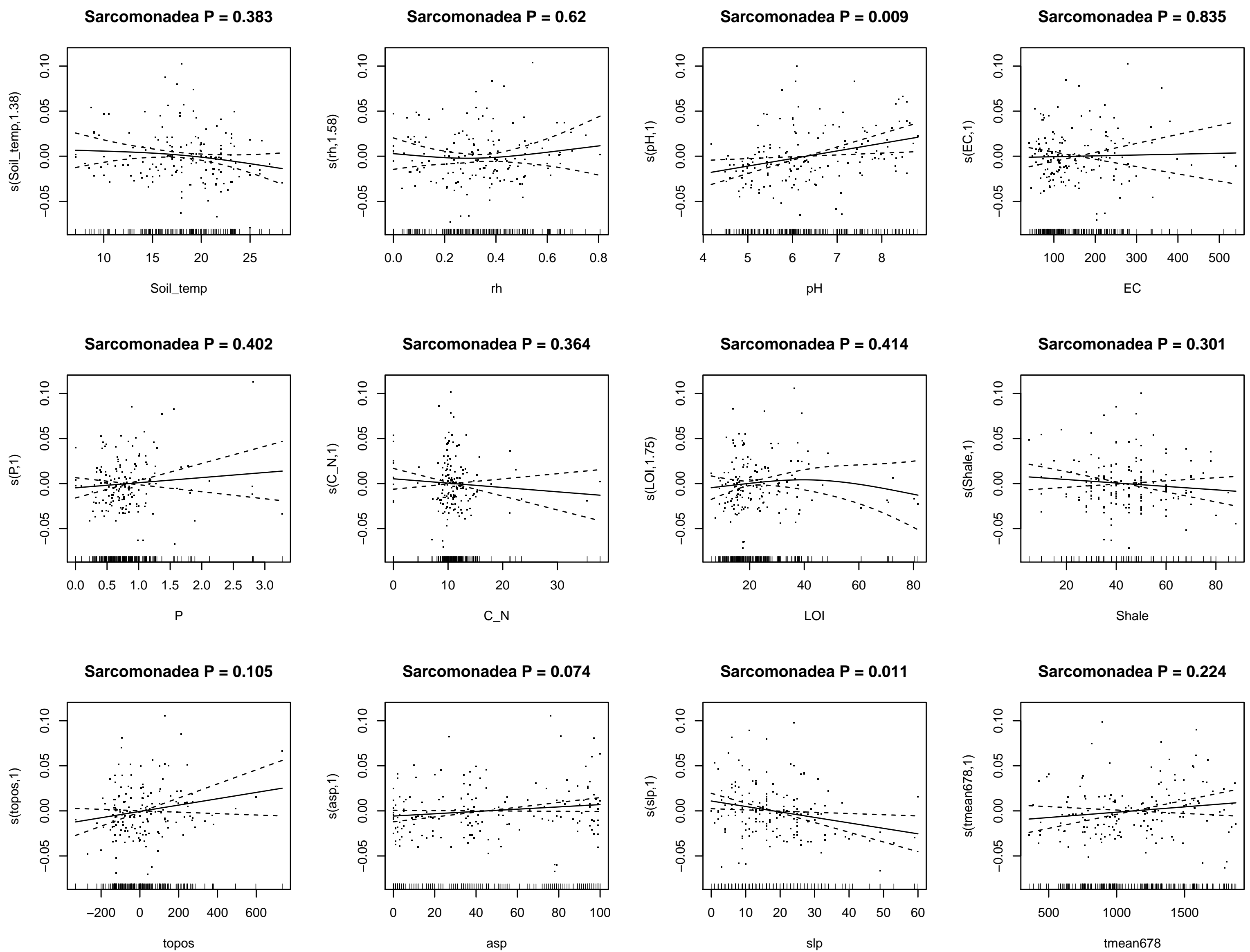


Figure S3.3: continuation

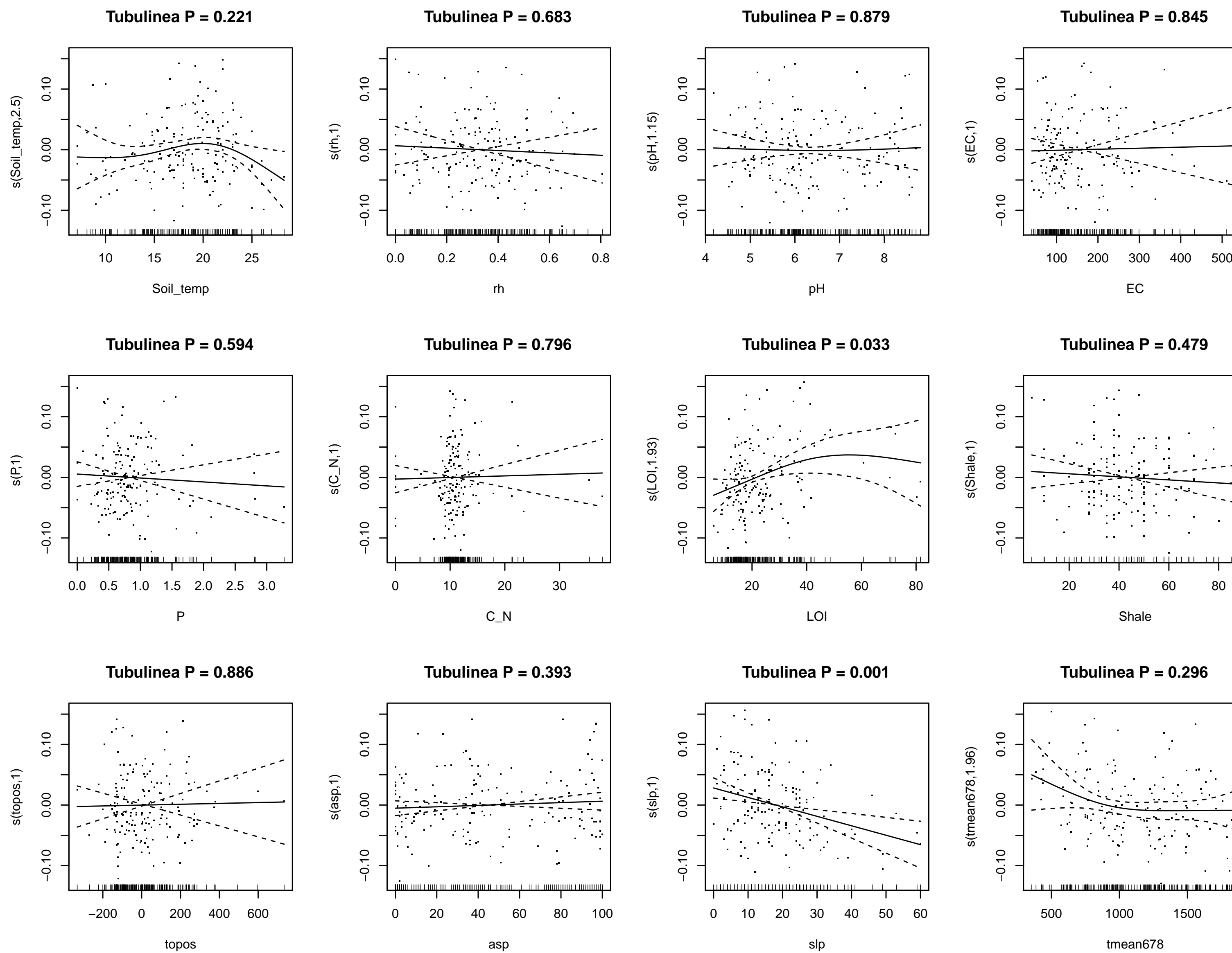


Figure S3.3: continuation

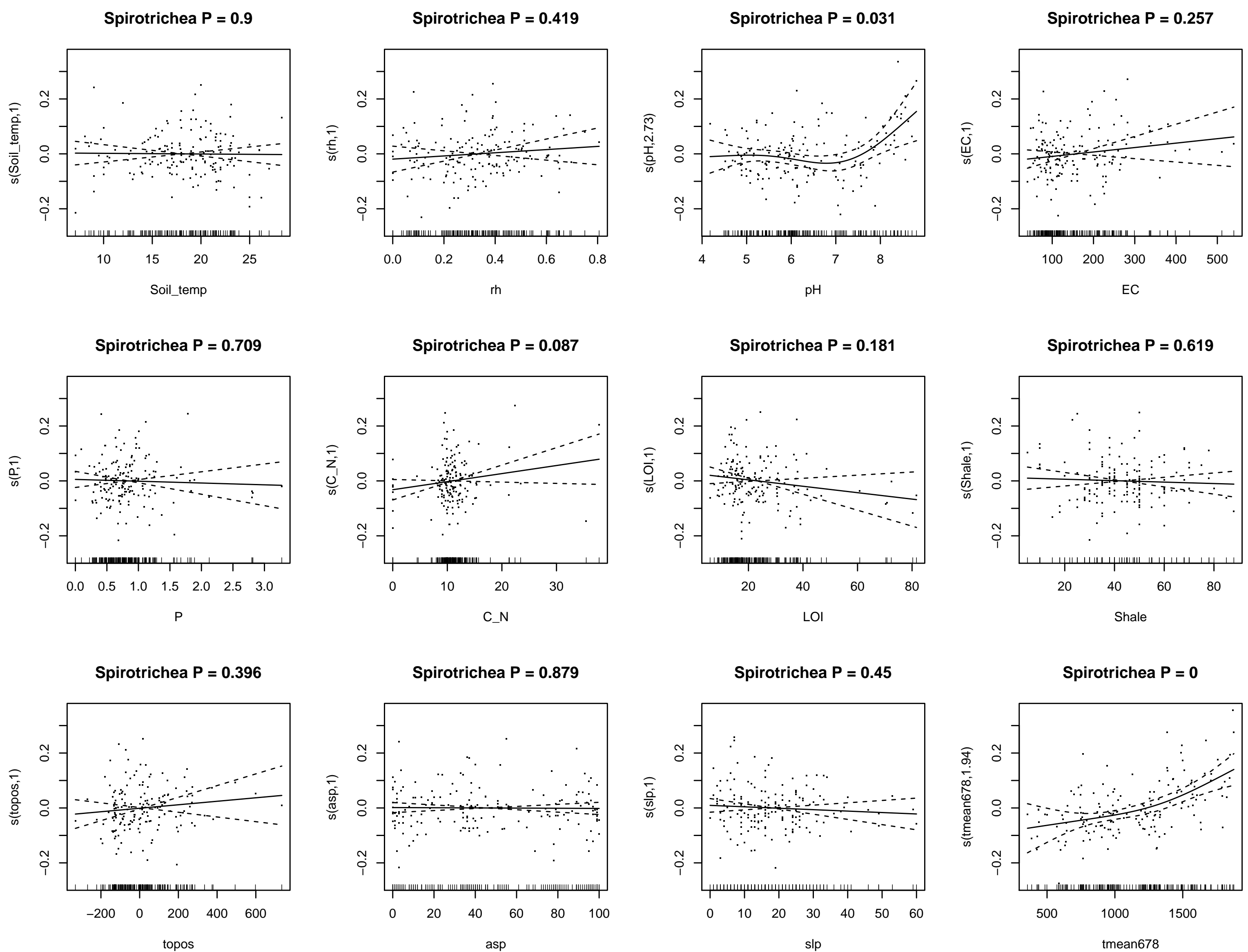


Figure S3.3: continuation

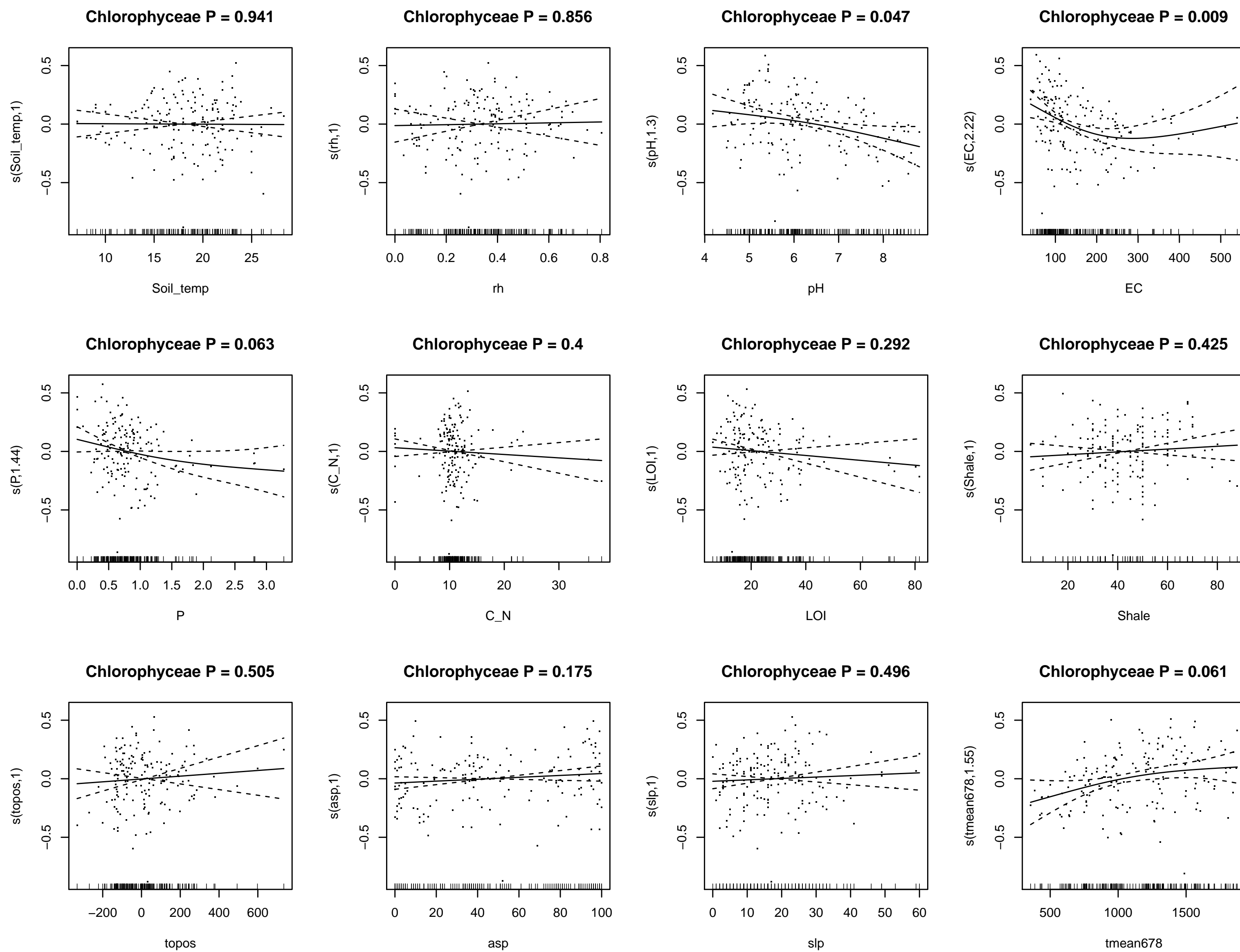


Figure S3.3: continuation

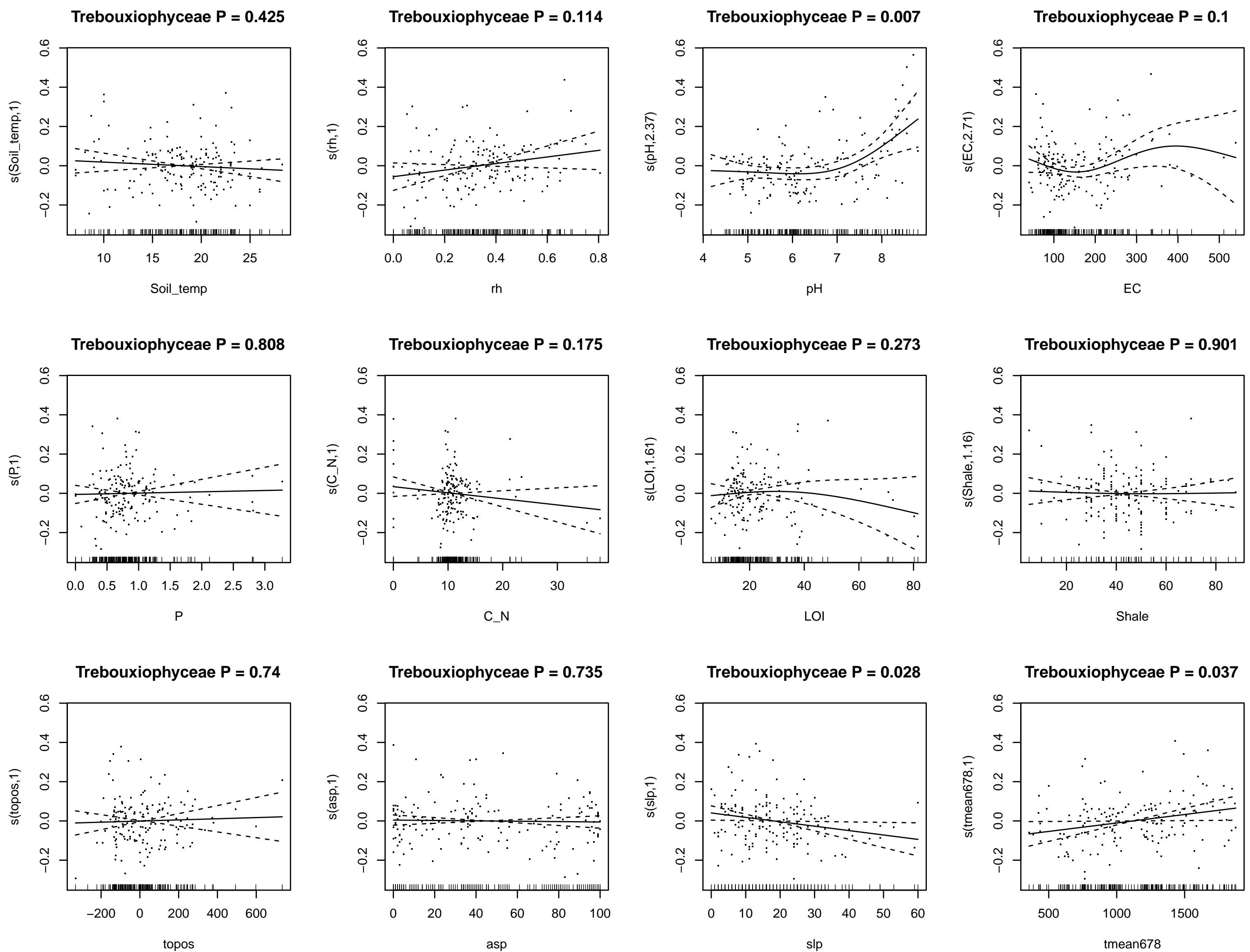


Figure S3.3: continuation

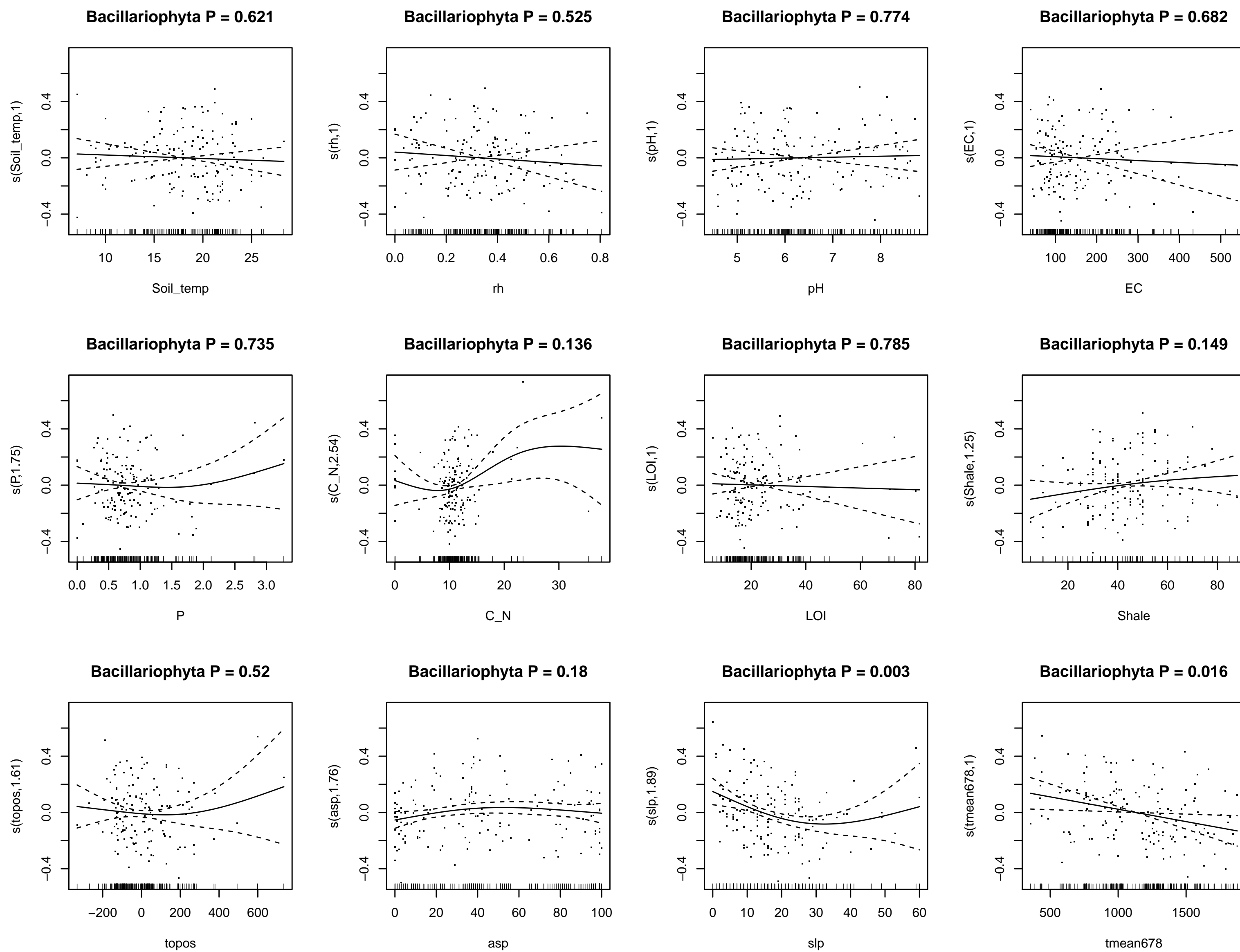


Figure S3.3: continuation

Table S3.1: Significance of topo-climatic predictors on the diversity modelled (Generalized Additive Model) from the total micro-eukaryotic community and nine broad taxonomic groups from OTUs gathered from Swiss western Alps meadow soils. The + and - signs indicate 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: $P < 0.1$, one sign: $P < 0.05$, two signs: $P < 0.01$, three signs: $P < 0.001$). 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 at Figure S3.4.

	Topo-climatic				R^2
	topos	asp	slp	tmean678	
Total	(+)	(-)		(++--)	0.09
Apicomplexa				(+)	0.05
Oomycota				-	0.03
Phytomyxea			+	(+)	0.08
Sarcomonadea		-	+	+-	0.08
Tubulinea			++	(+)	0.08
Spirotrichea				+++---	0.15
Chlorophyceae	--		---+++	--	0.22
Trebouxiophyceae	++--		++	(+)	0.11
Bacillariophyta		(-)	++	++	0.13

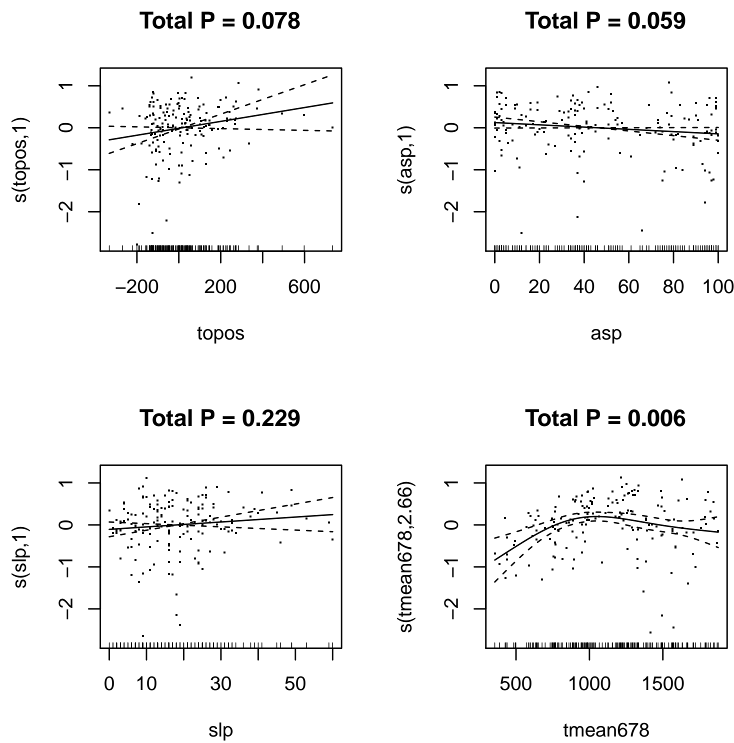


Figure S3.4: Diversity of the total protist community and all nine broad taxa in function of four topo-climatic predictors (topography, slope southness, slope steepness, average temperature from June to September) through Generalized Additive Models.

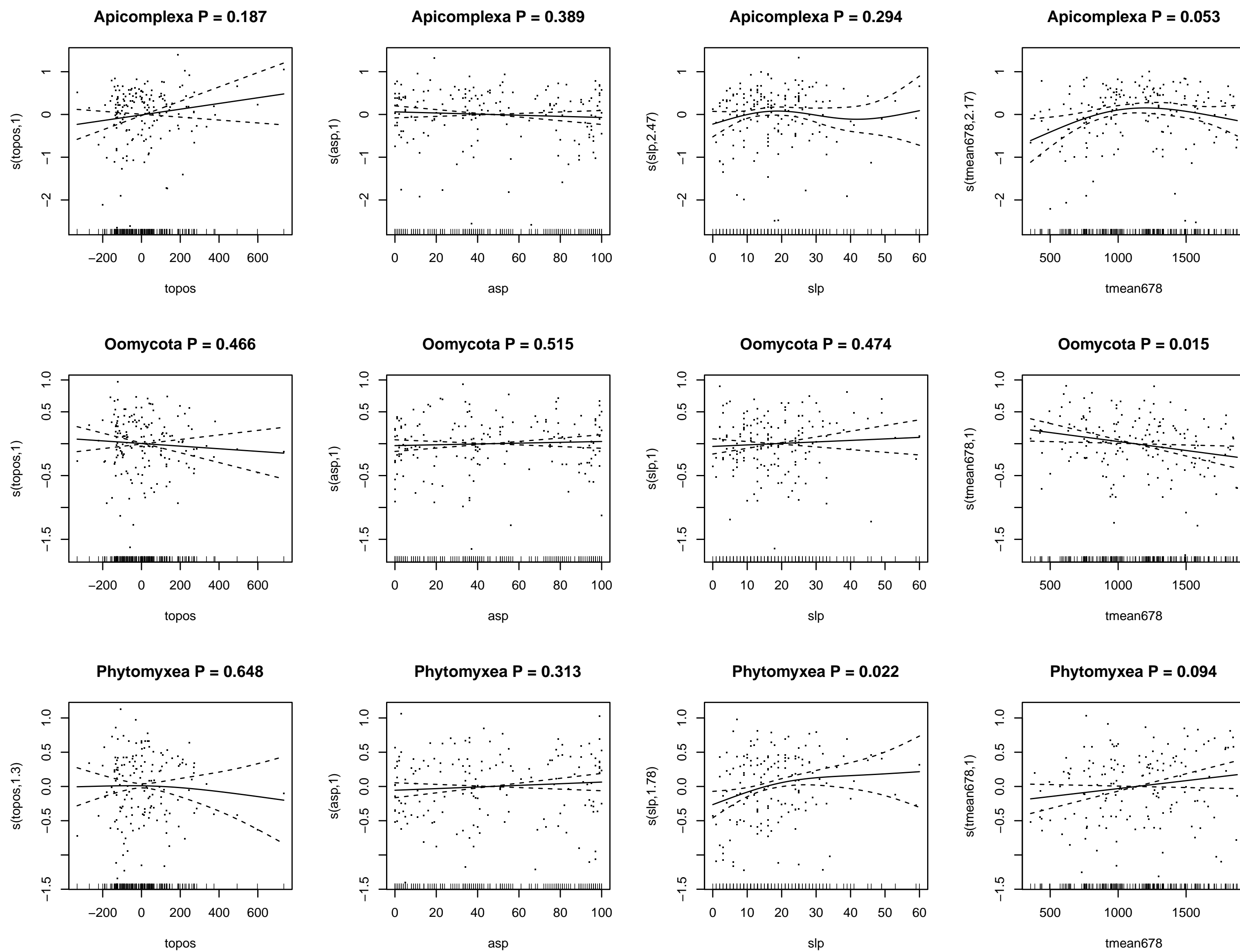


Figure S3.4: continuation

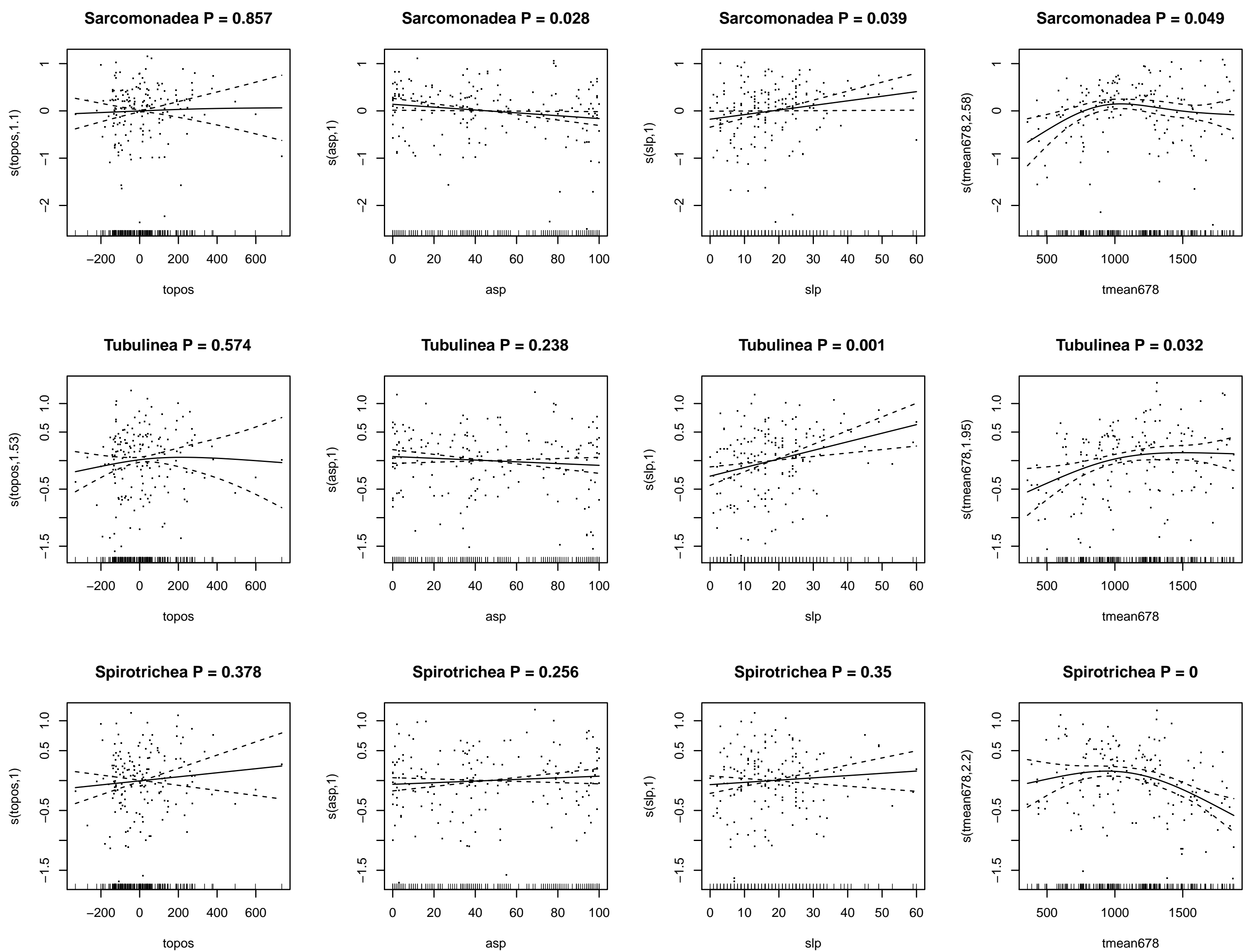


Figure S3.4: continuation

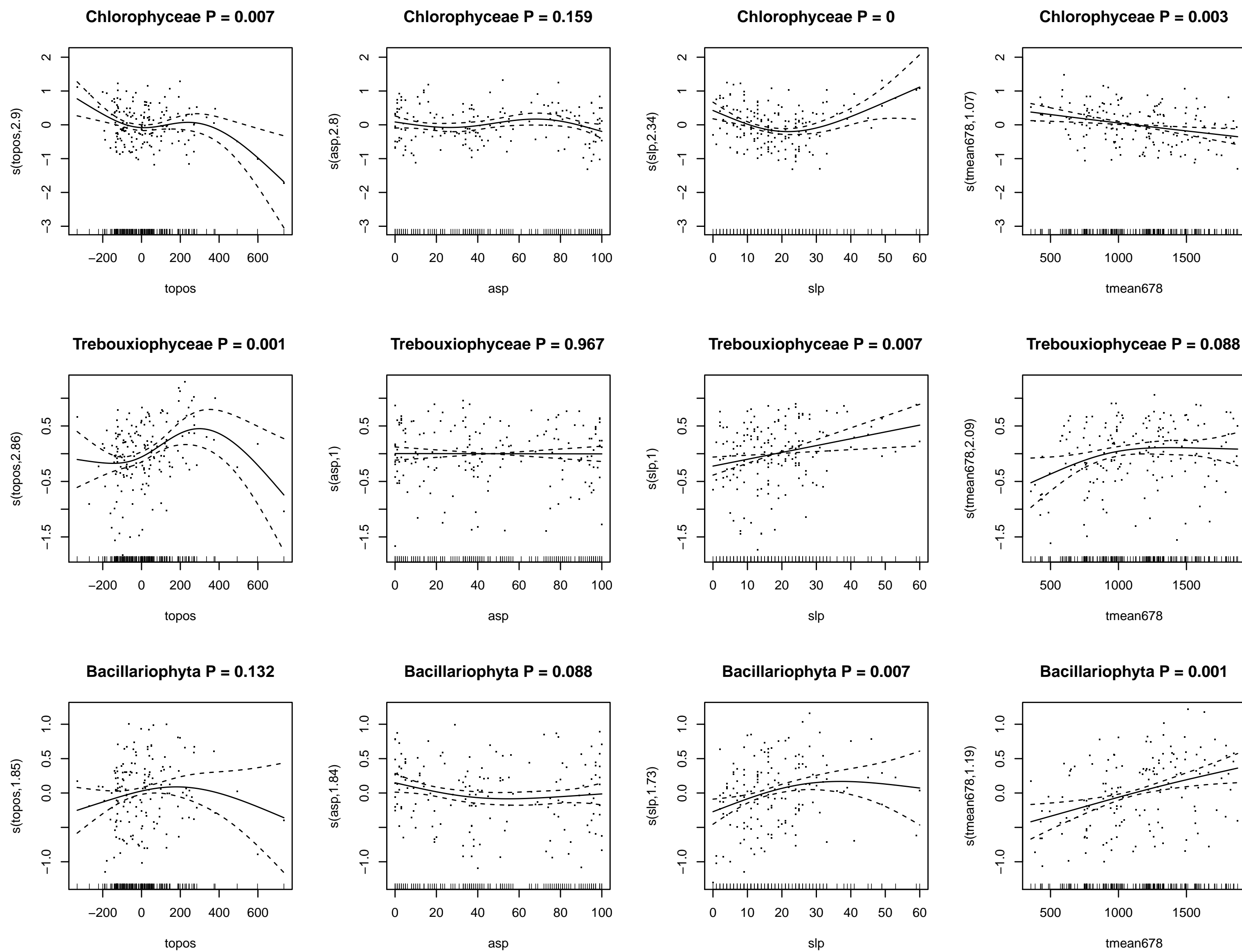


Figure S3.4: continuation

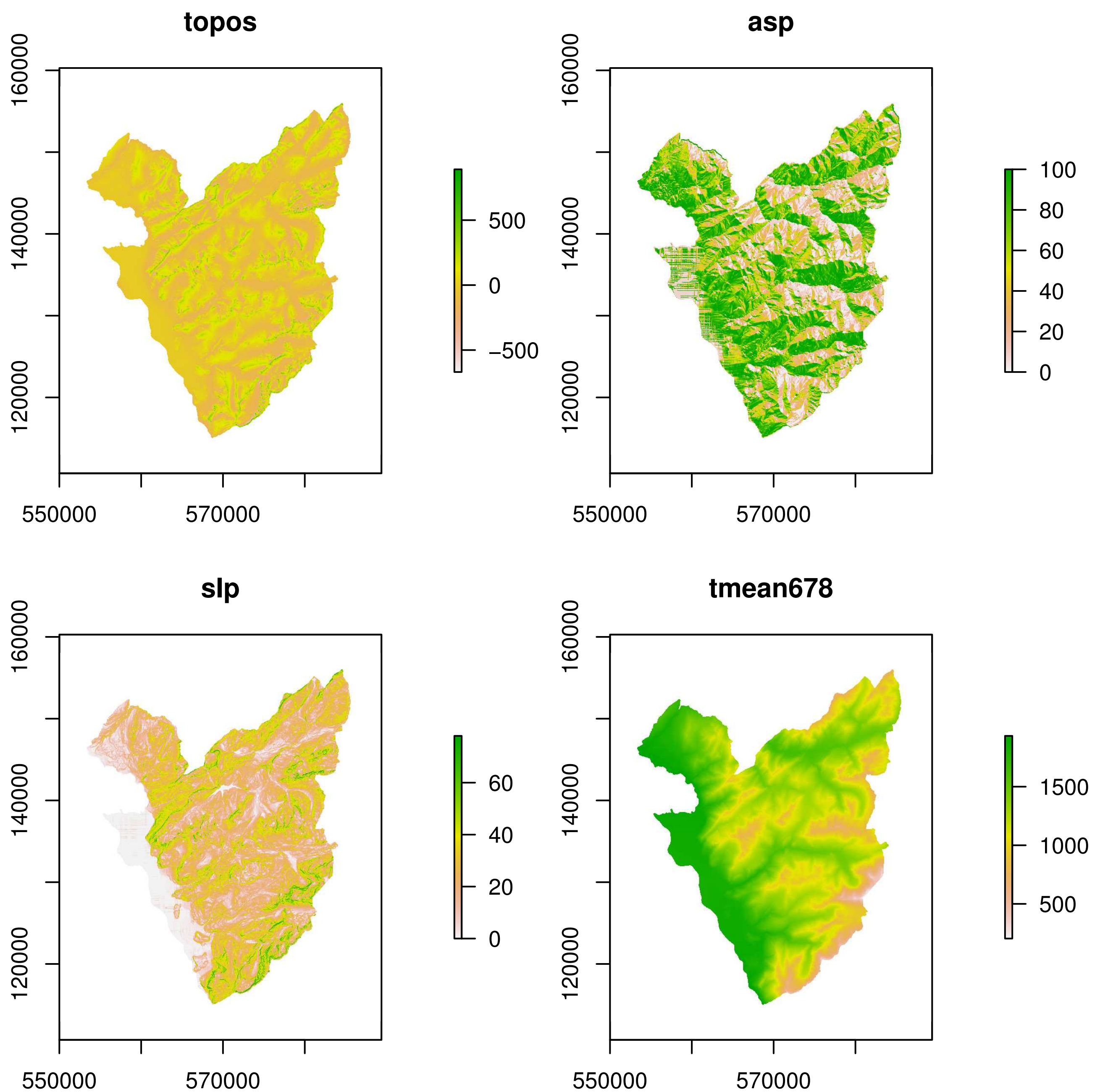


Figure S3.5: Four topo-climatic predictors (topography: topos, slope southness: asp, slope steepness: slp, average temperature from June to August from 1981 to 2010), spatialized on the area of the Swiss western Alps. These four predictors were used to model the diversity of the total protist community and nine broad taxa through Generalized Additive Models.

Soil Biol Biochem; (2017)
doi:10.1016/j.soilbio.2017.05.002

Chapter 4

Distribution patterns of soil microbial eukaryotes suggests widespread algivory by phagotrophic protists as an alternative pathway for nutrient cycling

Christophe V.W. Seppey[1*], David Singer[1], Kenneth Dumack[2], Bertrand Fournier[3] Lassaâd Belbahri[1], Edward A.D. Mitchell[1,4], and Enrique Lara[1]

[1]Laboratory of Soil Biodiversity, University of Neuchâtel, Rue Émile-Argand 11, 2000 Neuchâtel, Switzerland
[2]Department of Terrestrial Ecology, University of Cologne, Zülpicher Strasse 47b, Köln, Germany
[3]Department of Biology, Concordia University, 7141 Sherbrooke Street West, Montreal, Canada
[4]Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, CH-2000 Neuchâtel, Switzerland
[*]**Corresponding author:** christophe.seppey@unine.ch

Abstract: High-throughput sequencing (HTS) of soil environmental DNA (eDNA) allows assessing the full diversity of soil micro-eukaryotes. The resulting operational taxonomic units (OTUs) can be assigned to potential taxonomic and functional identities using increasingly complete reference databases. HTS of soil eDNA is revealing a high diversity and abundance of potential eukaryovorous protists, thus challenging the paradigm of the predominantly bacterivorous function of soil phagotrophic protists (i.e. microbial loop).

Using Illumina sequencing of soil eDNA and targeting the V9 region of the SSU rRNA gene, we investigated the taxonomic and functional diversities, distribution and co-occurrence patterns of soil micro-eukaryotes in three land-use categories: forests, meadows and croplands located in Switzerland. Each OTU was assigned to a broad functional category (phototrophs, phagotrophs, osmotrophs, or parasites).

Total OTU richness was similar in the three land-use categories, but community composition differed significantly between forests and other land-uses. The proportion of fungal sequences (especially Basidiomycota) was highest, and phototroph (i.e. soil microalgae) sequences least abundant in forests. Seven OTUs representing phagotrophic protists, together accounting for >25% of all phagotroph sequences, were significantly correlated to the total number of phototroph sequences, thus suggesting algivory. At least three of these OTUs corresponded to known algal predators.

These results suggest that beyond plants, soil microalgae represent a functionally significant but rarely considered input of carbon in soils that should be taken into account when modelling soil nutrient cycling.

Keywords: protist . eukaryotic micro-algae . phagotrophic protists . carbon cycling . V9 region of the SSU rRNA gene . high-throughput sequencing

4.1 Introduction

Our perception of the diversity and functional roles of protists is rapidly changing due mainly to the application of high-throughput sequencing (HTS) of environmental DNA (eDNA). HTS has revealed the extent of the huge unknown protist diversity in the photic zone of the world's oceans and shown that a large fraction of this diversity corresponded to mutualistic and parasitic symbionts (de Vargas et al., 2015). Likewise, studies performed on terrestrial habitats are revealing similarly high diversity of protists with a dominance of saprotrophs and parasites (Dupont et al., 2016; Geisen et al., 2014b, 2015a). These studies also revealed that many protists feed on eukaryotes, thus questioning the long-held view that soil phagotrophs fed mainly on bacteria (i.e. soil microbial loop) (Dumack et al., 2016a,b; Geisen et al.,

2015b; Geisen, 2016; Geisen et al., 2016).

Soil microbial eukaryotes, including protists and fungi, are involved in numerous biotic interactions and recognised as key actors of biogeochemical cycling (Verni and Gualtieri, 1997; van der Wal et al., 2013), and are thus considered a key element in soil fertility. However, the first (and still often the only) recognised functional role of soil protists was grazers of bacteria leading to the "soil microbial loop" paradigm, according to which phagotrophic grazing on soil bacteria releases labile compounds such as ammonium that stimulate plant growth (Bonkowski and Clarholm, 2012; Clarholm, 1985). Although feeding on bacteria is unquestionably widespread in phagotrophic microbial eukaryotes, there is increasing evidence that eukaryovory (i.e. the act of feeding partially or exclusively on other eukaryotes) is also common (Dumack et al., 2016a,b; Geisen, 2016). This implies that soil nutrient cycles are likely more complex than generally assumed.

Recent studies focusing on soil invertebrates have also questioned the origin of the carbon source feeding the soil communities, suggesting that very few soil invertebrates depend on litter (Pollierer et al., 2009) and suggesting that soil algae represent a functionally relevant source of soil carbon (Schmidt et al., 2016). The latter experimental study showed that autotrophic microbes contributed up to 17% of the body carbon of collembolan and 3% of earthworms over one week. However it is yet unclear to what extent this input is direct or if algae are first ingested by microbial grazers such as soil phagotrophs.

Several soil protists are known to be highly specialised predators of eukaryotes. For example, grossglockneriid ciliates feed exclusively on fungi (Petz et al., 1985). Parasitoids are also frequent in soils, including the widespread but still poorly studied Rozella group (also known as "Rozellida"; Lara et al., 2010 or Cryptomycota; Jones et al., 2011) which prey on chytrids, oomycetes and green algae and also include endo-nuclear parasites of Amoebozoa that ultimately cause cell death and lysis (Corsaro et al., 2014). In those cases, nutrient release by protists does not rely on bacterivory, implying pathways for nutrient cycling alternative to the microbial loop. It is unclear how quantitatively relevant this pathway is but one way to assess this is to study the diversity and abundance of taxa involved in these trophic relationships using the now available data from massive sequencing of soil environmental DNA.

The true diversity of soil protists has long been poorly known, mainly due to methodological limitations for their isolation, culture and subsequent identification (Ekelund and Ronn, 1994; Foissner, 1999b). Metabarcoding (environmental DNA amplicon based identification) of high-throughput sequencing data is now the golden standard for environmental screening of microbial diversity (Pawlowski et al., 2016). HTS data may also inform on the functioning of ecosystems based on the genetic identification of the organisms and knowledge on their lifestyles (de Vargas et al., 2015; Lara et al., 2015; Massana et al., 2014). The next step is to infer the biotic relationships between these organisms, which can be hypothesized when OTUs co-occur systematically across many samples, as can now be assessed by HTS. In practice, the nature of these relationships (i.e. trophic, but also symbiosis, competition, etc.) is not known, and co-occurrence data can thus be difficult to interpret in biological terms. Examples of known relationships taken from the literature can however illustrate well-supported co-occurrence and clarify the true nature of these relationships between organisms. Examples are manifold: predation of ciliates on fungi (Petz et al., 1985), of cercozoa on chlorophytes (Dumack et al., 2016a; Hess et al., 2012; Hess and Melkonian, 2013) but also symbioses, like between trebouxiophytes and testate amoebae (Gomaa et al., 2013). Putative relationships inferred from metabarcoding studies can also be explored by conducting new observations and experiments.

Phototrophic protists (i.e. eukaryotic algae) in soils include mostly exclusive free-living phototrophs (e.g. Bacillariophyta, Chrysophyceae, Xanthophyceae) and photosymbionts as in lichens (e.g. Trebouxiophyceae). Soil eukaryotic algae constitute an important part of the so-called cryptogamic crusts, which represent a significant carbon input in arid ecosystems (Elbert et al., 2012; Freeman et al., 2009; Frey et al., 2013). They are however also widespread in more humid soils but their functional role there is less well known and, consequently, has not been considered in the classical model of the soil microbial loop (Berard et al., 2005).

In order to assess the patterns of micro-eukaryotic taxonomic and functional diversities and address questions such as the possible role of soil algae as a carbon source it is useful to compare contrasted terrestrial ecosystems. Here we describe and compare the overall diversity and community structure of soil micro-eukaryotes in forest, meadow and cropland soils from 44 sites in Switzerland based on Illumina sequencing of the V9 region of SSU rRNA gene. Based on these data, we explored more specifically the abundance patterns of phototrophs and the co-occurrence patterns with their potential phagotroph predators. This trophic link was also explored by direct microscopic observations.

4.2 Materials and methods

Sampling

We collected 44 soil samples in permanent plots of the Swiss Biodiversity Monitoring program which aims to assess biodiversity all over Switzerland (BDM `http://www.biodiversitymonitoring.ch/en/home.html`). The sites included three land-uses which cover most of the Swiss territory (16 forests, 16 meadows and 12 croplands) (Figure 4.1, Table S4.1) and spanned a diversity of soil types that could be arguably considered as representative of the entire country. Likewise, samples were collected in a range of altitudes covering most of the Swiss territory (excepted alpine sites). We expect therefore to cover a significant part of the microeukaryotic diversity present in Swiss soils. In this purpose, each sample was

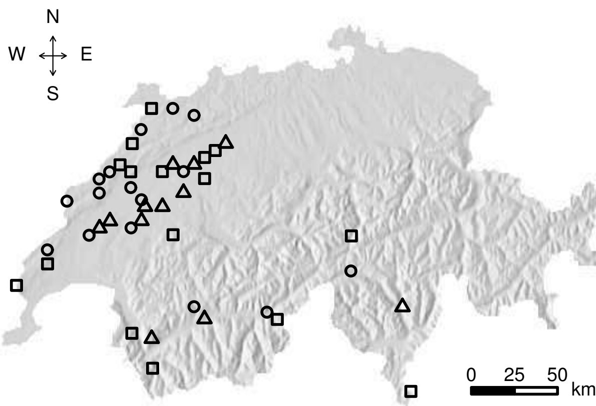


Figure 4.1: Location of the 44 sampling sites in Switzerland. The squares, circles and triangles indicate forests, meadows, and croplands, respectively.

characterized using the typology of Swiss natural habitats (Delarze et al.) (Table S4.1). Forests included both coniferous (e.g. *Picea abies*), or broadleaved trees (e.g. *Fagus sylvatica*). Most meadows were amended and used to produce fodder. Croplands were used for maize, cereals or tobacco cultivation. Meadows and croplands were designed as open habitats as much more light reach their soil surface than in forests. Sampling was performed over one month between September 27th 2012 and October 31th 2012. At each site, three topsoil cores (5cm diameter x 5cm depth) were taken along a circle of 1 m radius in the same land-use and pooled. Soil samples were kept cool (in an icebox) and DNA was extracted within 2-3 days.

DNA extraction, amplification and sequencing

DNA was extracted using the MoBio PowerSoil extraction kit (Carlsbad, CA, USA) according to the manufacturer instructions. The SSU rRNA V9 region was amplified using the broad spectrum eukaryotic primers 1380F/1510R (`CCCTGCCHTTTGTACACAC` / `CCTTCYGCAGGTTCACCTAC`; Amaral-Zettler et al., 2009). We used the smaller V9 region in this study instead of V4 because (1) we expected to have less taxonomical biases by using a fragment whose length is almost constant in all eukaryotes (as opposed to, for instance, the V4 region; (de Vargas et al., 2015)) and (2) because the fragment is shorter, the probability of generating artefactual diversity (i.e. chimeras) is lower (Valentini et al., 2009). PCR reactions were run in triplicates with a PTC-200 Peltier Thermo Cycler (BioConcept, Allswill, Switzerland) with 1 ng of environmental DNA, 6 µL of 10 x PCR buffer, 0.6 µL of each primer, 0.6 µL of each dNTP 400 µM (Promega, Dübendorf, Switzerland) and 0.2 µL of 0.05 U/µL GoTaq (Promega, Dübendorf, Switzerland). The volume was adjusted to 30 µL with ultra-pure water. Amplification was conducted with the following conditions: denaturation at 94 °C for 3 min, 30 cycles at 94 °C for 30 s, 57 °C for 60 s and 72 °C for 90 s and final extension at 72 °C for 10 min (Amaral-Zettler et al., 2009). PCR products were purified using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and pooled together at the same concentration prior to sequencing. A DNA library was prepared using the New England Biolabs's kit NEBNext DNA Sample Prep Master Mix Set 1. Illumina HiSeq sequencing was done by Fasteris (Geneva, Switzerland) using an Illumina HiSeq 2000 technology to obtain paired-end reads (2 X 100bp).

Bioinformatic analyses

The PR^2 database (Guillou et al., 2013) was used as the reference database for a first taxonomic assignation of reads and OTUs; only sequences containing complete forward and reverse primers described above were retained. SSU sequences from bacteria and archaea were also added to the PR^2 database from the Silva database (Pruesse et al., 2007) in order to identify and remove eventual prokaryotic sequences from the analysis. Prokaryotic SSU sequences were truncated from the general primer 1389F (`TTGTACACACCGCCC`; Amaral-Zettler et al., 2009) to the end of the SSU rRNA sequence and kept as ortholog of the eukaryotic V9 fragment. The truncated prokaryotic sequences were then de-replicated before being added to the PR^2 database.

Reads were merged using the program `flash` (v. 1.2.9; Magoc and Salzberg, 2011) and demultiplexed into samples using the program `sabre` (`https://github.com/najoshi/sabre`). Only sequences containing complete forward and reverse primers described above were kept. Good quality sequences were selected according to the method used in de Vargas et al. (2015). Chimeric sequences were then discarded using the software `usearch` (v. 7.0.1090; Edgar et al., 2011) by comparing reads against the PR^2 (Guillou et al., 2013) and Silva (Pruesse et al., 2007) databases and against reads within the sample. In order to remove artefactual sequences, we kept only those that were found at least three times in two samples (de Vargas et al., 2015).

OTUs were clustered using the software `swarm` (v. 1.2.5) (Mahé et al., 2014) with the default set-up. OTUs were then taxonomically assigned by aligning the dominant sequence of every OTU against the PR^2 database using `ggsearch` (`Fasta` package v. 36.3.6; Pearson, 2000). The OTUs were considered as undetermined eukaryotes if their percentage identity with sequences of PR^2 was lower than 80% as in de Vargas et al. (2015). We also removed sequences belonging to prokaryote, Metazoa or Embryophyceae. In order to homogenize the number of reads present in all samples for further numerical analyses, we randomly selected 50'000 for each sample.

Assignation to functional groups and numerical analyses

We selected 41 taxa with well-characterized trophic function (i.e. 5 osmotrophs, 5 parasites, 6 phototrophs, 25 phagotrophs) in the list of divisions, classes and orders

of the PR[2] assignation for the diversity analyses (Table S4.2).

As a first comparison of community composition, we calculated the Shannon index and performed a non-metric multidimensional scaling (NMDS) analysis on OTU abundances. We assessed the difference in diversity among land-use types with a non-parametric multiple comparisons Nemenyi test (Hollander and Wolfe, 1999) (`posthoc.kruskal.nemenyi.test` function, package `PMCMR` v. 4.1 (Pohlert, 2014)), and also calculated NMDSs for each pairs of land-use and tested the community difference by a permutation test (`envfit` function `vegan` package v. 2.0-10, (Oksanen et al., 2013)). P-values were multiplied by three to take into account multiple tests adjustment (Holm, 1979).

We then assessed in which environment sequences belonging to phototroph organisms were most abundant using a Nemenyi test. To retrieve putative algae consumers, we measured the correlation between each of the 100 most dominant phagotroph OTUs and the total abundance of phototrophs, taking also into account land-use as second environmental variable in linear models (LMs). To normalize the distribution of both phagotroph OTUs and total phototroph abundance we log transformed their sequences abundances (`decostand` function, `vegan` package v. 2.0-10, (Oksanen et al., 2013)). The two environmental variables (i.e. total phototroph abundance, land-use) were tested independently in the LMs as none of the model tested showed significant interaction. We finally adjusted the p-values of the two environmental variables for the 100 models according to Holm (1979). We also verified if each of the LM respected conditions of residuals normality and homoscedasticity by performing a Shapiro test on model residuals and non-constant variance test (function `shapiro.test` and `ncvTest`, packages `car` v. 2.0-20; (Fox and Weisberg, 2011), and stats v. 3.1-0; (R_Core_Team, 2014) respectively). OTUs respecting the LM conditions and showing a significant correlation with phototroph abundance, were selected as putative alga consumers and their taxonomy was verified on GenBank by using Blast with the default parameters.

Isolation of protists and microscopic observation on algivorous behaviour

To illustrate the trophic interactions among selected protists and test if identified co-occurrences indeed could be interpreted in terms of trophic relationships, we documented by microscopical observations organisms from the same genus/species as the OTUs whose abundances were positively and significantly correlated with those of algae. *Rhogostoma* sp. was isolated from leaf surfaces (Cologne, Germany), *Leptophrys vorax* was isolated from a freshwater puddle (Cologne, Germany), and *Trinema* sp. appeared as a contamination in such protist cultures. All protists were morphologically determined. The illustrated organisms were identified morphologically based on unmistakable criteria, which were corroborated by taxonomic literature (Howe et al., 2011; Hess et al., 2012; Lara et al., 2007b) but were not sequenced in the frame of this study.

The pictures of *Leptophrys vorax* were obtained from an individual directly taken from a natural sample. Other organisms were cultured in Waris-H medium (McFadden and Melkonian, 1986) at room temperature on a window bench and enriched with *Characium* sp. and an undetermined coccoid green alga. The cultures were checked for potential algal ingestion after three days of incubation, using an inverted microscope (Nikon Eclipse TS-100, Japan) at 100x and 400x magnification. Pictures were taken with a Nikon digital sight DS-U2 camera (program: NIS-Elements v 4.13.04) and a Nikon Eclipse 90i (DIC, up to 600x magnification).

4.3 Results

Data quality and overall diversity

The full dataset contained 15'365'116 raw reads, of which 93.9% passed the quality check, 87.5% were found at least three times in two samples, 87.4% were not considered as chimeras, and 77.4% were not considered as Metazoa, Embryophyceae or prokaryotes. Therefore, a total of 11'893'592 reads were left for further analyses. In the dataset adjusted to 50'000 sequences by sample, we retrieved a total of 18'586 OTUs, of which 87% could be taxonomically assigned unambiguously according to the assignation threshold; altogether, representing 97% of the reads and 75% of the OTUs (Figure S4.1, S4.2). The most abundant supergroup of eukaryotes in all samples were Opisthokonta (Fungi), followed by Rhizaria (Cercozoa) and Stramenopiles.

The most noticeable difference in relative abundance of taxa could be observed between open and forest habitat, and was mostly due to a divergence in the abundance of Basidiomycota (Figure S4.1). In contrast, richness did not differ deeply between land-use types, and varied between 2371 and 3516 OTUs. Richness was dominated by both Fungi and Rhizaria, more or less in equal proportions, followed by Stramenopiles (Figure S4.2).

Shannon diversity and micro-eukaryotic community composition differed significantly between forest and open habitats (meadows and croplands) (Nemenyi test, and permutation test on NMDS after correction, $P < 0.001$) while diversity and communities did not differ significantly between meadows and croplands ($P > 0.05$; Figure 4.2, S4.3).

Diversity and abundance patterns

Fungi (the large majority of the osmotrophic taxa) represent more than 50% of the overall abundance and 43%, 51% and 67% of all reads in croplands, meadows and forests respectively (Figure S4.1). The dominance of Fungi in forest samples was mostly due to the presence of a single Basidiomycota OTU (X3), which accounted for 38% of the totality of all reads in forests. Although taxonomic resolution of the SSU rRNA is too low to differentiate between fungal species, the OTU X3 could be

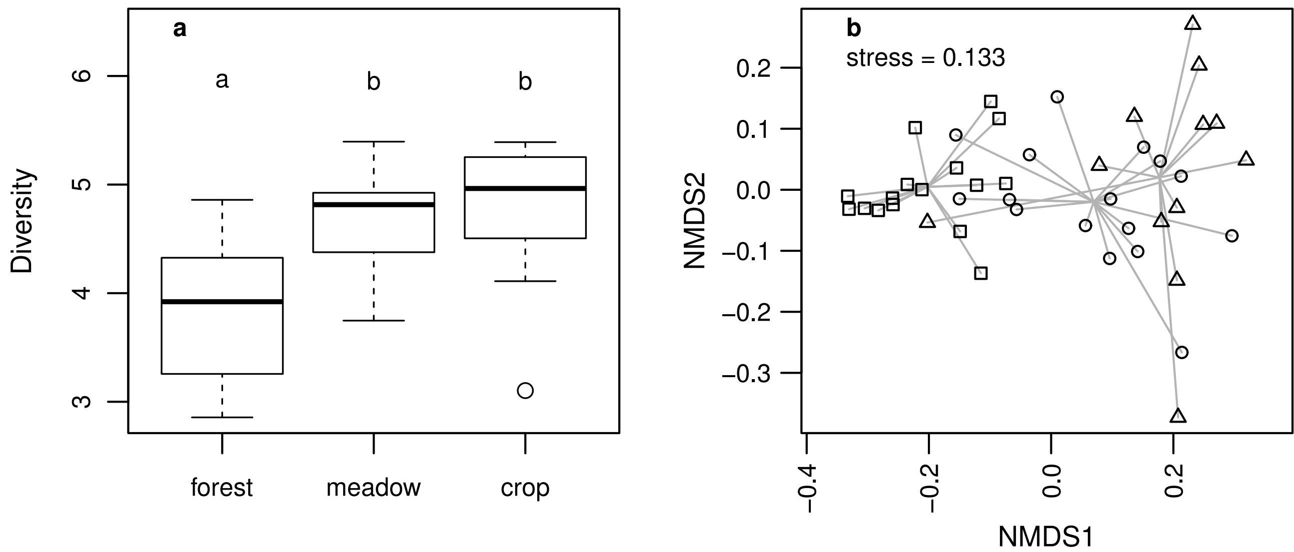


Figure 4.2: Distribution of OTUs Shannon diversity for each environments (a) and non-metric multidimentional scaling (NMDS) based on Bray-Curtis dissimilarities of 44 soil samples from Switzerland. The letters above the boxplots represent groups of environments expressing significant different diversity distribution according to a Nemenyi test (P < 0.05). The three land-uses are denoted by squares (forests), circles (meadows) and triangles (croplands).

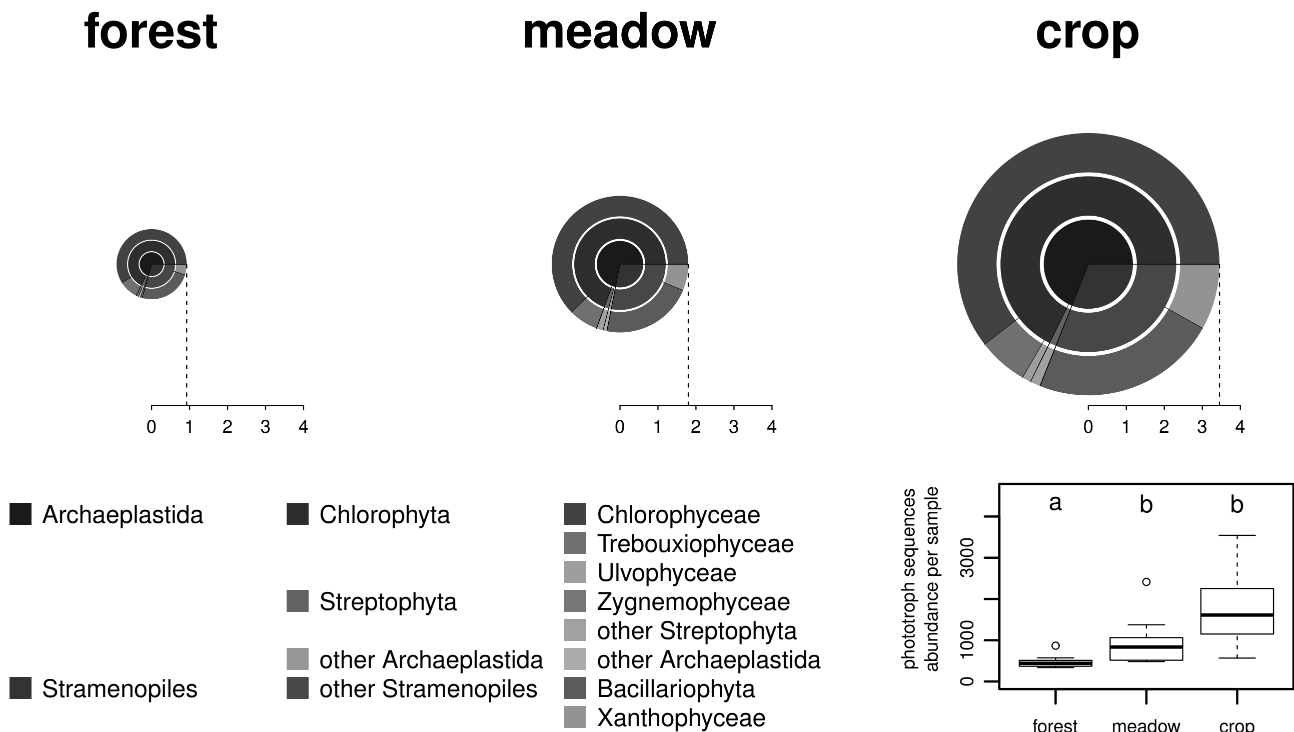


Figure 4.3: In the upper part, the relative abundance of phagotroph taxon sequences in the three land-use types. The radius of the pie-chart represents the percentage of phototroph sequences in each land-use. Taxa representing less than 1% of the land-use are represented in the "other Streptophyta" or "other Archaeplastida" section. On the bottom right part, boxplots representing the abundance of phototroph sequences according to the land-use. Letters on the top part of the boxplots represent the groups of land-use formed according to the Nemenyi test (P < 0.05) on the phototroph sequences abundances.

assigned (100% match) to a wide array of Agaricomycotina, (e.g. *Leucopaxillus*, *Ampulloclitocybe*). In addition to Basidiomycota, the two next dominant groups of fungi were the Ascomycota and the Mucoromycota.

Potential parasites reached 3.9% of the overall abundance representing 2.7%, 4.4% and 4.8% of the reads of croplands, meadows and forests respectively. Among these numbers, Oomycota represented the large majority of parasites abundance regardless of the land-use (75%, 61%, 57% in croplands, meadows and forests respectively). The five most abundant Oomycota OTUs were assigned with confidence ($\geq$ 94% of identity) to genera Aphanomyces (X99), *Pythium* (X53, X8), *Pythiopsis* (X31) and *Saprolegnia* (X30). Oomycota were followed in abundance by Mesomycetozoa, Gregarinasina and Phytomyxea, whose relative abundance varied depending on

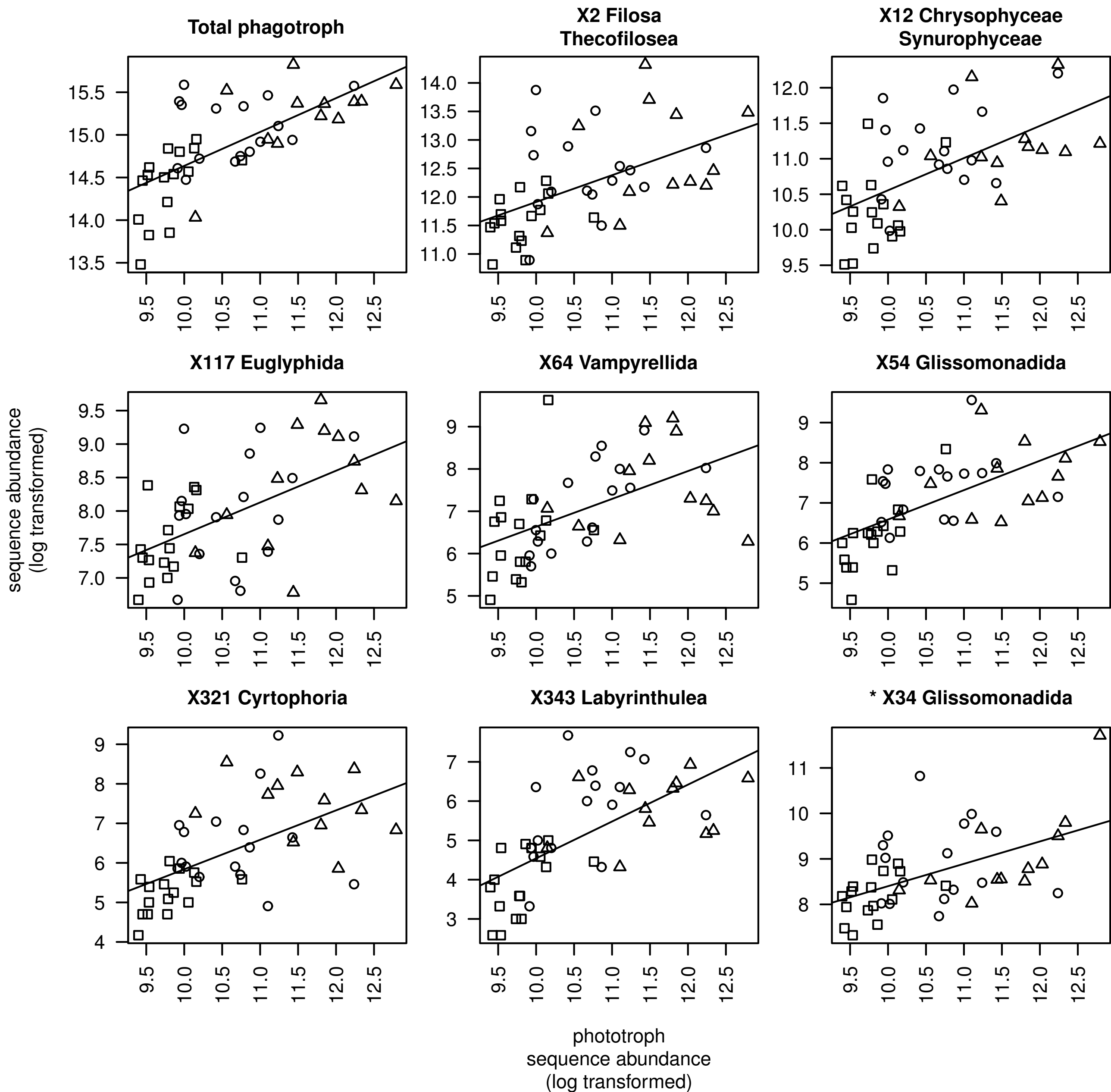


Figure 4.4: Biplots showing the regression between the abundance of total phagotroph and eight OTUs and phototroph abundance. The identifiant of each OTU is shown on the top of each graphic. The asterisk indicate the OTU which belong to the ten most dominant phagotroph OTU and respond significantly to the phototroph despite an heteroscedastic distribution.

the land-use (8% to 21%, 4% to 17% and 5% to 11% respectively).

Sequences belonging to OTUs assigned to phototrophic organisms accounted for 1.9% (42'475 sequences) of all sequences. This proportion was highest in open habitats, representing 3.5%, 1.8% and 0.9 % of the sequences found in croplands, meadows and forests, respectively (Figure S4.1). As for overall community patterns, this difference was statistically significant between forests and the other two land-uses (Nemenyi test after correction, $P < 0.01$; Figure S4.1). The diversity of phototrophic micro-eukaryotes was largely dominated by Chlorophyceae, followed by diatoms (=Bacillariophyta), and Trebouxiophyceae or Xanthophyceae; the rest being shared by other typical subaerial algae like Ulvophyceae and other Archaeplastida (Figure S4.1).

OTU assigned to phagotrophic organisms accounted for 32% (704'308 sequences) of all sequences. Seven of the 100 most dominant phagotroph OTUs showed a positive correlation to total phototroph sequence abundance, and respected the conditions of residuals normality and homoscedasticity (Figure 4.4, S4.4 and Table S4.4). These OTUs belong to Cercozoa (X2, X117, X64, X54), Ciliophora (X321) and Stramenopiles (X12, X343) and together account for 27% of the phagotroph sequence abundance and 8% of the total abundance of all sequences of the dataset (Table S4.3). Apart from one OTU assigned to Labyrinthulea (X343), all other OTU sequences obtained a good match ($\geq$ 97%) with sequences from the GenBank database (Table S4.3). In addition to these seven OTUs we observed that X34 (Glissomonadida, group of Viridiraptor) was correlated to the abundance of eukaryotic algae and among the ten most

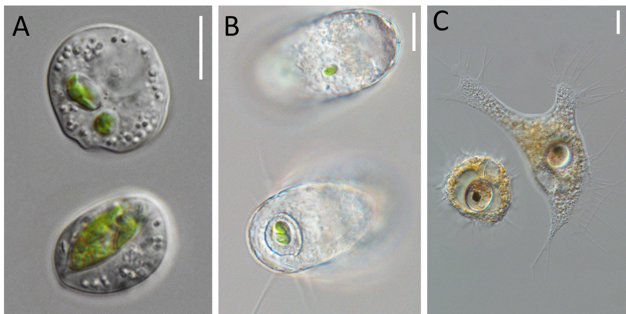


Figure 4.5: Pictures of three selected organisms, closely related to the found OTUs correlating to the phototroph sequence abundances (*Rhogostoma* sp. (a), *Trinema* sp. (b), *Leptophrys vorax* (c)). The scale bar represent 10µm.

abundant phagotrophic OTUs despite the fact that the linear model did not respect the conditions of homoscedasticity (Figure 4.4, S4.4 and Table S4.3, S4.4). In addition to these eight OTUs, seven other OTUs were also correlated to total phototroph sequence abundance but were rare and without a homoscedastic distribution (Table S4.4).

Microscopic observations

We screened environmental samples to find organisms corresponding to the taxonomic assignation of OTUs significantly correlated to the sequence abundance of phototrophic protist taxa and found three (i.e. X2, X64, and X117) out of the seven identified by our analyses. All three organisms were observed with most likely ingested algal material, either in the natural samples (*Leptophrys vorax*, potentially linked with X64), or when incubated with algal cells (*Rhogostoma* sp., *Trinema* sp., potentially linked with X2 and X117) (Figure 4.5). *Rhogostoma* spp. are characterized by the presence of a hyaline theca with a cleft-like apertural opening (not shown) and filopodia. Similar to *Rhogostoma* spp., *Trinema* spp. exhibit filopodia, but in contrast bear large circular scales and a subterminal, ovoid or round aperture. The filose genus *Leptophrys vorax* is characterized by being naked, sometimes with slightly orange cytoplasm, the ingestion of diverse groups of algae and the transformation between the isodiametric and expanded morphotype. Since all these features were observed in our isolates, we determined them as such.

4.4 Discussion

Metabarcoding studies are revealing not only a huge unknown diversity but also unsuspected trophic interactions in every studied environment (de Vargas et al., 2015). The predation of phototrophs by heterotrophic protists suggested by our data implies a carbon input to the soil ecosystem that was not taken into account by the traditional microbial loop model and is in line with a recent study focusing on soil invertebrates (Schmidt et al., 2016).

Overall diversity and community patterns

The diversity patterns of individual micro-eukaryotic groups across the three land-use types is coherent with the contrast among these habitats. Fungi dominate micro-eukaryotic communities in forest soils (Behnke et al., 2011; Geisen et al., 2015a; Glaser et al., 2015; Lesaulnier et al., 2008). In our data, this dominance was explained by the presence of the OTU X3, which is assigned to Fungi that build ectomycorrhiza (e.g. *Leucopaxillus*, *Ampulloclitocybe*) (Cairney and Chambers, 1999) and thus most likely establishes symbiotic relationships with trees. Alternatively, some Fungi represented also by OTU X3 (e.g. Auriculariaceae, *Panaeolus*) (Boddy et al., 2007) are known as wood decomposers, which would explain their high abundances in forests.

The next most represented groups (Stramenopiles, Rhizaria) comprise also organisms that can be encountered often in soils, such as Oomycota, Cercomonadida and Chrysophytes (Lesaulnier et al., 2008). Fungi, despite of being by far, the most abundant microbial eukaryotes in soils, had a richness which was comparable to Rhizaria (Figure S4.2). This is most probably due to the highly ramified hyphae, and high biomass, in comparison to the mostly small and unicellular Rhizaria.

The distribution of parasites follows the different land-use characteristics. The deceasing abundance of Arthropod parasites (Mesomycetozoa and Gregarines) from forests to meadows to is in line with the corresponding decline in plant biomass, habitat complexity and diversity of ecological niches for their hosts. Additionally, pesticide use is highest in cropland and further reduces insect diversity and biomass (Lachat et al., 2011). Croplands are associated with an increase in Oomycota abundance, where OTUs assigned with acknowledged crop diseases (X99: 100% identity with *Aphanomyces euteiches*; X53: 100% identity with a crop pathogen group of a *Pythium* sp.).

Phagotroph vs phototroph

The abundance of eight OTUs representing phagotrophs was strongly and significantly correlated to the total abundance of phototroph OTUs (Figure 4.4). Six of these could be assigned with confidence to known genera, their sequences having over 97% identity with cultured organisms (Table S4.3). Amongst these taxa, four (*Rhogostoma*, *Platyreta*, *Trinema* and *Pseudochilodonopsis*) are relatively large sized protists (>20 µm) - and thus potential predators of micro-eukaryotes, including phototrophs. We illustrate three of these species in the act of predating algae (Figure 4.5). *Rhogostoma* spp. (Rhogostomidae, Thecofilosea) (X2) are closely related to recently characterized eukaryvores that have been shown to avoid feeding on bacteria (Dumack et al., 2016a,b). Although some strains of *Rhogostoma* spp. can live exclusively on a bacterial diet (Howe et al., 2009), we could show that at least certain species of genus *Rhogostoma* do feed on algae. OTU X64 is a member of the exclusively eu-

karyvorous Leptophryidae (Vampyrellida). Members of this family have been reported to be algal predators to a large extent (see Figure 4.5, represented by the closely related *Leptophrys vorax*) (Bass et al., 2009; Gong et al., 2015; Hess et al., 2012). Co-occurrence patterns and observational data provide two lines of evidence that suggest that X64 actually feeds on algae. The same conclusions can be drawn for *Trinema* spp. (Euglyphida) (X117) where larger members of the genus feed to a large extent on micro-algae (Cyanobacteria and/or pigmented Eukaryotes) (Meisterfeld, 2000a; Santibanez A et al., 2011). Our observations confirmed the ingestion of algal material (Figure 4.5). *Pseudochilodonopsis* (X321) are considered as exclusive algivores specialized on diatoms (Hamels et al., 2004).

Labyrinthulomycetes branching within the Amphifilidae (X343) are a diverse group (Pan et al., 2016) including bacterivores such as *Sorodiplophrys stercoraria* and *Amphifila marina* (Anderson and Cavalier-Smith, 2012; Tice et al., 2016). The taxonomic as well as functional diversity of this group is however only marginally documented, and the existence of algivorous forms is thus possible. The group of *Spumella*-like Chrysophyte (X12) is composed of small phagotrophic flagellates having lost their photosynthetic abilities secondarily. However, it has been shown that transitions between phagotrophic and phototrophic strategies occurred often in the evolutionary history of Chrysophytes. It is possible therefore that the *Spumella*-like Chrysophyte X12 is actually mixotrophic like many Chrysophyceae (Boenigk et al., 2005), and therefore shares higher light requirements with other phototrophs. Alternatively, it is possible that the *Spumella*-like Chrysophyte X12 feeds preferentially on bacteria that are associated to phototrophs and their exudates. Bacterial communities associated to algae are highly influenced by the host in aquatic systems (Sapp et al., 2007). A similar explanation could possibly be given for *Allapsa* (X54), a genus of small Cercozoan flagellates formerly collectively classified under the name "*Heteromita globosa*" (Howe et al., 2009).

To the contrary, OTU X34 is assigned to the Viridiraptoridae, a family of highly specialised Cercozoans feeding as yet known exclusively on phototrophic organisms (Hess and Melkonian, 2013). The linear model obtained for this OTU did not respect the conditions of homoscedasticity because of its high abundance in two samples. Such high sequence abundance may correspond to local blooms of these small flagellates, which are reported as frequent (Hess and Melkonian, 2013).

Altogether, phagotroph sequences belonging to an OTU co-occurring with phototrophs reached 26.9% of all phagotrophs (28.1% if X34 is considered). Thus, if only those organisms that we observed eating algae are actually playing that role, then 19.8% of all phagotrophs could actually feed (to various degrees) on phototrophs. Based on this, we estimate the total proportion of algal-feeders to account for between one fifth and one third of all phagotrophic sequences, an amount which is far from being negligible. It is noteworthy that, out of the 100 best represented phagotrophic OTUs, only seven were robustly correlated to phototroph abundance. This low number suggests that the correlations observed are probably highly specific, as demonstrated for X34 (Viridiraptoridae; (Hess and Melkonian, 2013)). Although it is known that organisms such as *Trinema* can feed opportunistically on various eukaryotes such as fungal conidia (Santibanez A et al., 2011), the food regime of all members of this very diverse genus (Lara et al., 2016) has not been surveyed and it is still possible that some members are specialized in eating algae, at least to a certain extent. At this point, only experimental evidence can demonstrate if the selected OTUs represent organisms that are exclusive algal predators or not.

Whatever percentage of environmental sequences from phagotrophic organisms interacting with phototrophs is taken as a reference, the corresponding number of phototroph sequences is by far lower. As rRNA gene sequence numbers can be considered to providing reasonably accurate estimations of the relative biomass of the organisms in DNA environmental surveys (Giner et al., 2016), this suggests that the standing biomass of soil microalgae is lower than that of their predators. By analogy to aquatic ecosystems, this can be explained by the faster turnover of phototrophs. Indeed, most potential algal predators are large protists and can therefore be expected to have relatively longer generation times.

Trophic relationships inferred from correlative analysis of metabarcoding data need to be further explored, possibly with new statistical tools and datasets including other climatic zones and soil types. Nevertheless, we argue that what is now most needed is to characterise the many unknown OTUs, and conducting good observations and experimentation on these organisms to provide useful natural history background needed for sound interpretation of HTS data. As suggested by our study, we believe that future studies providing exact identities of the huge amount of unknown OTUs and revealing their life styles and ecology will provide sound interpretation of the ever-increasing massive sequencing data.

Acknowledgements

We would like to thank Kathleen Hasler for the field work, and Marion Quartier and Christophe Paul for the laboratory analysis. We would also want to thank the BDM program for providing basic data about the sites. This study was partly funded by Swiss National Science Foundation projects no. 310003A 143960 to EL. LB is supported by the European Union's Seventh Framework Programme under grant agreement 245268 (ISEFOR). The authors declare no conflict of interest.

Supplementary materials

Figure S4.1: Relative abundance of micro-eukaryotic sequences in forest, meadow, and cropland soils in Switzerland. Only taxa representing at least 1% of the total OTU abundance for any given land-use are represented.

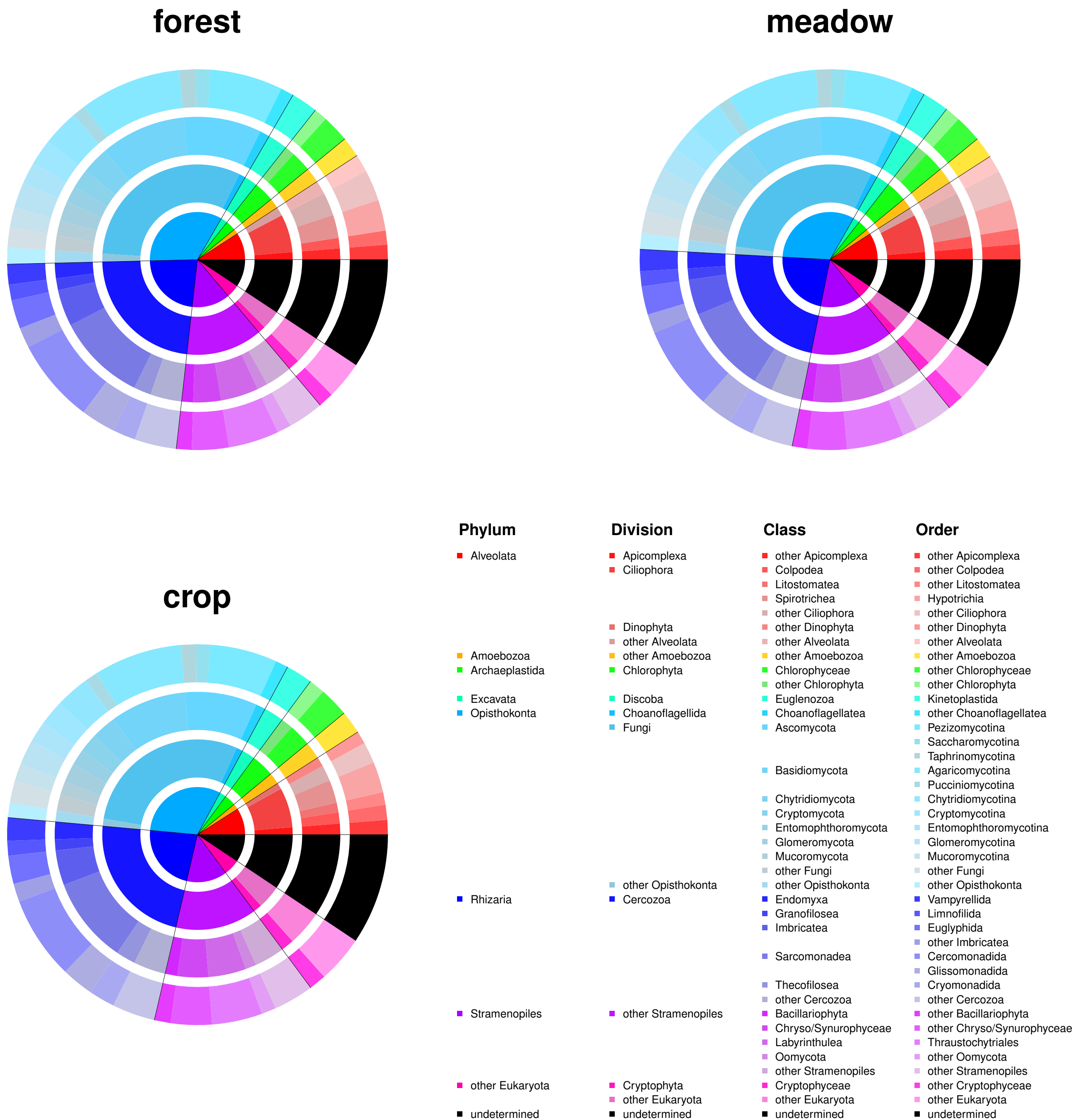


Figure S4.2: Relative abundance of micro-eukaryotic OTU richness in forest, meadow, and cropland soils in Switzerland. Only taxa representing at least 1% of the total OTU number for a land-use are represented.

Table S4.1: Land-use, location (municipality), geographical coordinates and altitude of the 44 sampling sites.

Land-use	Location	Latitude (N)	Longitude (E)	Altitude (m a.s.l.)
forest	Zielebach	47° 08' 56" N	07° 35' 02" E	458
	Wassen	46° 42' 40" N	08° 36' 09" E	481
	Muggio	45° 54' 51" N	09° 23' 59" E	489
	Limpach	47° 06' 47" N	07° 30' 17" E	505
	Siseled	47° 02' 27" N	07° 11' 20" E	548
	Seleute	47° 21' 52" N	07° 06' 27" E	580
	Moosseedorf	47° 00' 18" N	07° 30' 16" E	600
	Les Bois	47° 11' 03" N	06° 57' 49" E	772
	Berolle	46° 34' 04" N	06° 19' 48" E	799
	Savagnier	47° 02' 23" N	06° 57' 07" E	877
	Oberschrott	46° 43' 01" N	07° 16' 07" E	1195
	Saint-Cergue	46° 27' 27" N	06° 05' 52" E	1262
	Fontaines	47° 04' 32" N	06° 52' 20" E	1308
	Vérossaz	46° 12' 44" N	06° 57' 39" E	1560
	Orsières	46° 01' 59" N	07° 06' 57" E	1573
	Ried-Brig	46° 17' 00" N	08° 02' 54" E	1735
meadow	Avenches	46° 53' 47" N	07° 01' 55" E	380
	Saint-Léonard	46° 20' 54" N	07° 25' 32" E	504
	Schüpfen	47° 02' 27" N	07° 20' 48" E	569
	Develier	47° 21' 53" N	07° 15' 60" E	582
	Essertines-sur-Yverdon	46° 42' 52" N	06° 38' 27" E	618
	Châtonnaye	46° 45' 08" N	06° 57' 16" E	640
	Rebeuvelier	47° 19' 44" N	07° 25' 32" E	688
	Le Bémont	47° 15' 22" N	07° 01' 45" E	971
	Birgisch	46° 19' 13" N	07° 58' 15" E	978
	Le Noirmont	46° 57' 33" N	06° 57' 09" E	1014
	La Sagne	47° 02' 21" N	06° 47' 38" E	1077
	L'Abbaye	46° 38' 23" N	06° 19' 43" E	1113
	Chaux-du-milieu	47° 00' 10" N	06° 42' 56" E	1180
	Les Verrières	46° 53' 21" N	06° 28' 30" E	1190
	Airolo	46° 31' 53" N	08° 35' 55" E	1262
	Provence	46° 55' 51" N	06° 42' 59" E	1344
crop	Giubiasco	46° 20' 49" N	08° 59' 04" E	195
	Kappelen	47° 04' 37" N	07° 16' 03" E	431
	Donneloye	46° 45' 03" N	06° 43' 08" E	482
	Aeschi	47° 11' 05" N	07° 39' 46" E	502
	Rapperwil	47° 04' 38" N	07° 25' 32" E	552
	Düdingen	46° 51' 39" N	07° 11' 22" E	575
	Prez-vers-Noréaz	46° 47' 19" N	07° 01' 57" E	642
	Bern	46° 55' 59" N	07° 20' 48" E	647
	Murist	46° 47' 14" N	06° 47' 49" E	681
	Pierra fortscha	46° 51' 39" N	07° 03' 32" E	753
	Cortébert	46° 11' 04" N	07° 06' 32" E	777
	Montana	46° 17' 07" N	07° 30' 13" E	780

Table S4.2: List of micro-eukaryotic taxa identified based on the high-throughput sequencing of 44 soil samples from forest, meadow, and cropland soils in Switzerland. Taxonomic assignation follows the PR^2 database (Guillou et al., 2013).

Trophic group	Division	Class	Order
Osmotroph	Fungi	Ascomycota	
		Basidiomycota	
		Chytridiomycota	
		Glomeromycota	
		Mucoromycota	
Parasite	Apicomplexa	Apicomplexa_X	Gregarines
	Fungi	Blastoclamidiomycota	
	Mesomycetozoa		
	Cercozoa	Phytomyxea	
	Stramenopiles	Oomycota	
Phototroph	Chlorophyta	Chlorophyceae	
		Trebouxiophyceae	
		Ulvophyceae	
	Streptophyta (without Embryophyceae)		
	Stramenopiles	Bacillariophyta	
		Xanthophyceae	
Phagotroph	Ciliophora	Colpodea	
		Heterotrichea	
		Litostomatea	
		Oligohymenophorea	
		Phyllopharyngea	Cyrtophoria
		Spirotrichea	
	Conosa	Mycetozoa-Dictostelea	Dictyosteliida
		Mycetozoa-Myxogastrea	
		Variosea	
	Lobosa	Discosea-Flabellinia	Dactylopodida
		Discosea-Longamoebia	Thecamoebida
		Tubulinea	
	Discoba	Euglenozoa	Kinetoplastida
		Heterolobosea	
	Opisthokonta	Choanoflagellida	
	Cercozoa	Endomyxa	Vampyrellida
		Filosa-Granofilosea	
		Filosa-Imbricatea	Euglyphida
			Thaumatomonadida
		Filosa-Sarcomonadea	Cercomonadida
			Glissomonadida
		Filosa-Thecofilosea	
	Stramenopiles	Bicoecea	
		Chrysophyceae	
		Synurophyceae	
		Labyrinthulea	

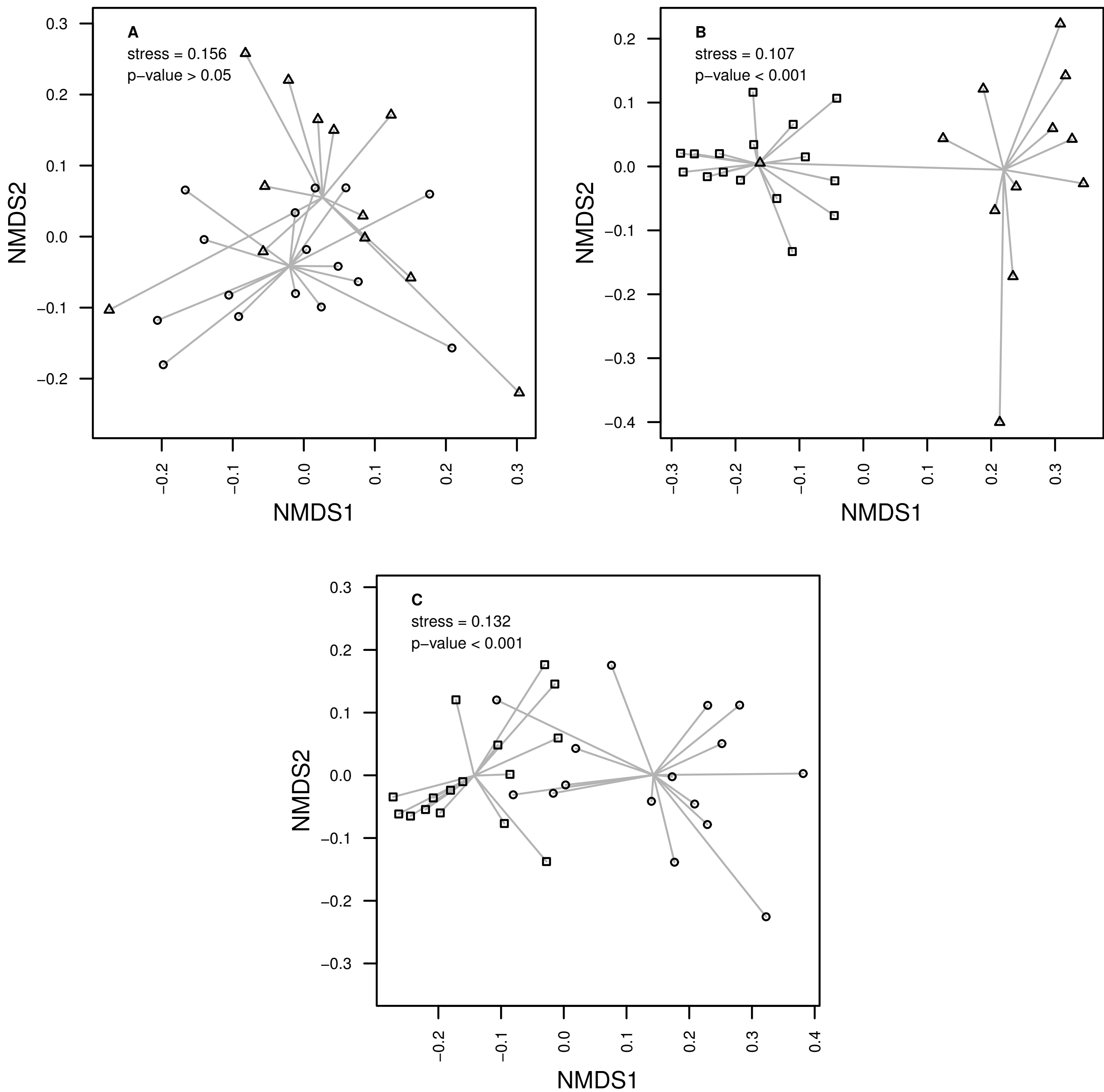


Figure S4.3: Non-metric multidimensional scaling (NMDS) ordinations based on Bray-Curtis dissimilarities for each pair of land-uses (meadow-cropland: A, forest-cropland: B and forest-meadow: C). The land-use types are denoted by squares (forests), circles (meadows), and triangles (croplands). P-values from the permutation tests show the probability that the two environments have the same communities.

Table S4.3: Taxonomic assignation and abundance of eight OTUs which were correlate with total abundance of phototroph OTUs. Column 1 : high taxonomic level given by the ggsearch assignation on the PR² database. Columns 2-4 : finer taxonomy retrieved on GenBank. The top seven OTUs respect residual normality and homoscedasticity. The asterisk identifies an OTU which is known as exclusive phototroph consumer and is among the ten dominant phagotroph OTU but does not respect the condition of homoscedasticity of the linear model. Environmental sequences corresponding to these eight OTUs are given below the table.

	Taxonomic information				OTUs importance in the dataset	
	PR²	GenBank verification	GenBank identifiant	Percent identity (%)	Number of sequences	Proportion of phagotroph (%)
X2	Filosa-Thecofilosea	*Rhogostoma* sp.	HQ121460.1	99	123696	18
X12	Chrysophyceae-Synurophyceae	*Spumella*-like sp.	AB585967.1	100	45073	6.4
X117	Euglyphida	*Trinema* sp.	AJ418792.1	97	6322	0.9
X64	Vampyrellida	*Platyreta* sp.	AY941201.1	99	3984	0.57
X54	Glissomonadida	Allapsidae	AM114807.1	100	3766	0.53
X321	Cyrtophoria	*Pseudochilodonopsis* sp.	KR611083.2	97	2499	0.35
X343	Labyrinthulea	*Sorodiplophrys* sp.	KU728176.1	82	1112	0.16
*X34	Glissomonadida	*Viridiraptor* sp.	KF207869.1	95	11493	1.63

```
>X2
CGCCCGTCGCTACTACCGATTGAATGGCTTAGTGAGCTCTCTGGACTGTTGCGTGTTAGGGGCAACTTTGACACGCAAACGGGAAGGAGATCAAACTTGATCATTTAGAGGAAGTAAAAGTCGTAACAAGGTTTCC
>X12
CGCCCGTCGCACCTACCGATTGAATGATTCGGTGAAACTTTCGGACTGTGGTCTGGACGCCTCACGGCGACTAGATCGTAGGAAGTTATTTAAACCTCATCATTTAGAGGAAGGTGAAGTCGTAACAAGGTTTCC
>X117
CGCCCGTCGCTACTACCGATTGGATGGTTTAGTGAGGTCCACGGATTGTTGCTCTTGGCCGGTTCTCCGGTCCTGCGTATACGAGAAGTGGGACAAACTTGATCATCTAGAGGAAGTAAAAGTCGTAACAAGGTTTCC
>X64
CGCCCGTCGCTCCTACCGATTGAATGTTCCGGTGAATCCTTCGGACCTATTTCTAGAGGTGGGGAAACTCATTTTTAGAAGATTGGGAAGTTGTGTAAACCTTAACATTTAGAGGAAGGAGAAGTCGTAACAAGGTTTCC
>X54
CGCCCGTCGCTACTACCGATTGAATGGATTAGTGAGCTTCAGGGATTGTGGCCTCGTTGTGGCAACACAATGCTGTGTCGCGAGAACTGAATCAAACTTGCTCATTTAGAGGAAGTAAAAGTCGTAACAAGGTCTCC
>X321
CGCCCGTCGCTCCTACCGATTGAGTGATCCGGTGAACTTTCTGGACTCCTGGTTAACGCCGAGGGGGAAGTTAAGTAAACCTTATCACTTAGAGGAAGGAGAAGTCGTAACAAGGTTTCC
>X343
CGCCCGTCGCACCTACCGATTGGACGATCCGGCAAGATGTTTGGAGATGGCAATTTTGAGCTTGCTTAAAATTGTTGTCAAAGCTCCTCAAACCTTATCGTCTAGAGGAAGGTGAAGTCGTAACAAGGTCTCC
>X34
CGCCCGTCGCTACTACCGATTGAATGAATTAGTGAGCTCCAGAGATCGAGCTGTTTCGGGCAACCGGGGCAGATTGAGAACTGGATCAAACTTGCTCATTTAGAGGAAGTAAAAGTCGTAACAAGGTTTCC
```

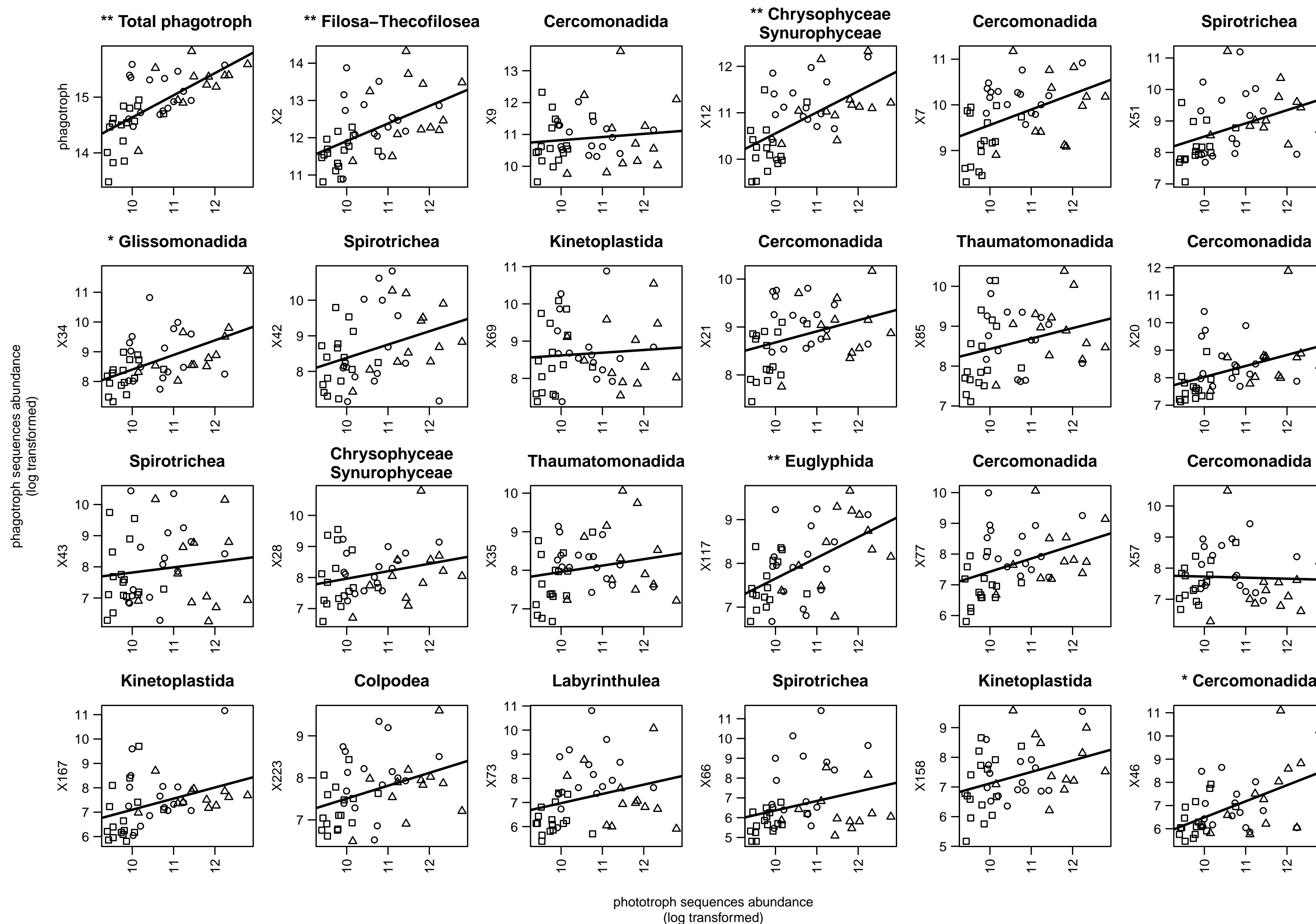


Figure S4.4: Biplots of the total abundance of phagotroph OTUs and the abundance of each of the 100 most abundant phagotroph OTUs versus the total abundance of phototroph sequences in forest, meadow, and cropland soils in Switzerland. For each biplot, the title shows the high-level taxon to which the OTU belongs. OTUs which show a positive and significant correlation with phototroph abundance and respect the two conditions of the linear model are indicated with two asterisks. OTUs which show a positive and significant correlation with total phototroph OTUs abundance but shows a heteroscedastic distribution are indicated with one asterisk. The land-use types are represented by squares (forests), circles (meadows) and triangles (croplands). Linear regressions on the pooled data (three land-use types) are shown for each biplot.

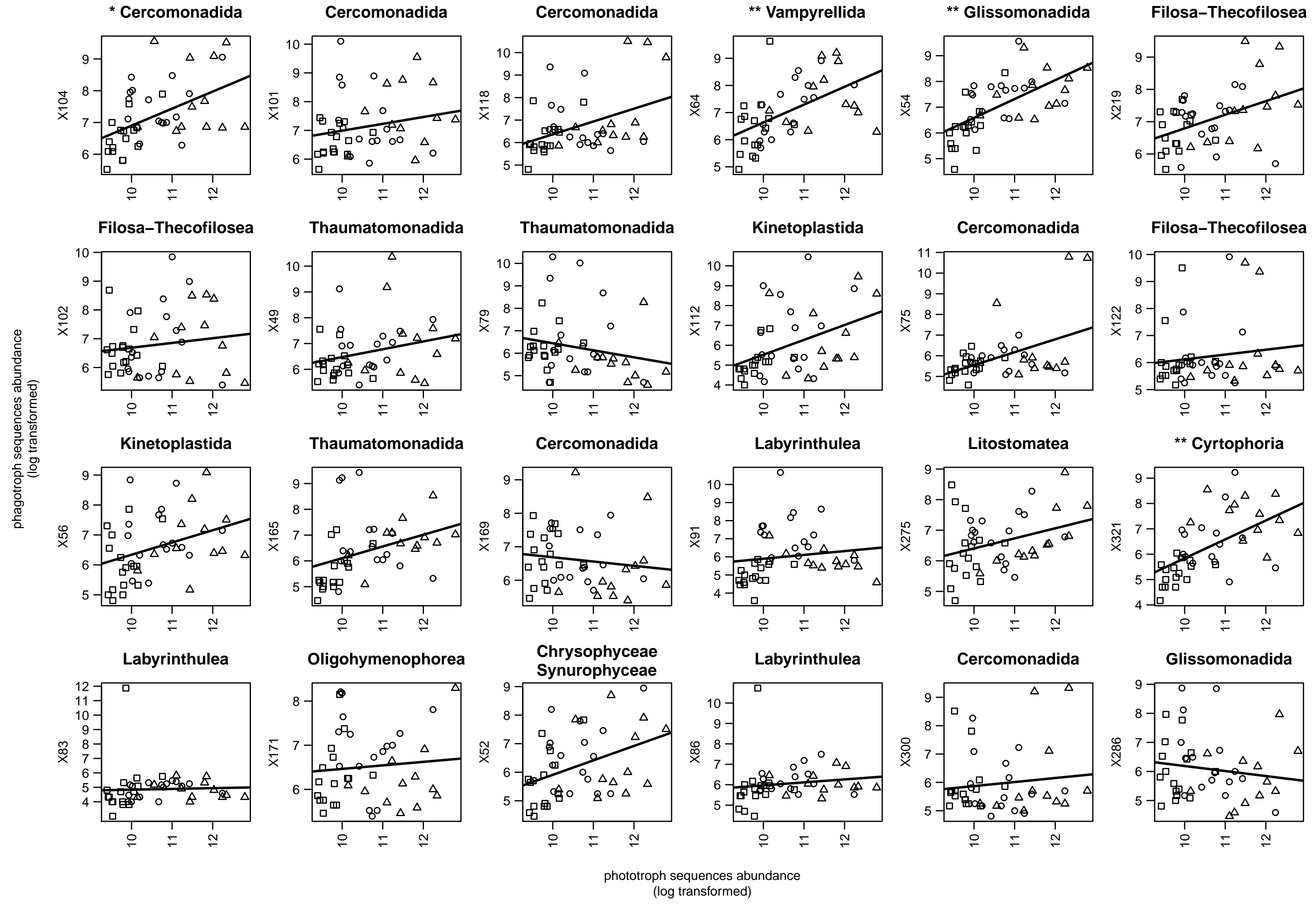


Figure S4.4: (continuation)

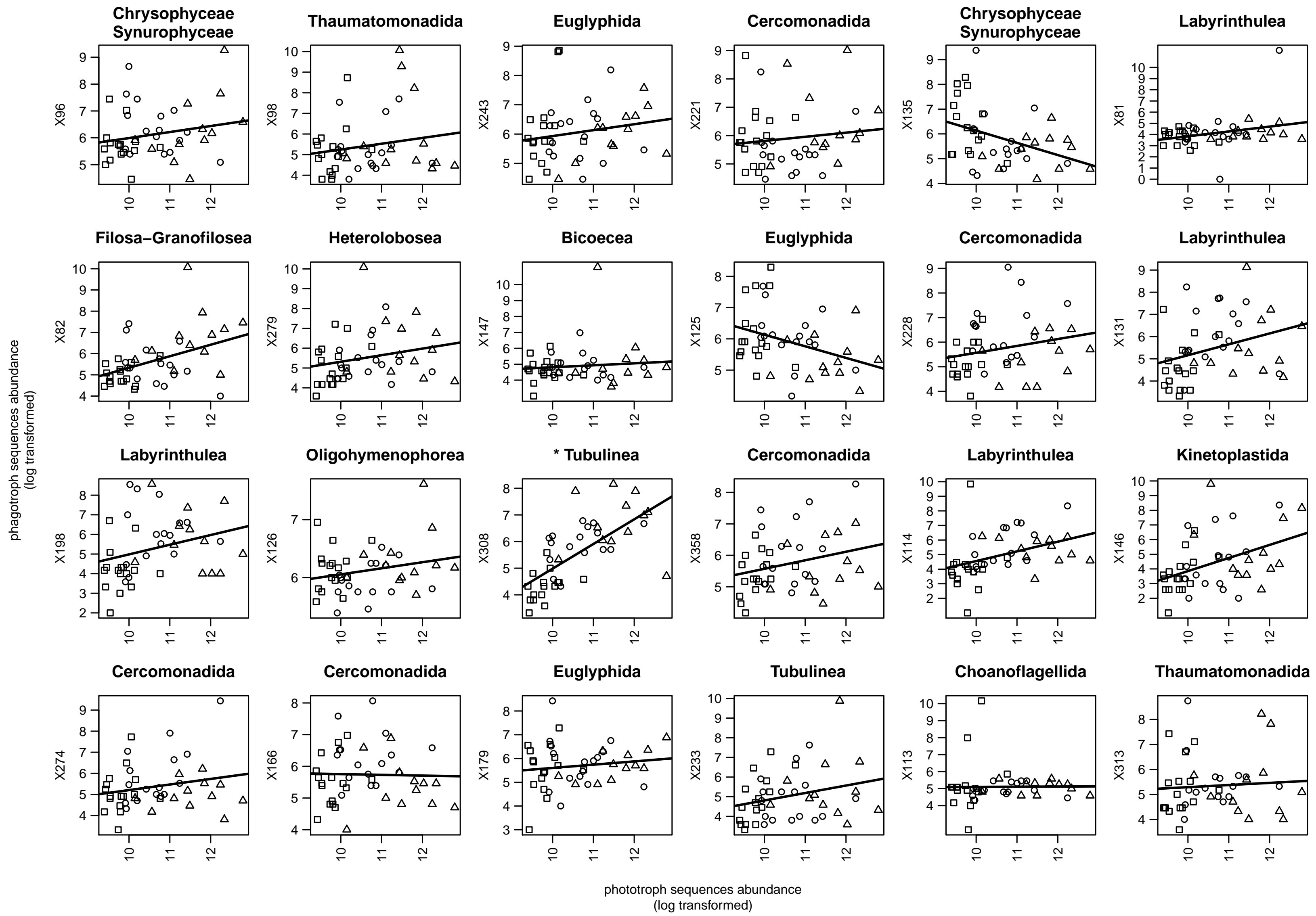


Figure S4.4: (continuation)

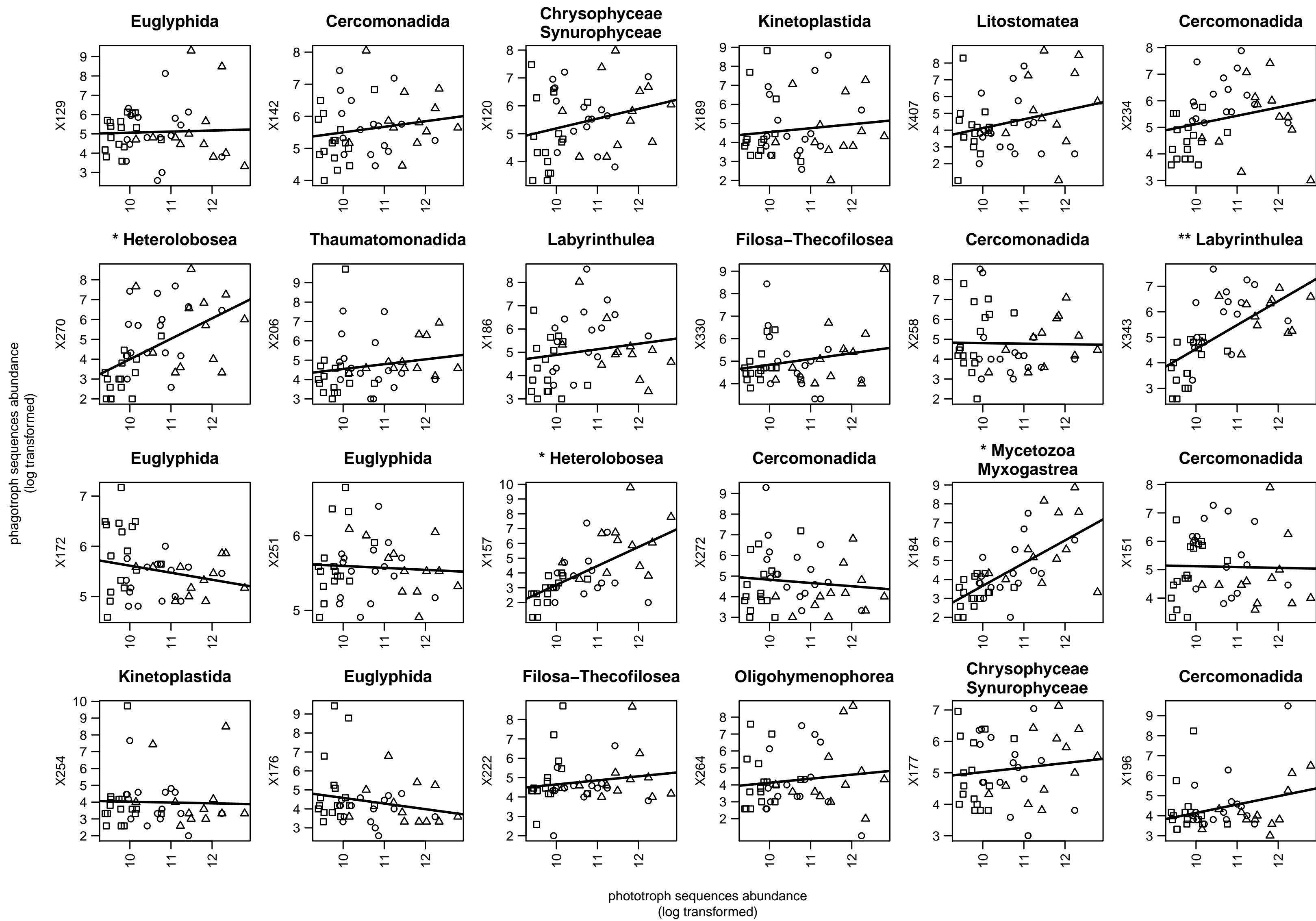


Figure S4.4: (continuation)

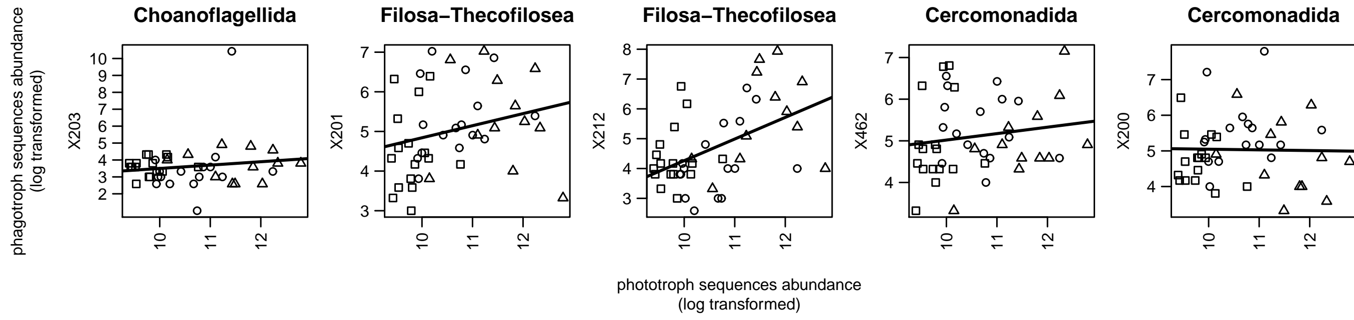


Figure S4.4: (continuation)

Table S4.4: Summary results of linear models calculated on the total abundance of phagotrophs and the abundance of the 100 most abundant phagotroph OTUs according versus the total abundance of phototroph OTUs in forest, meadow, and cropland soils in Switzerland. Both phagotroph and phototroph abundance were log transformed prior to the analysis. Results are shown for two sets of analyses, either not taking into account the interaction between land-uses and abundance of phototroph or including this interaction. All p-values (phototroph and land-use effect, and interaction) were adjusted according to Holm (1979). The residuals normality and homoscedasticity are also shown for every model. The last column show the coefficient of determination of the models calculated only against total phototroph abundance.

	linear models without interaction				linear models with interaction					R^2 of the general model
	phototroph effect	land-use effect	Shapiro test	ncv test	phototroph effect	land-use effect	interaction	Shapiro test	ncv test	
phagotroph	*1.7e-07	5.3e-02	*9.5e-01	*1.3e-01	*1.5e-07	*4.9e-02	2.0e-01	*4.1e-01	*2.2e-01	4.6e-01
X2	*2.2e-02	1	*5.6e-01	*8.0e-02	*2.8e-02	1	1	*7.9e-01	*5.3e-02	2.6e-01
X9	1	1	1.7e-02	*8.2e-01	1	1	1	4.9e-03	*6.2e-01	1.3e-02
X12	*9.6e-04	1	*2.1e-01	*9.6e-01	*1.5e-03	1	1	*1.9e-01	*1	3.5e-01
X7	6.9e-02	1.9e-01	*4.7e-01	*3.3e-01	8.4e-02	2.3e-01	1	*4.1e-01	*2.6e-01	1.9e-01
X51	4.7e-01	1	3.0e-02	*2.4e-01	5.1e-01	1	1	4.2e-02	*1.8e-01	1.6e-01
X34	*2.9e-02	1	*5.6e-02	2.2e-03	*2.0e-02	1	1	*1.1e-01	2.9e-02	2.7e-01
X42	1	1	*6.2e-02	*1.8e-01	1	1	1	3.1e-02	*4.0e-01	1.2e-01
X69	1	1	2.7e-02	*9.9e-01	1	1	1	4.9e-02	*7.1e-01	5.7e-03
X21	1	1	*1.4e-01	*9.7e-01	9.2e-01	1	1	*2.7e-01	*9.3e-01	1.2e-01
X85	1	1	*2.9e-01	*7.1e-01	1	1	1	*3.3e-01	*2.4e-01	8.1e-02
X20	4.5e-01	1	2.0e-05	1.3e-02	3.5e-01	1	1	1.2e-04	7.4e-04	1.6e-01
X43	1	1	1.8e-02	*9.2e-01	1	1	1	1.3e-02	*7.5e-01	1.7e-02
X28	1	1	3.6e-03	*2.5e-01	1	1	1	7.3e-03	*5.6e-01	6.9e-02
X35	1	1	*3.9e-01	*8.6e-01	1	1	1	*3.3e-01	*8.1e-01	4.0e-02
X117	*2.1e-02	1	*1	*3.0e-01	*2.9e-02	1	1	*1	*2.5e-01	2.9e-01
X77	1.7e-01	1.3e-01	4.3e-02	*3.2e-01	1.9e-01	1.5e-01	1	8.8e-03	*5.3e-01	1.6e-01
X57	1	1	2.0e-03	*4.2e-01	1	1	1	3.0e-03	*9.5e-01	1.4e-03
X167	6.7e-01	1	1.3e-04	*5.8e-01	6.4e-01	1	1	3.7e-04	*4.1e-01	1.5e-01
X223	6.7e-01	1	*7.6e-01	*1.3e-01	7.5e-01	1	1	*7.7e-01	*8.3e-02	1.4e-01
X73	1	*2.9e-02	3.2e-02	4.4e-02	1	*3.2e-02	1	8.8e-03	*9.9e-02	7.8e-02
X66	1	*2.0e-02	*3.0e-01	8.4e-04	1	*1.9e-02	1	*5.1e-02	1.2e-03	8.4e-02
X158	1	1	*2.8e-01	*9.7e-01	1	1	1	*2.5e-01	*7.2e-01	1.3e-01
X46	*3.6e-02	1	*9.8e-02	9.8e-04	*1.5e-02	1	1	*1.2e-01	5.2e-04	2.7e-01
X104	*4.3e-02	1	*2.1e-01	2.2e-02	*4.1e-02	1	1	*1.6e-01	1.2e-02	2.4e-01
X101	1	1	*2.3e-01	5.3e-03	1	1	1	*4.9e-01	3.2e-03	4.4e-02
X118	5.7e-01	1	9.5e-05	3.0e-03	2.0e-01	1	2.0e-01	8.7e-04	1.5e-03	1.6e-01
X64	*2.6e-02	1	*1.8e-01	*3.4e-01	*2.0e-02	1	1	2.6e-02	*6.4e-01	2.8e-01
X54	*2.2e-04	1	*7.2e-02	*5.2e-01	*1.4e-04	1	1	4.0e-02	*4.2e-01	4.0e-01
X219	3.4e-01	1	*1	1.8e-02	2.9e-01	1	1	*9.4e-01	4.9e-02	1.9e-01
X102	1	1	*1.1e-01	9.1e-03	1	1	1	*1.4e-01	1.3e-03	1.9e-02
X49	1	1	4.4e-03	4.9e-03	1	1	1	3.7e-03	6.6e-03	7.0e-02
X79	1	1	5.6e-04	9.8e-03	1	1	1	1.6e-03	1.8e-03	4.4e-02

Table S4.4: (continuation)

	linear model without interaction				linear model with interaction					R^2 of the general model
	phototroph effect	land-use effect	Shapiro test	ncv test	phototroph effect	land-use effect	interaction	Shapiro test	ncv test	
X112	3.5e-01	1	2.9e-02	*5.3e-02	4.1e-01	1	1	4.6e-02	2.3e-02	1.7e-01
X75	1.7e-01	1	5.3e-05	1.2e-11	1.5e-01	1	1	2.3e-04	1.3e-09	2.1e-01
X122	1	1	1.6e-08	*9.9e-02	1	1	1	1.3e-08	*1.1e-01	1.9e-02
X56	1	1	*5.3e-02	*8.9e-01	1	1	1	*6.5e-02	*8.6e-01	1.2e-01
X165	9.4e-01	1	6.8e-03	*1.5e-01	5.7e-01	1	1	1.2e-02	2.1e-03	1.3e-01
X169	1	1	4.7e-03	*1.6e-01	1	1	1	2.0e-03	*7.0e-02	1.7e-02
X91	1	*8.8e-05	2.1e-02	2.9e-04	1	*7.7e-05	1	2.7e-03	4.9e-05	2.1e-02
X275	1	1	*1.8e-01	*4.2e-01	1	1	1	*3.3e-01	*8.7e-01	1.1e-01
X321	*4.1e-04	*3.7e-02	*7.2e-02	*1.6e-01	*4.3e-04	*3.7e-02	1	*9.9e-02	*1.1e-01	3.2e-01
X83	1	1	2.4e-09	4.9e-02	1	1	1	8.4e-10	*9.0e-02	1.5e-03
X171	1	1	*1.2e-01	*1.6e-01	1	1	1	*7.5e-01	*6.2e-02	7.8e-03
X52	6.0e-01	1	*5.9e-02	*2.9e-01	6.1e-01	1	1	4.2e-02	*2.4e-01	1.5e-01
X86	1	1	3.7e-08	5.3e-03	1	1	1	3.6e-08	*6.9e-02	2.0e-02
X300	1	1	1.1e-06	*5.1e-02	1	1	1	1.6e-05	5.5e-03	1.3e-02
X286	1	1	6.4e-03	*6.3e-01	1	1	1	*5.2e-02	*1.6e-01	2.1e-02
X96	1	1	2.6e-02	*1.0e-01	1	1	1	*3.4e-01	1.2e-02	4.2e-02
X98	1	1	4.1e-05	8.0e-03	1	1	1	3.2e-05	5.2e-03	3.1e-02
X243	1	1	*6.3e-02	9.3e-03	1	1	1	*9.8e-02	9.9e-04	3.4e-02
X221	1	1	4.3e-04	*2.6e-01	1	1	1	3.0e-04	*3.0e-01	1.4e-02
X135	6.8e-01	1	2.5e-02	*1.6e-01	7.0e-01	1	1	*7.0e-02	4.3e-02	1.5e-01
X81	1	1	8.1e-07	8.5e-14	1	1	1	1.0e-04	1.8e-12	7.9e-02
X82	1.3e-01	6.0e-01	2.3e-02	7.4e-04	9.5e-02	4.4e-01	1	3.6e-03	4.2e-04	1.8e-01
X279	1	1	1.6e-02	3.4e-02	1	1	1	*7.6e-02	8.7e-04	5.4e-02
X147	1	1	2.5e-07	1.2e-06	1	1	1	3.2e-07	2.9e-07	9.0e-03
X125	1	1	*1.1e-01	*3.7e-01	1	1	1	*9.0e-02	*2.1e-01	1.3e-01
X228	1	1.1e-01	*4.5e-01	*6.7e-02	1	1.5e-01	1	*4.7e-01	*6.4e-02	5.4e-02
X131	1	1	*1.2e-01	*3.7e-01	1	1	1	*1.8e-01	*4.5e-01	1.0e-01
X198	1	7.8e-01	*1.0e-01	*1.9e-01	1	9.0e-01	1	*5.2e-02	*3.2e-01	8.2e-02
X126	1	1	*5.5e-02	2.8e-02	1	1	1	4.1e-02	4.2e-02	5.4e-02
X308	*8.4e-07	*2.1e-02	*3.2e-01	1.1e-02	*1.1e-06	*2.3e-02	1	*7.6e-01	2.4e-02	4.6e-01
X358	1	1	4.2e-02	*6.8e-02	1	1	1	*5.7e-02	*1.3e-01	7.3e-02
X114	6.0e-01	1	8.9e-04	*9.7e-02	4.4e-01	1	1	3.5e-05	*5.1e-02	1.5e-01
X146	4.3e-01	1	4.1e-02	*1.9e-01	3.7e-01	1	1	*1.8e-01	4.9e-02	1.8e-01

Table S4.4: (continuation)

	linear model without interaction				linear model with interaction					R^2 of the general model
	phototroph effect	land-use effect	Shapiro test	ncv test	phototroph effect	land-use effect	interaction	Shapiro test	ncv test	
X274	1	6.9e-01	*1.5e-01	1.9e-02	1	4.7e-01	1	*1.7e-01	*3.8e-01	4.5e-02
X166	1	6.1e-01	*6.8e-02	*8.7e-01	1	7.3e-01	1	*2.3e-01	*6.1e-01	6.3e-04
X179	1	1	*5.9e-01	*3.5e-01	1	1	1	*6.6e-01	*8.0e-01	1.8e-02
X233	1	1	1.5e-02	1.0e-02	1	1	1	1.2e-03	*1.6e-01	6.8e-02
X113	1	1	6.3e-08	1.6e-06	1	1	1	2.2e-07	4.6e-08	8.4e-05
X313	1	1	2.4e-04	*9.2e-01	1	1	1	4.8e-04	*2.7e-01	4.7e-03
X129	1	1	1.3e-02	*5.9e-02	1	1	1	8.5e-03	*4.2e-01	1.5e-03
X142	1	1	*1.8e-01	*3.5e-01	1	1	1	*2.9e-01	*2.8e-01	3.2e-02
X120	1	1	*3.0e-01	*6.7e-01	1	1	1	*1.2e-01	*2.7e-01	6.9e-02
X189	1	1	2.6e-04	*7.0e-01	1	1	1	1.4e-04	*6.9e-01	1.3e-02
X407	1	1	*1.2e-01	*1.0e-01	1	1	1	*1.2e-01	*9.0e-02	6.7e-02
X234	1	*4.9e-02	*5.7e-01	*6.4e-01	1	7.0e-02	1	*5.9e-01	*6.8e-01	6.1e-02
X270	*1.0e-02	1	*9.6e-01	3.5e-02	*1.0e-02	1	1	*9.5e-01	1.3e-02	2.7e-01
X206	1	1	6.8e-06	*1.2e-01	1	1	1	2.9e-05	*9.8e-01	2.9e-02
X186	1	1	*3.4e-01	*4.3e-01	1	1	1	*7.3e-02	*4.5e-01	2.7e-02
X330	1	1	2.8e-04	*5.3e-02	1	1	2.5e-01	*1.3e-01	1.1e-03	4.1e-02
X258	1	1	7.0e-03	*3.4e-01	1	1	1	*6.5e-02	*1.7e-01	2.9e-04
X343	*2.4e-05	3.1e-01	*8.5e-01	*4.8e-01	*3.2e-05	3.3e-01	1	*7.7e-01	*3.9e-01	4.2e-01
X172	1	1	*6.9e-01	2.4e-03	1	1	1	*7.1e-01	5.0e-03	5.2e-02
X251	1	1	*4.7e-01	*3.5e-01	1	1	1	*1.1e-01	*6.6e-01	3.6e-03
X157	*1.8e-04	1	*6.2e-02	6.5e-03	*2.0e-04	1	1	1.6e-02	1.1e-02	4.0e-01
X272	1	1	1.1e-02	*3.9e-01	1	1	6.5e-01	4.7e-02	*1.2e-01	1.1e-02
X184	*1.1e-04	1	*5.4e-02	6.5e-05	*1.9e-04	1	1	*8.3e-02	1.3e-04	4.4e-01
X151	1	1	*2.3e-01	*9.1e-01	1	1	1	*6.8e-02	*9.2e-01	5.3e-04
X254	1	1	3.4e-07	*9.3e-01	1	1	1	5.2e-07	*3.3e-01	8.3e-04
X176	1	1	2.4e-05	5.0e-04	1	1	1	4.4e-05	1.8e-04	4.1e-02
X222	1	1	3.2e-04	2.0e-02	1	1	1	2.3e-04	4.0e-04	2.5e-02
X264	1	1	1.7e-03	*8.0e-02	1	1	1	1.0e-03	*2.4e-01	1.6e-02
X177	1	1	*3.3e-01	*9.2e-01	1	1	1	*4.9e-01	*5.4e-01	1.5e-02
X196	1	1	1.6e-04	2.3e-04	1	1	1	3.7e-04	1.5e-03	9.1e-02
X203	1	1	5.2e-08	6.1e-03	1	1	1	2.8e-07	2.6e-08	2.0e-02
X201	1	1	4.7e-02	*9.4e-01	1	1	1	*6.2e-02	*9.9e-01	6.5e-02
X212	*5.0e-02	1	*4.1e-01	1.5e-02	6.7e-02	1	1	*5.1e-01	1.5e-02	2.4e-01
X462	1	1	*6.5e-02	*9.6e-01	1	1	1	*8.5e-02	*4.3e-01	2.4e-02
X200	1	1	*7.2e-02	*3.9e-01	1	1	1	2.9e-02	*5.1e-01	3.5e-04

Chapter 5

Hyper-diverse, easily accessible and least well known: Micro-eukaryotic diversity is higher, more variable and less well identified in soils than in plankton

Christophe V.W. Seppey[1*+], David Singer[1+], Stefan Geisen[2], Alexandre Aebi[1], David Bass[db], Lassaâd Belbahri[1], Betty Benrey[bb], Quentin Blandenier[1], Didier Debroas[dd], Colomban de Vargas[cdv], Isabelle Domaizon[id], Micah Dunthorn[md], Jason Grant[jg], Irina Izaguirre[ii], Gabriela Mataloni[gm], Romina Schiaffino[rs], Edward A.D. Mitchell[1,eadm], Enrique Lara[1]

[1]Laboratory of Soil Biodiversity, University of Neuchâtel, Rue Émile-Argand 11, 2000 Neuchâtel, Switzerland
[2]Department of Terrestrial Ecology, Netherlands Institute of Ecology, Droevendaalsesteeg 10, 6708 Wageningen, Netherlands
[db]Department of Life Sciences, The Natural History Museum, Cromwell Road, London SW7 5BD, United Kingdom
[db]Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Barrack Road, Dorset DT4 8UB, United Kingdom
[bb]Laboratory of Evolutive Entomology, University of Neuchâtel, Rue Émile-Argand 11, 2000 Neuchâtel, Switzerland
[dd]Laboratory of Microorganisms: Genome and Environment, Centre National de Recherche Scientifique (CNRS), avenue des Landais 11, 63177 Aubière, France
[cdv]Roscoff Biological Station, Pierre and Marie Curie University (UPMC) and Centre National de Recherche Scientifique (CNRS), Place Georges Teissier, 29680 Roscoff, France
[id]Alpine Center on Trophic Network and Lake Ecosystems (CARRTEL), Savoie Mont Blanc University and National Institute of Agricultural Research (INRA), Avenue de Corzent 75, 74200 Thonon-les-bains, France
[md]Department of Ecology, University of Kaiserslautern, Erwin-Schroedinger Strasse, 67663 Kaiserslautern, Germany
[jg]Laboratory of Evolutionary Genetics, University of Neuchâtel, Rue Émile-Argand 11, 2000 Neuchâtel, Switzerland
[ii]Departament of Ecology, Genetic and Evolution, University of Buenos Aires, Calle Viamonte 430, 1053 Buenos Aires, Argentina
[gm]Biodiversity, Limnology and Conservation Biology Group, National University of San Martrìn, Calle de Mayo y Francia 25, 1650 San Martìn, Argentina
[rs]North west Research and Transfer Center, National University of the North west of Buenos Aires, Calle Monteagudo 2772, 2700 Buenos Aires, Argentina
[eadm]Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland
[*]**Corresponding author:** christophe.seppey@unine.ch
[+]**Co-first authors**

Abstract: Soils are more heterogeneous habitats when compared to aquatic ecosystems, and water chemistry varies more among lakes and rivers than across the world oceans. As environmental abiotic conditions are the main drivers of biodiversity, alpha and beta microbial diversity should be highest in soils and lowest in the ocean. High throughput sequencing of environmental DNA now makes it possible to assess this.

We analysed the micro-eukaryotic diversity of 163 contrasted samples (59 marine, 37 freshwater and 67 soil) by high-throughput sequencing of the V9 region of the small sub-unit ribosomal RNA. Using the same bioinformatics pipeline for all samples we computed alpha and beta diversity metrics and explored the community patterns among the three categories of ecosystems.

Overall taxonomic richness predicted from species accumulation curves, evenness and beta diversity were highest in soils. From the 56 major eukaryotic taxa studied here, the extrapolated OTU richness was highest in soils (66%), followed by the oceans (25%) and was lowest in freshwater (4%).

Micro-eukaryotic communities were significantly different among the three ecosystem types, the clearest difference being between ocean and continental ecosystems (soil + freshwater). Dominant taxa for each ecosystem were identified at the higher level such as Dinophyta and Radiolaria in oceans, Cryptophyta and Ochrophyta in freshwaters and Fungi in soils.

Our results suggest that micro-eukaryotic diversity is driven by abiotic factors like 1) salinity barrier (oceans vs. continental) or drought (soils vs. aquatic), and 2) habitat heterogeneity and dispersal limitations - explaining the higher variability in soil and freshwater ecosystems. Our diversity estimates suggest that soil micro-eukaryotic diversity likely represents the major component of eukaryotic diversity on Earth and possibly the largest knowledge gap in this global biodiversity.

Keywords: micro-eukaryotic diversity . metabarcoding . V9 fragment ribosomal RNA . alpha diversity . beta diversity . soils . freshwater . oceans

5.1 Introduction

Estimating global diversity, how it compares among broad taxonomic groups, how it varies across continents, biomes and ecosystems and understanding the factors that drive this diversity are among the oldest (Wallace, 1876), but still open questions in biology (Mora et al., 2011; Gaston, 2000; Tittensor et al., 2010; Wall et al., 2001). These questions have remained largely unresolved because of the lack of common methodological approaches for diversity assessment but also due to imbalances in research efforts among ecosystems and taxonomic groups (Geisen et al., 2017). As a result, opinions differed on the taxonomic richness of soils, freshwater and marine habitats (Andre et al., 2002; Brandt et al., 2007), between macroscopic and microscopic organisms, between fungi and protists (Pawlowski et al., 2012) and within microbes between protists and bacteria (Grattepanche et al., 2014).

Although estimates of micro-eukaryotic diversity are still very uncertain, high throughput sequencing studies clearly show that they represent a huge component of global diversity in aquatic ecosystem (de Vargas et al., 2015) as well as in soils (Mahé et al., 2017).

Micro-eukaryotes (protists and Fungi) dominate eukaryotic diversity in all major ecosystems on Earth (Pawlowski et al., 2012), and possibly also overall diversity as shown for ocean (de Vargas et al., 2015). They play essential roles in biogeochemical cycles (Berner and Berner, 2012) including carbon sequestration (Falkowski, 2002; Arrigo, 2005; Landsberg and Sands, 2010), organic matter decomposition (Davidson and Janssens, 2006; Caron and Hutchins, 2012), and enhance nutrient remobilisation by preying on bacteria and other microorganism (Middelburg et al., 1993; Berg, 2000; Bonkowski and Clarholm, 2012; van der Wal et al., 2013). As predators or parasites they apply top-down selection pressure on their preys or hosts (Gonzalez et al., 1990; Hahn and Höle, 2001; Murase et al., 2006; Dumack et al., 2016b; Geisen et al., 2015b,c, 2016). They also constitute an essential prey resource for higher trophic levels (Turner, 2004).

While micro-eukaryotes are key players in all ecosystems it is clear that major ecosystem types strongly differ in the phylogenetic and taxonomic makeup of communities and contribution to ecosystems functions. For example, the dominant components of phytoplankton are Dinophyceae (Alveolata) in oceans and Ochrophyta (Stramenopiles) in fresh waters (Izaguirre et al., 2004). Environmental DNA studies are however revealing the existence of some taxonomic groups in habitats where they were not believed to occur, such as foraminifera in soils (Lejzerowicz et al., 2010) or unsuspected diversity of taxonomic and functional groups such as parasitic Apicomplexa in soils (Mahé et al., 2017). These findings call for a reassessment of the contribution of individual micro-eukaryotic groups to taxonomic and functional diversities across ecosystem types (Grossmann et al., 2016). This effort should also contribute to filling the existing knowledge gap in micro-eukaryotic diversity between aquatic and soil habitats and between protists and other groups (Geisen et al., 2017; Mahé et al., 2017).

A reassessment of micro-eukaryotic diversity across ecosystems, made possible by the development of high throughput sequencing technology, provides the opportunity to explore the role of environmental factors in shaping community composition and determining overall and within-group diversity. The world's major ecosystem types represented by oceans, freshwater and terrestrial ecosystem differ in key factors such as salinity, water availability, temporal stability and the degree of habitat heterogeneity (Chapin III et al., 2011). Each of these factors are known to structure communities and require adaptations that

have evolved to various degrees depending on the phylogenetic group considered (Xu et al., 2008; Geisen et al., 2014a; Jiang and Pu, 2009).

Of these three general types of ecosystems soils are the most heterogeneous in space - at the micro to the macro-scales - and variable temporally - from short term such as a rain event to seasonality. This complexity offers a very high diversity of contrasted conditions potentially allowing a huge diversity of organisms to co-occur - although strictly speaking not all at the exact same place and time (Ettema and Wardle, 2002). By contrast aquatic habitats are generally more homogeneous and for this reason should theoretically be less diverse. However, population dynamics driven by biotic interaction allow a high diversity to occur, and thus explain the paradox of the plankton (Hutchinson, 1961). However, given the current sampling bias favouring aquatic ecosystems, it is not possible to provide any reasonable comparison among ecosystems types. A reassessment of diversity and community patterns among major ecosystem types requires a common methodology which is what we have aimed to do.

We assessed the micro-eukaryotic diversity in soils, marine and freshwater plankton from a worldwide dataset comprising 163 samples (59 marine, 37 freshwater, 67 soils) using high throughput sequencing metabarcoding of the V9 region of the small sub-unit (SSU) rRNA. Our aims were to assess how these three major ecosystem types differed in 1) taxonomic composition of micro-eukaryotic groups, 2) diversity of individual major eukaryotic groups, 3) alpha and beta diversities within samples, 3) community structure and heterogeneity among samples. Considering the likely role of habitat complexity and spatial and temporal heterogeneity we hypothesised that diversity would increase from the ocean to freshwater and to soil habitats.

5.2 Materials and methods

Sampling

Our dataset was composed of a worldwide collection of 163 samples: 59 marine plankton, 37 freshwater plankton and 67 soil samples, hereafter referred to as "ecosystems" (Table S5.1). Soil samples consisted of ca. 500 g of the upper organic horizon. Samples were taken in sterile conditions and kept at low temperature during transport to the laboratory. Freshwater samples (50 litres) were filtered either through a single mesh (0.2 μm) or sequential meshes (20 μm and 0.2 μm). The filtrates were sent to the laboratory at low temperature (-20 °C) for DNA extraction. Marine samples were taken at two depths (surface and Deep-Chlorophyll Maximum: DCM) (`http://taraoceans.sb-roscoff.fr/EukDiv/`). For each sample, 100 to 150'000 litres of marine water were sequentially filtered through meshes of 2000 μm, 180 μm, 20 μm, 5 μm and 0.8 μm, retrieving the 180-2000 μm, 20-180 μm, 5-20 μm and 0.8-5 μm fractions. The filtrates were preserved separately in liquid nitrogen until DNA extraction.

Molecular analysis

Soil and freshwater DNA was extracted with the Mo-Bio PowerSoil extraction kit (Carlsbad, CA, USA) according to the manufacturer instructions. Marine DNA was extracted with the DNA Elution buffer kit (Macherey-Nagel) (`http://taraoceans.sb-roscoff.fr/EukDiv/`). SSU rRNA V9 fragments were amplified as described in Seppey et al. (2017) for the Swiss soil samples, Singer et al. (2016) for freshwater and the other soil samples, and de Vargas et al. (2015) for marine plankton. The amplicons were sequenced with a HiSeq Illumina sequencer for both soils and freshwater, and with a Genome Analyser for marine samples.

Bioinformatic analysis

Database

A dereplicated rRNA SSU V9 fragments reference database (PR2_Silva) of eukaryotic (PR$^{2a;}$ Guillou et al., 2013, downloaded the 14 September 2016) and prokaryotic (Silva; **?**, downloaded the 8 August 2016) sequences was constituted. The eukaryotic sequences were trimmed according the eukaryotic primers use by Amaral-Zettler et al. (2009) (1380f/1510r: `CCCTGCCHTTTGTACACAC` / `CCTTCYGCAGGTTCACCTAC`). Prokaryotic V9 fragments were obtained by trimming the prokaryotic Silva sequences from the universal primers 1389f `TTGTACACACCGCCC` (Amaral-Zettler et al., 2009) to the end of the rRNA SSU.

Pipeline

Environmental rRNA SSU V9 reads were first trimmed to remove the eukaryotic primers. We then removed low quality reads defined as those containing a 50 nucleotides window with an average phred score below 20. Chimera were removed by comparing the reads within each sample and against the PR2_Silva database. Sequences belonging to Metazoa, Embryophyceae of prokaryotes were removed after environmental reads were taxonomically assigned by aligning them to the PR2_Silva database with the program `ggsearch` (v. 36.3.6; Pearson, 2000). Rare reads found less than three times in two samples were removed in order to avoid noise caused by rare sequences (de Vargas et al., 2015). We then randomly selected 50'000 reads from each sample to homogenize the sampling effort. To obtain comparable diversity patterns from each sample taken in many fractions (e.g. marine or freshwater) we randomly selected, for each fraction, a number of reads corresponding to 50'000 divided by the number of fractions and pooled the selected reads into a single 50'000 reads sample. The resulting reads were clustered into OTUs using the program `swarm` (v. 2.1.8; Mahé et al., 2015). Each OTU was taxonomically assigned by aligning the dominant read against the PR2_Silva database with the program `ggsearch` (v. 36.3.6; Pearson, 2000). The OTUs having a percentage of identity lower then 80% with the reference database were

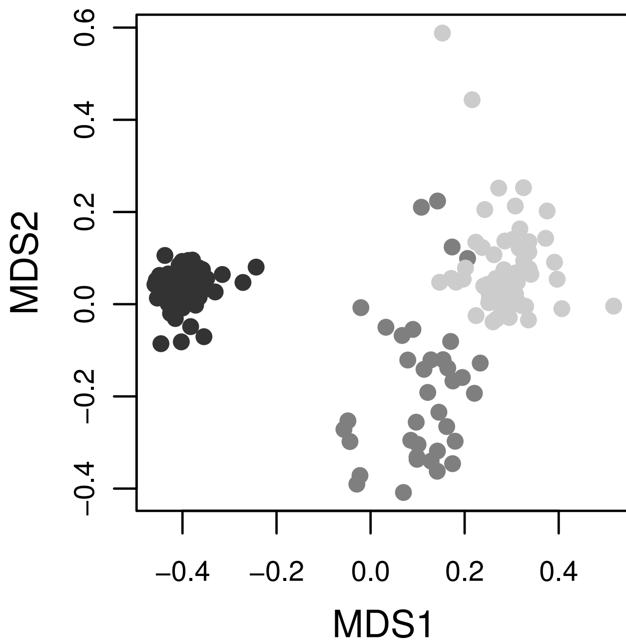


Figure 5.1: Ordination plot (non-metric multidimensional scaling) of OTU micro-eukaryotic communities found in marine water (dark grey), freshwater (grey) and soil (pale grey) samples from around the world.

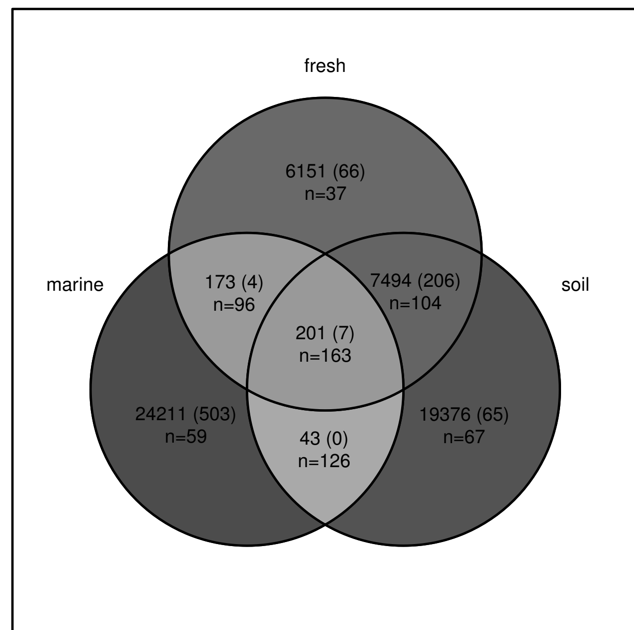


Figure 5.2: Number of micro-eukaryotic OTUs retrieved in marine water, freshwater and soil samples from around the world. The grey shades represent this number normalized by the number of samples. All OTUs were averaged according to the ecosystems. Between parenthesis, the number or OTUs respecting the thresholds of 1) have a minimum ecosystem average corresponding to at least 5% of the ecosystem averages sum and 2) have at least 10 sequences in the full dataset.

considered as undetermined eukaryotes (de Vargas et al., 2015).

Numerical analysis

We assessed the similarity patterns among communities of the 163 samples by non-metric multidimensional scaling (NMDS) and the significance of differences between pairs of ecosystems by permutation tests (10'000 permutations; functions `metaMDS` and `envfit`, package `vegan`; Oksanen et al., 2013).

We compared the number of OTUs present in each ecosystem, or group of ecosystems to measure the specificity of the communities. We specifically identified ubiquitous OTUs (i.e. present in the three ecosystems types) to assess which organisms can potentially cross abiotic constrains like the salinity barrier or live in both soil and aquatic habitat. The ubiquity of each OTU was assessed based on its average abundance in each ecosystem. An OTU was considered as ubiquitous if 1) the lowest average among the three ecosystems was over 5% of the sum of these averages and 2) if the sum of these averages was over 10 sequences on the full dataset. The first threshold avoids considering extremely rare OTU in a given ecosystem (which would not be ecologically relevant in that ecosystem) and the second avoids considering overall rare OTUs (which would not be ecologically relevant globally).

We computed accumulation curves for the total community as well as for each major eukaryotic taxon, and extrapolated the total number of OTUs by bootstrapping for each ecosystem (functions `specaccum` and `specpool`, package `vegan`; Oksanen et al., 2013). For each ecosystem, we assessed the relative sequences abundance and diversity of the major taxa.

We computed α and β diversity metrics for each ecosystem. For α diversity we reported the OTU richness (R), Shannon index (H) and Pielou's evenness (J). For β diversity we calculated the Sorensen β diversity and its turnover and nestedness components (Baselga, 2010). We assessed if the three β diversity metrics differed among ecosystems by bootstrapping 100 times one quarter of the samples of each ecosystem (function `beta.sample`, package `betapart`; Baselga et al., 2013). We then tested for differences among ecosystems for each diversity index and β diversity metric by pairwise tests for multiple comparison of mean rank sums (Nemenyi test, $P < 0.05$; function `posthoc.kruskal.nemenyi.test`, package `PMCMR`; Pohlert, 2014).

5.3 Results and discussion

Data results

We retrieved a total of 450'000'000 quality-checked sequences from the 163 samples. Of these, 99% were not chimera and 53% were not assigned to a macro-organism (Metazoa, Embryophycea) or a prokaryote. From these ca. 240'000'000 sequences over 210'000'000 were found at least three times in at least two samples and were retained for further analysis. The 8'150'000 random sequences (163 X 50'000) resulting from the sample-size normalization were clustered into 57'649 OTUs. Of these, 12% could not be taxonomically assigned at or above the 80% identity threshold with our PR2_Silva database.

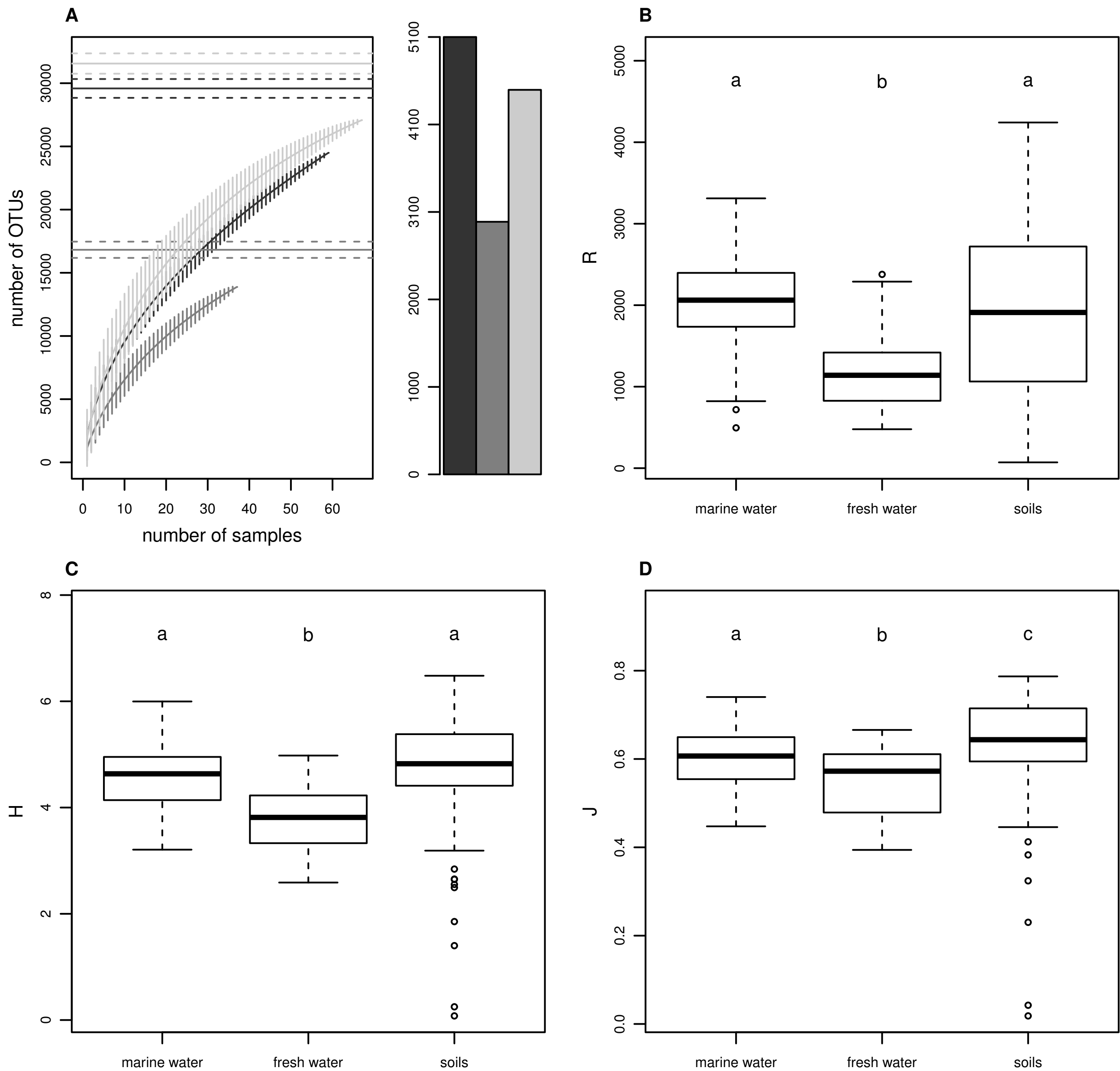


Figure 5.3: Species accumulation curves (A) and ecological metrics (B-D) calculated from micro-eukaryotic OTU communities retrieved from marine (dark grey), freshwater (grey) and soil (pale grey) samples taken around the world. The horizontal plain and dashed lines above the species accumulation curves (A) indicate the predicted number of OTUs (from bootstrap) and standard error associated for each environment respectively. The barplots on the right side (A) show the proportion of OTUs undiscovered (number of OTUs predicted by bootstrapping minus number of OTUs found in the dataset). The ecological metrics represented are the Richness (R:A), Shannon index (H:B) and Pielou's evenness (J:C). Letters above the distributions represent the significantly different groups (Nemenyi test $P < 0.05$).

Community patterns among and within ecosystems

Micro-eukaryotic communities differed significantly among the three ecosystems (Figure 5.1, S5.2). Marine samples were more clustered in the ordination space, reflecting more similar communities as compared to soils or freshwater. This pattern matches the difference in connectivity as well as a possible range of abiotic characteristics among the three ecosystems. Indeed, oceans are highly connected through global marine circulation and abiotic conditions are more homogeneous than lakes and soils. Lakes cover a broader range of conditions, especially water chemistry. Soils are even more heterogeneous both spatially at fine- and broad-scales and temporally (e.g. dry-wet cycles strongly modifying both water availability and soil water chemistry).

General taxonomic and functional composition of communities

The relative proportions of different eukaryotic supergroups in the communities clearly differed among ecosystems (Figure S5.1). Dinophyceae (Alveolata) and to a lower extent, Bacillariophyta (Stramenopiles) dominated marine phototrophs. However, as a relatively large proportion of marine Dinophyceae OTUs corresponded to unknown (31%) or non-phototrophic organisms

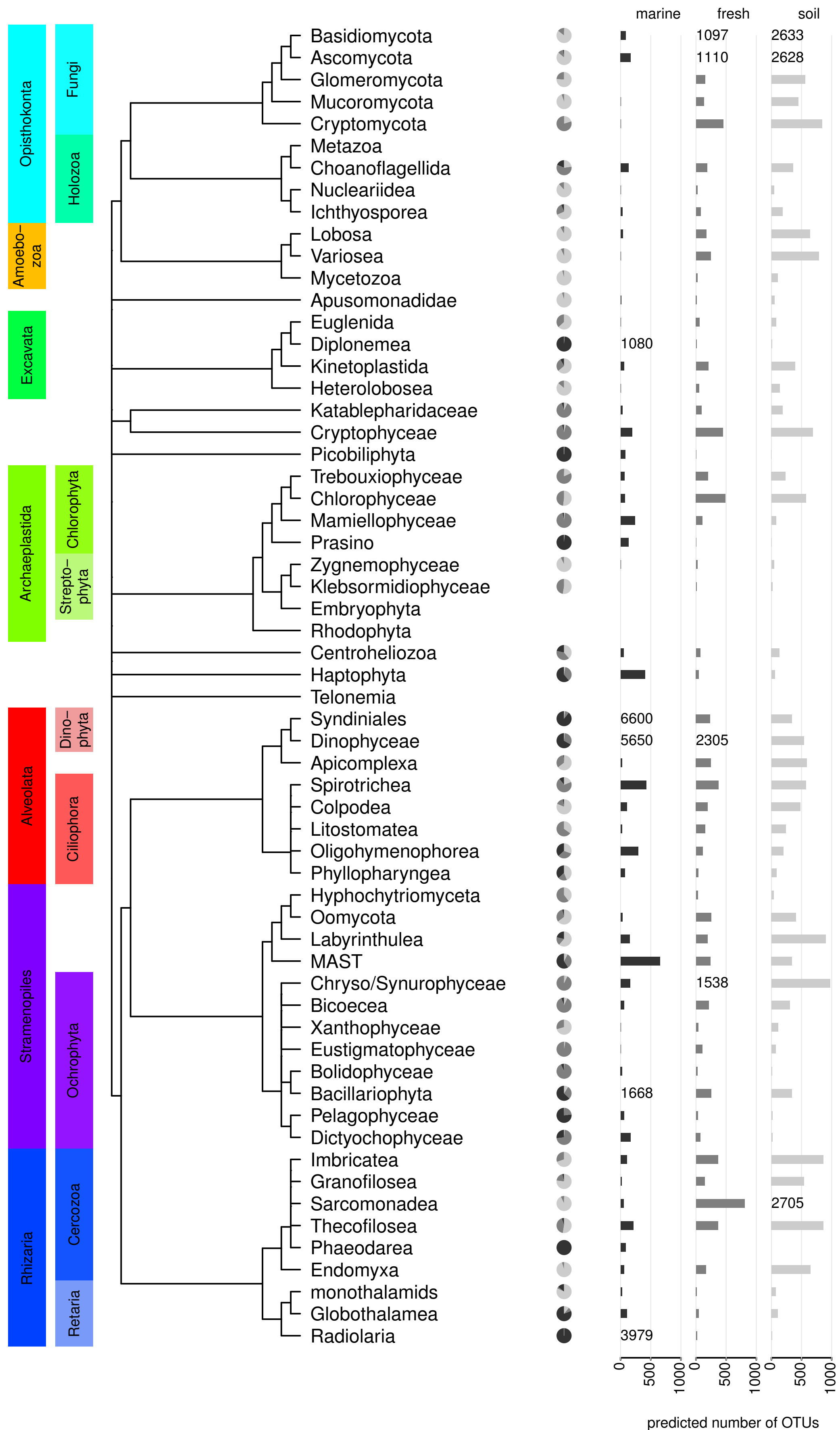


Figure 5.4: Schematic phylogenetic tree of the main micro-eukariotic taxa's relative abundance according to three ecosystems (marine water: dark grey, freshwater: grey, and soil: pale grey). The relative abundances are normalised by the number of samples per ecosystem. Barplots represent the OTU richness of each taxon predicted by bootstraping in each environment. Predictions higher then 1000 OTUs are written numerically.

(e.g. Blastodinium; 6%), it is hard to assess the ratio between phototrophy and other possible functions (i.e. parasitism, phagotrophy) for this taxon and to accurately infer its contribution to each functional group. The dominant phototroph micro-eukaryotes in freshwater and soils were Chryso/Synurophyceae (Stramenopiles), Archaeplastida and Cryptophyceae (freshwater only). Furthermore, the proportion of sequences assigned to phototrophic organisms was lower in soils. However, here too the real proportion of phototrophy cannot be inferred precisely as these dominant taxa contain a significant proportion of OTUs with uncertain taxonomic assignment or corresponding to non-phototrophic organisms (e.g. Goniomonadales; Novarino and Lucas, 1995, *Oikomonas*; Boenigk, 2008).

Phagotrophs were dominated by Radiolaria (Rhizaria) in oceans, and Ciliata (Alveolata) and Cercozoan (Rhizaria) in soils and freshwater samples. Fungi (Opistoknota) were mainly composed by Dicarya (Ascomycota and Basidiomycota) and represent the principal group of saprotrophs in continental ecosystems. OTUs assigned to Dicarya clearly dominated terrestrial ecosystems, accounting for over half of all soil eukaryotic sequences, in agreement with the known abundance of fungi in soils (Seppey et al., 2017).

Parasites mainly consisted of the planktonic Syndiniales (in marine and freshwater samples), and the Apicomplexa (Alveolata) and Oomycota (Stramenopiles) in soils. Sequences assigned unambiguously to parasites were much more abundant in oceans (e.g. Syndiniales: 12%) than in freshwater (e.g. Syndiniales: 1%) or soils (e.g. Apicomplexa + Oomycota: 3%). While micro-eukaryotes were long thought to mainly consist of phagotrophs and phototrophs it is increasingly clear that parasites constitute a major component of micro-eukaryotic diversity and likely play major roles in trophic networks and biogeochemical cycling (Guillou et al., 2008; Mahé et al., 2017).

Salinity and drought barriers

The analysis of the number of OTUs present in each ecosystem, or group of ecosystems, showed a clear abiotic conditions cleavage between marine and continental ecosystems which corresponds to the salinity barrier (Figure 5.2). By contrast, many OTUs were found in both freshwater and soils, suggesting that drought represents a comparatively weaker ecological filter. Only 201 OTUs were found in all three ecosystems, of which seven were considered as ubiquitous and respected the two thresholds of 1) presence in all ecosystems and 2) sufficient abundance in the whole dataset. These seven OTUs included two Kinetoplastida (X957: *Bodo* sp., X1020: *Neobodo* sp.), one Bacillariophyta (X3529: *Navicula* sp.) and four Fungi (X444: Dothideale, X785: Debaryomycetaceae, X1086: Agaricale, X493: *Rhodotorula* sp.). *Bodo* and *Neobodo* were indeed previously reported from both freshwater and marine ecosystems (Koch and Ekelund, 2005; Simpson et al., 2006). The presence of X3529, which matched perfectly *Navicula* sp. (100%), in both marine and terrestrial ecosystems is also in agreement with the literature. Indeed, many species belonging to this genus seem to tolerate a broader range of salinity than other diatoms (Scholz and Liebezeit, 2012a,b). Three of the four fungi OTUs present in the three habitats, were assigned at 100% identity with acknowledged halotolerant organisms (X444: *Hortaea werneckii*; X785: Derbaryomyces sp., *Candida* sp., *Pichia* sp.; X493: *Rhodotorula* sp.) (Gunde-Cimerman and Zalar, 2014). The fourth ubiquitous fungal OTU was assigned with many Agaricales, including many yeasts known to be able to live in oceans (Jennings, 1983; Passarini et al., 2015). It is also noteworthy that a single V9 fragment matched perfectly with many Agaricales species reported from different ecosystems (Seppey et al., 2017).

Major taxa OTU richness and abundances

The total OTU richness inferred from our extrapolation analyses was higher for soils than for aquatic habitats (Figure 5.3). This was also the case for 66% (37/56) of all major taxa (Figure S5.4). This suggests that soils are the most diverse ecosystems on Earth. The richness extrapolation predicted the existence of ca. 5100 more OTUs in the marine plankton and ca. 4800 in soils. However, the number of undiscovered OTUs was higher in soils for 52% (29/56) of all major taxa (Figure S5.4). This difference between the overall and per taxa results can be explained by the high number of predicted undiscovered OTUs in very diverse taxa mostly found in marine plankton (e.g. OTU to discover: Syndiniales 845, Radiolaria 1032). In addition, 52% (29/56) of all major taxa were also more abundant in soils than in aquatic ecosystem (Figure 5.4). Of these, eight taxa (Fungi: Cryptomycota, Basidiomycota, Ascomycota; Cercozoa: Imbricatea, Sarcomonadea, Thecofilosea; Stramenopiles: Labyrinthulea, Oomycota) had more than 100 predicted undiscovered OTUs. Last, it is noteworthy that some taxa thought as mostly aquatic (Choanoflagellida, MAST) were found to be quite abundant and diverse in soils (Figure S5.1).

Diversity indices

α diversity indices (richness and Shannon's H per sample) were higher in soils and marine samples than in freshwater samples and Pielou's evenness (J) was highest in soil, intermediate in marine samples and lowest in freshwater samples (Figure 5.3B-D). The higher evenness in soils could be explained by the more variable micro-environmental conditions in both space and time as compared to aquatic ecosystems. Temporal variability in the scale of minutes (e.g. rain) selects for a high proportion of encysted protists while higher spatial heterogeneity increases the number of potential ecological niches at the microscale.

Sorensen β diversity increased from marine plankton to freshwater to soils (Figure 5.5). This pattern reflected the heterogeneity among samples in each ecosystem. Soil characteristics vary considerably in relation to bedrock, chemistry, texture, hydrology, climate and vegetation.

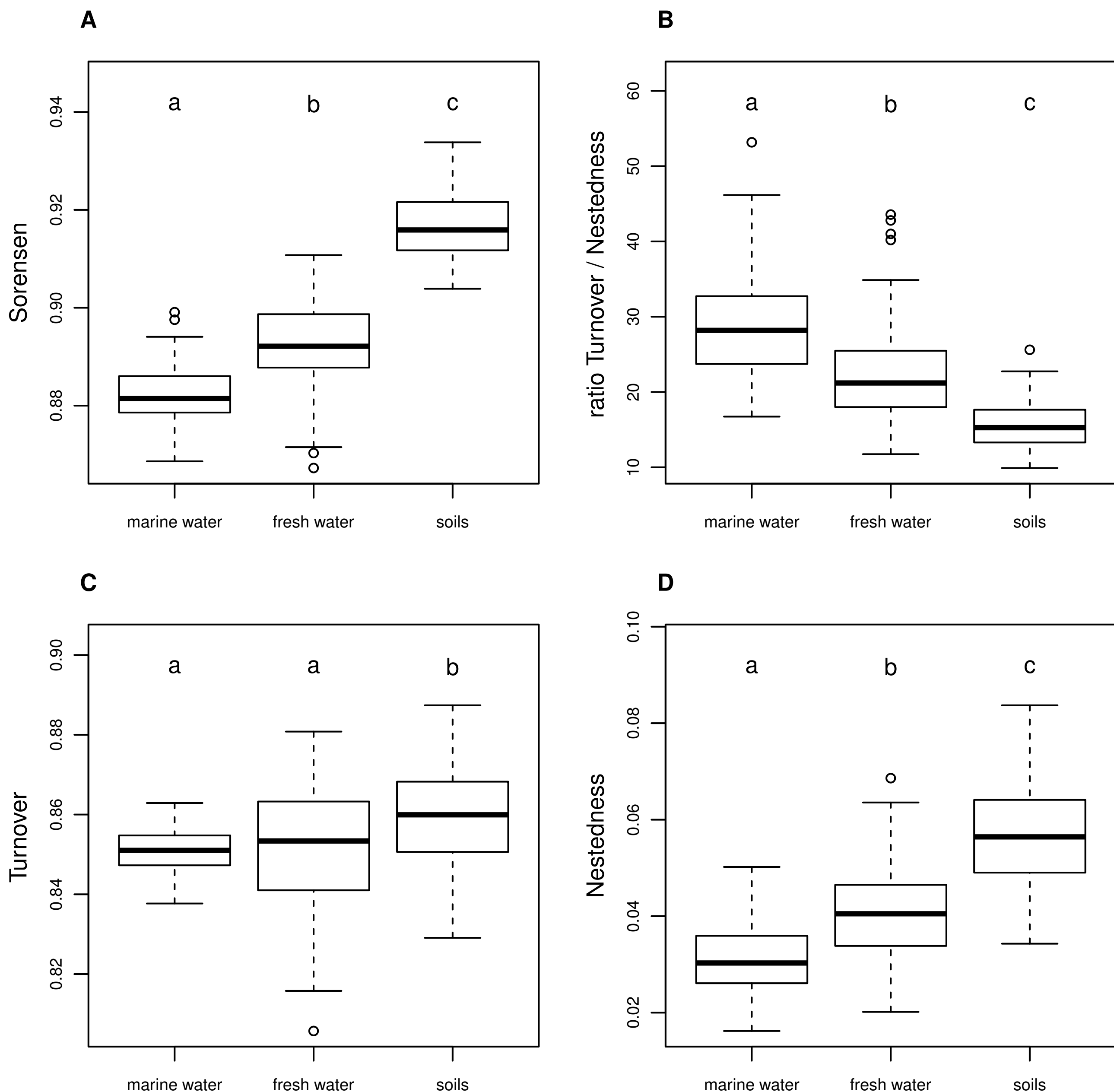


Figure 5.5: Sorensen beta diversity (A) as well as two of its components (Turnover: C, Nestedness: D) calculated from micro-eukaryotic OTU communities retrieved from marine water, freshwater and soils taken around the world. The ratio between Turnover and Nestedness are also shown (B). For each ecosystem, the distribution of beta diversity was retrieved from 100 bootstraps made on one quarter of the samples. Letters above the distributions represent the significantly different groups (Nemenyi test $P < 0.05$).

Freshwater bodies are less variable than soils but more than the ocean. In all ecosystems, turnover constitutes the main component of β diversity (Figure 5.5C-D). This is explained by the fact that the majority of OTUs are only found in one or two samples (Figure S5.3). Soils also show the lowest turnover to nestedness ratio (Figure 5.5B). This signifies that OTUs found in less diverse soils are also found in the richer ones, while OTUs found in just a few samples of planktonic communities are distributed more randomly among the samples. The higher nestedness in soils and to a lesser extent freshwater ecosystems also suggest an influence of dispersal limitation. Indeed, dispersal limitation would be favoured by the patchiness of both habitats (Baselga, 2010). Dispersal could be predicted to being especially low for soil organisms lacking any dispersal mechanism while freshwater plankton can more easily be transported by the water masses moves or passively by wind or birds from one water body to an other.

Conclusion and perspective

First, our results give an other line of evidence that abiotic factors like salinity or substrate type are significant drivers of micro-eukaryotic communities. This difference in communities is also illustrated by the evolutionary rigidity of some major taxa (e.g. Radiolaria, Diplonemea) which have not evolved the capacity to cross the salinity and/or drought barriers. Nevertheless, other taxa have filled the corresponding ecological niches in freshwater and soil habitats. Only few organisms, mostly halotolerant, can be found in the three ecosystem-types.

Second, the predictions made on the OTU communities suggest that there are probably not more OTUs in oceans than in soils. Although there would be more diversity to discover in oceans - mostly because of hyperdiverse taxa (Syndiniales, Radiolaria) - , most of the major groups of micro-eukaryotes are expected to have a higher unknown diversity in soils as compared to aquatic ecosystems.

This meta-analysis contain certain bias in term of geographical coverage and sample size. Nevertheless, these bias support certain of our conclusion. The geographical coverage did not include tropical soils which are likely to be very diverse if micro-eukaryotes show the same diversity pattern then macro-organisms. In term of sample size, if the same amount of soil would had been taken then for the marine water, it is very likely that the diversity would had been increased.

Acknowledgements

The authors would like to thank all the peoples who took part in the different samplings and the laboratory work. This study was partly funded by Swiss National Science Foundation projects no. 310003A 143960 to EL.

Supplementary materials

Table S5.1: Sampling location summary of 163 communities of micro-eukaryotes from marine and freshwater plankton, and soils

Environment	Continent	Country	Samples number	Published in
Marine Water	Mediterranean See	-	13	(de Vargas et al., 2015)
	Red See	-	4	
	South Atlantic	-	8	
	South Pacific	-	17	
	Indian Ocean	-	13	
	Antarctic Ocean	-	4	
Fresh water	Europe	France	21	(Schiaffino et al., 2016) (Lara et al., 2015)
	South America	Argentina	12	
	Asia	Russia	1	
	Antarctic	-	3	(Schiaffino et al., 2016)
Soil	Europe	Switzerland	11	(Seppey et al., 2017)
		Spain	5	
		United Kingdom	4	
		Portugal	4	
		Greece	3	
		Turkey	3	
		Denmark	2	
		Scotland	1	
	North America	Canada	6	
		Alaska	1	
	Central America	Mexico	3	
		Costa Rica	1	
	South America	Brazil	4	
		Argentina	2	
		Peru	2	
		Chile	1	
	Asia	China	3	
		Kashmir	3	
		Nepal	1	
		Thailand	1	
	Africa	Ghana	2	
		South Africa	1	
	Oceania	New Caledonia	1	
	Atlantic archipelago	Spain	2	

Figure S5.1: Relative abundance of taxon sequence abundances in three environments (marine water, freshwater and soil) from samples taken around the world. Only taxa representing at least 1% of the total abundance for a micro-habitat are represented.

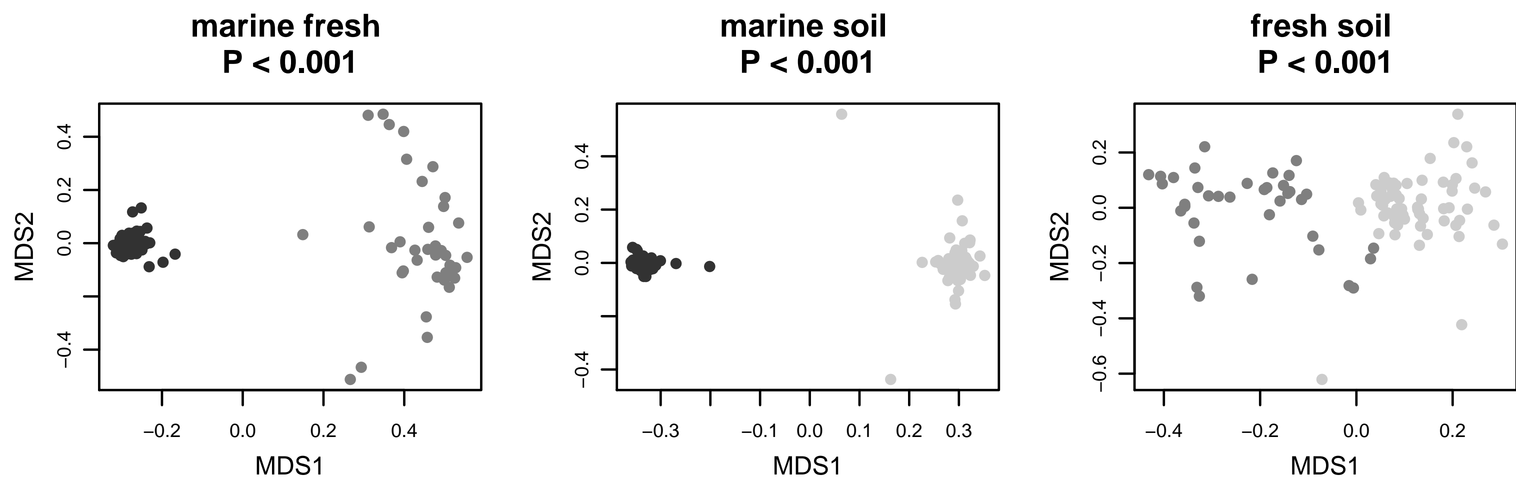


Figure S5.2: Ordination plot (non-metric multidimensional scaling) of OTU micro-eukaryotic communities for every pair of environment (marine water: dark grey, freshwater: grey, soil: pale grey) from samples taken around the world. The significance of the environmental effect is shown above the plot (permutation test: 10000 permutations).

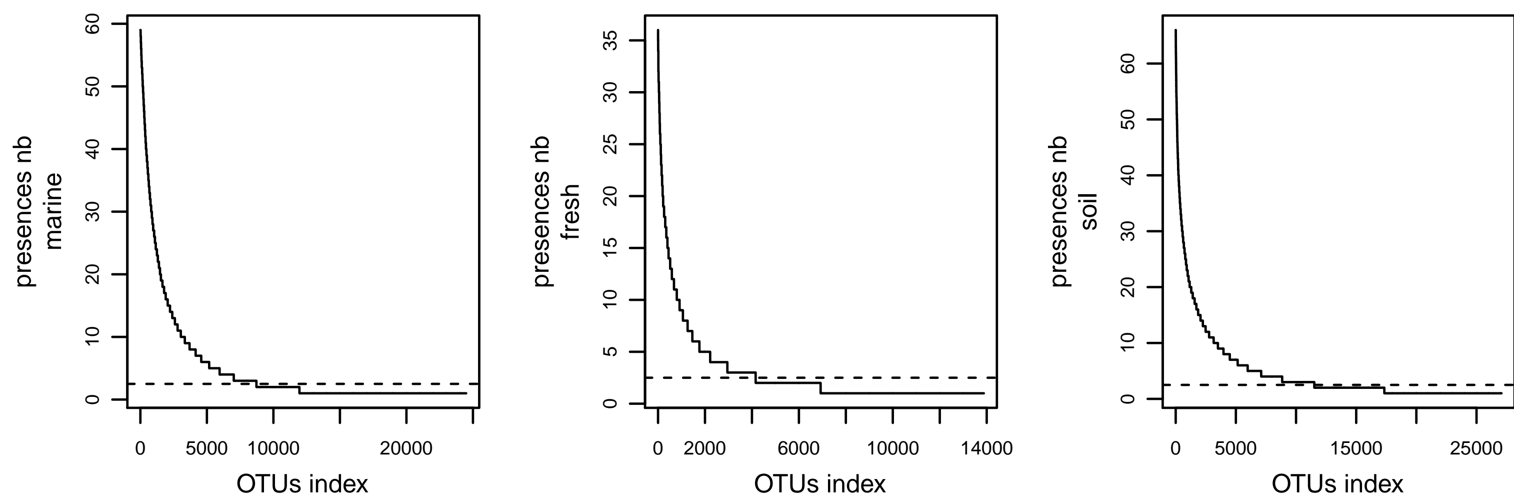


Figure S5.3: Rank abundance on the OTUs presence in the three ecosystem. Below the dashed line, the OTUs present in one or two samples.

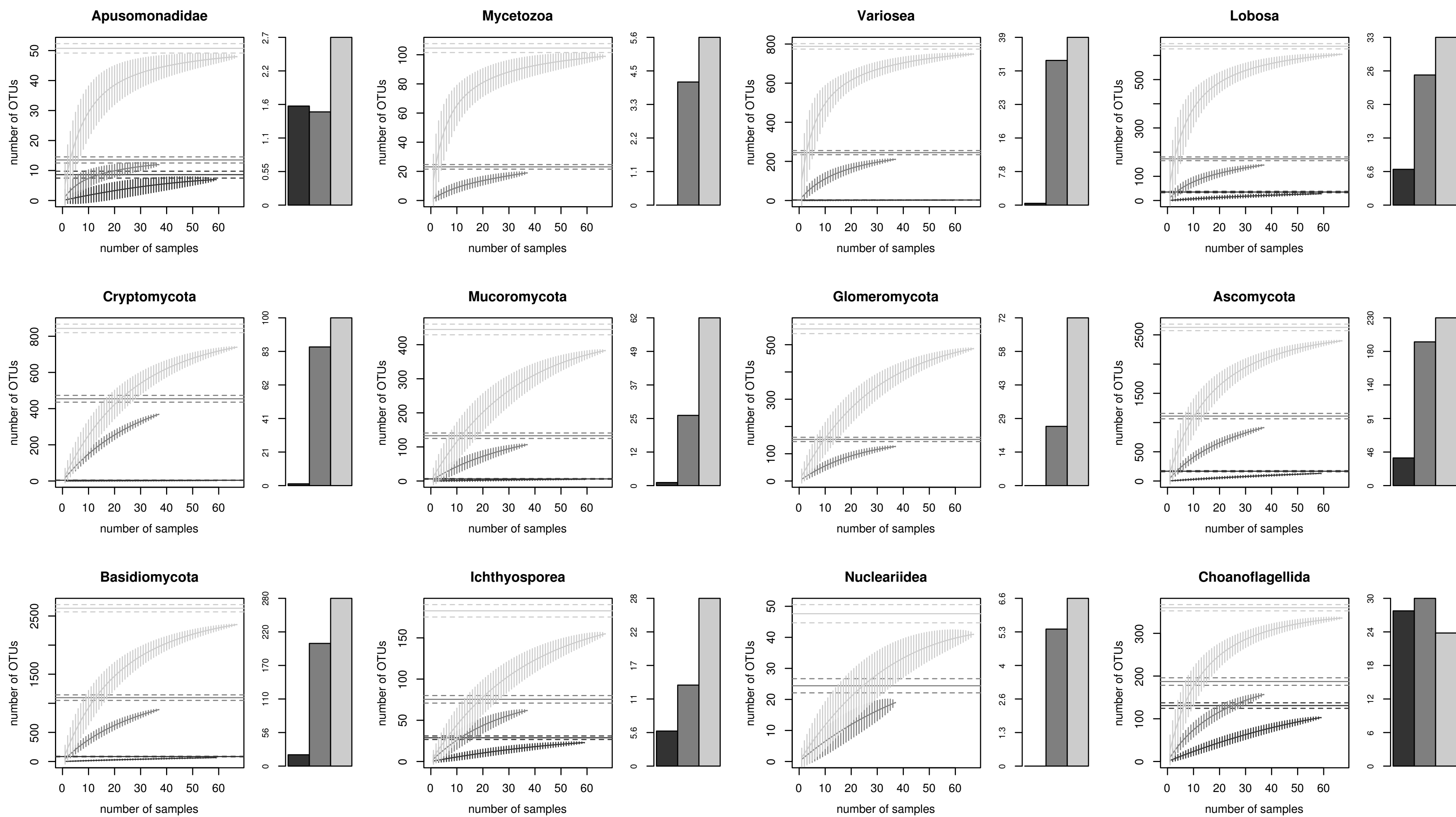


Figure S5.4: Species accumulation curves of micro-eukaryotic taxa calculated on OTU communities retrieved from marine water (dark grey), freshwater (grey) and soil (pale grey) samples taken around the world. The horizontal plain and dashed lines indicate predicted number of OTUs (from bootstrap) and standard error associated for each environment. On the right of each taxa, the barplots show the number of OTUs undiscovered (number of OTUs predicted by bootstrapping minus number of OTUs found in the dataset).

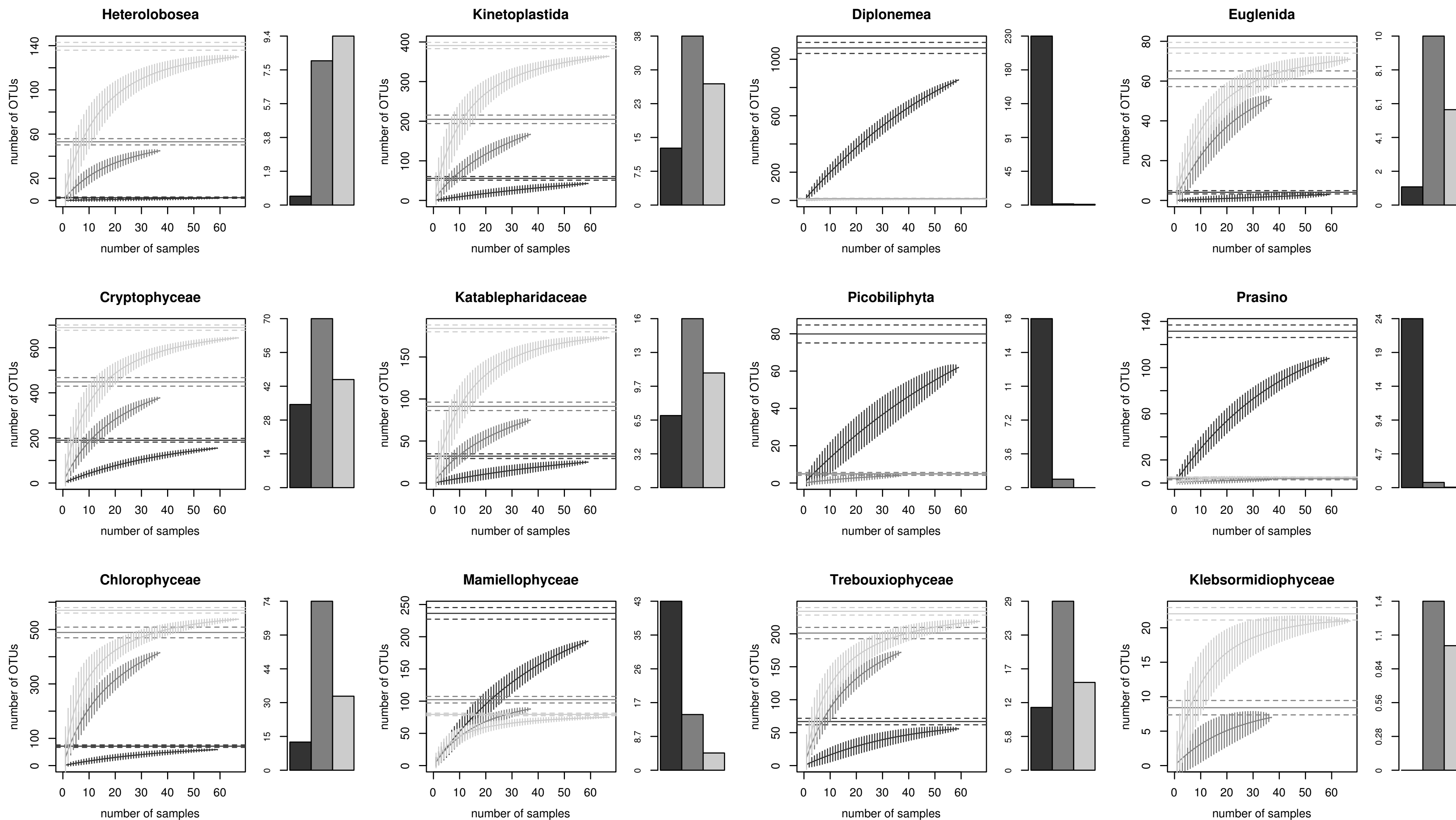


Figure S5.4: continuation

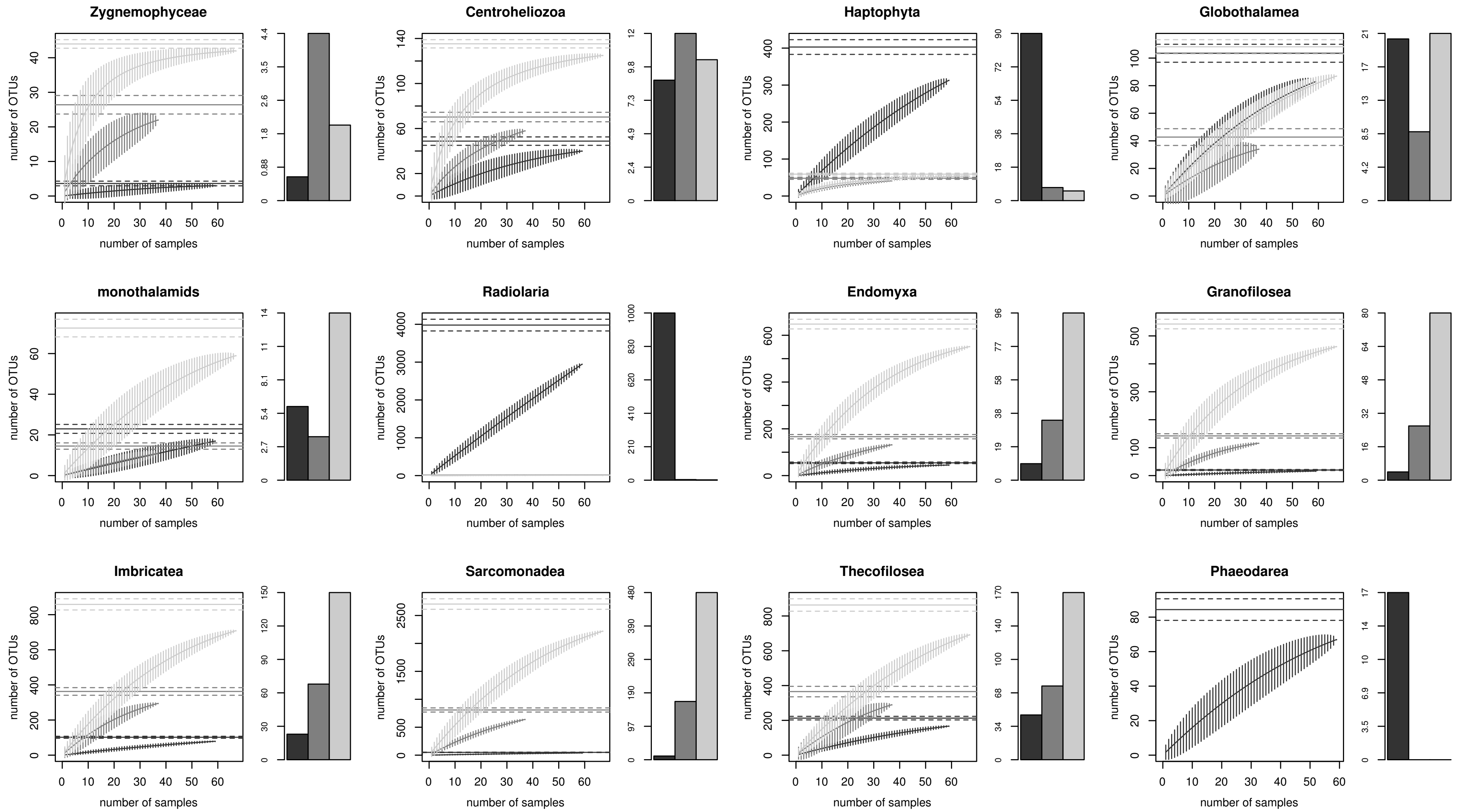


Figure S5.4: continuation

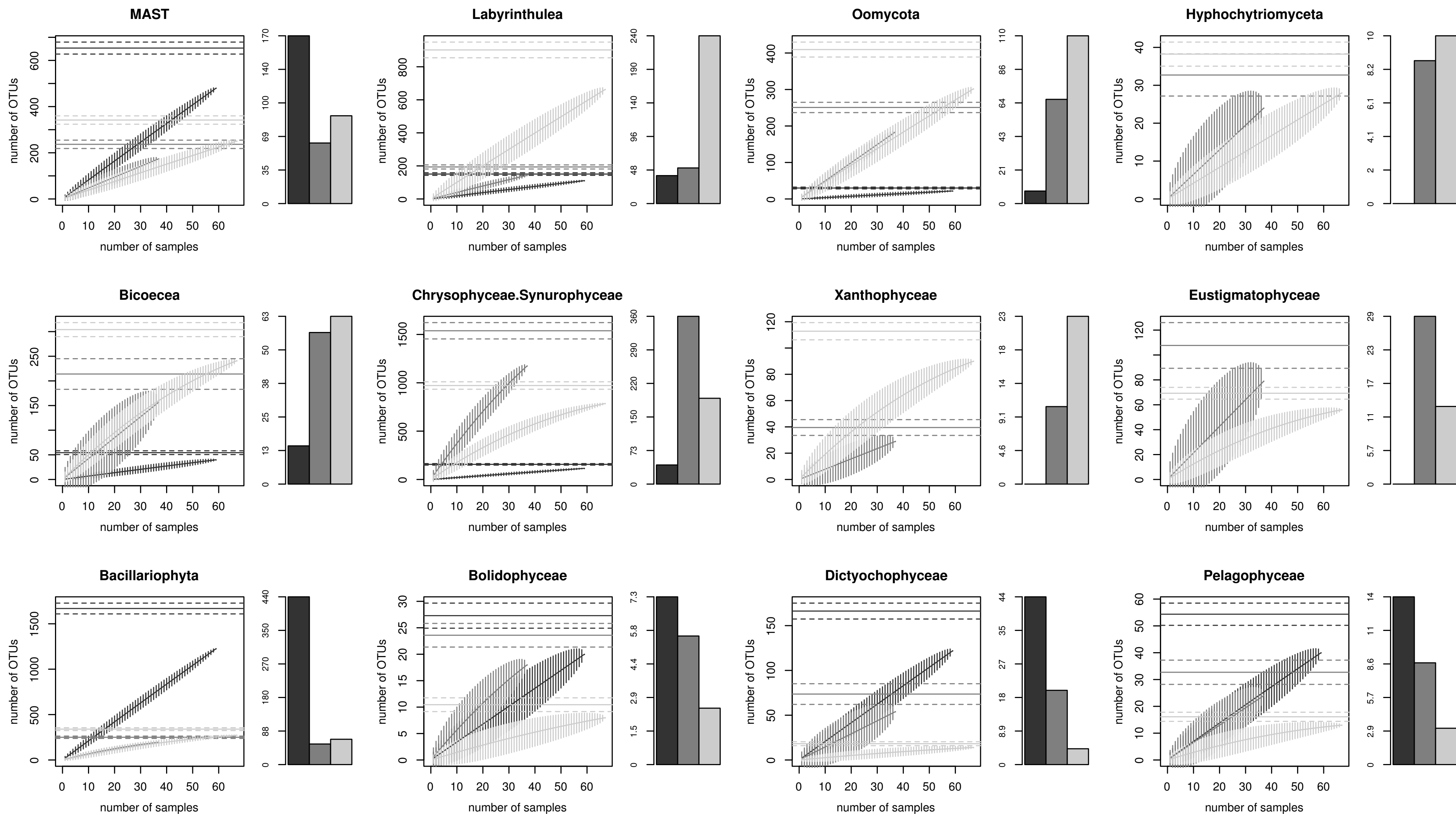


Figure S5.4: continuation

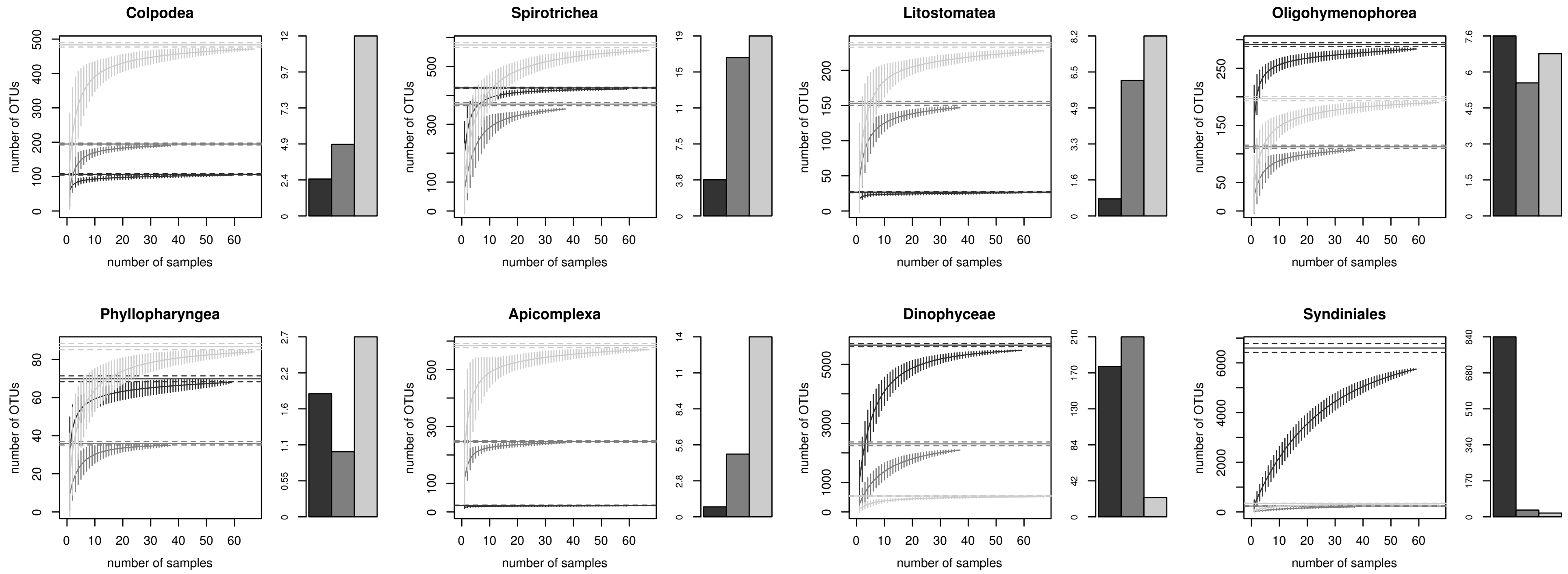


Figure S5.4: continuation

Discussion and conclusion

Christophe V.W. Seppey[1*]

[1]Laboratory of Soil Biodiversity, University of Neuchâtel, Rue Émile-Argand 11, 2000 Neuchâtel, Switzerland
*Corresponding author: christophe.seppey@unine.ch

Abstract: This thesis aimed to improve the knowledge on the ecology of micro-eukaryotes in soils using high-throughput sequencing metabarcoding. For the first time this methodology, developed in the 2010's, allows revealing a more or less complete picture of the huge environmental eukaryotic diversity of these organisms. When coupled with both classical and innovative ecological analyses, it revealed community patterns which allowed inferring fundamental mechanisms in ecosystem functioning. These conclusions could be drawn based on more than one century of knowledge accumulation on microbial functional ecology. Metabarcoding, coupled with the effort of morphological studies on micro-eukaryotes, is a promising field of research to understand the role and importance of micro-eukaryotes in soils.

1 Background

Micro-eukaryotes are key players in almost all environments on Earth. Some play fundamental roles as primary producers through photosynthesis, as decomposers, symbionts or parasites of different organisms or as predators. As such, they play considerable roles in nutrient cycling. Primary producers (eukaryotic algae) integrate CO^2 into their biomass in very large amounts: diatoms alone fix as much carbon as all rainforests on Earth (Smetacek, 1999; Gross, 2012). Fungi are major players in degradation of organic matter in soils, but other groups can be also relevant in different ecosystems like thraustochytrids in marine sediments, where they decompose highly refractory compounds which are inaccessible to bacteria (Bongiorni et al., 2005; Bongiorni, 2012). Protists have established obligate symbiotic relationships with animals like ruminants (Hungate, 2013) or termites (Watanabe et al., 1998), allowing them digesting cellulose, and sustain the growth of coral reefs (Little et al., 2004). As for parasites, malaria (*Plasmodium falciparum*) alone was responsible for 438'000 death in 2015 according to the World Health Organisation (`http://www.who.int/malaria/media/world-malaria-report-2015/en/`).

A better knowledge about these communities and their functioning is crucial for understanding ecological mechanisms at the global scale. Soils provide important ecosystem services such as nutrient recycling (directly linked to plant growth and fertility) and carbon storage, potentially mitigating greenhouse effects. Microbial eukaryotes are key players in all these processes. Fungi for instance increase the root surface of plants which significantly impacts the host growth (Harrison, 2005). As decomposers, they take up nutrients which are then made available to vascular plants through the predatory action of phagotrophic protists (Geisen et al., 2016). These predator protists also feed on prokaryotes (Clarholm, 1985). This release of nutrients in the soil system is complemented by phototrophic protists, which perform photosynthesis and are subsequently predated by phagotrophic organisms, as documented in Chapter 4 (Seppey et al., 2017). Through all these processes, microbial eukaryotes increase the soil pool of labile nutrients and are thus directly responsible for soil fertility. Carbon storage, in turn, depends on the efficiency of soil respiration. The importance of fungi in soil respiration has been long assessed (Anderson and Domsch, 1973). They play a major role as decomposers, arguably more important than bacteria (de Boer et al., 2005) and therefore are important contributors in carbon release. Typically, soil with low nutrient turnover like peatlands release less C than they store, and can be considered as C-sinks (Gorham, 1991). Typically, the speed and efficiency of carbon storage and respiration is multi-factorial, but fungi develop better in dry micro-habitats of the peatland (Jaatinen et al., 2007). Moreover, eukaryotic microbes influence other organisms through their position in microbial foodwebs. Parasitism is recurrent in soils, and apicomplexans (animal parasites) are so important in tropical soils that they are supposed to drive the evolution of arthropods through an arms race mechanism (Mahé et al., 2017). Plant parasites, in turn, correspond to the many non-bacterial agricultural pests reported, which may be even more diverse in natural than in

agricultural systems (Singer et al., 2016), what is understandable considering the respective diversity of the host in the two ecosystems.

Until recently, micro-eukaryotes living in the soil were considered as arduous to study for many reasons. Soil heterogeneity firstly complicates the spatial and temporal assessment of a microbial community, thus requiring larger experimental designs than those used in aquatic microbial ecology. Direct observation cannot always be used reliably to identify communities, because most transparent forms are hidden by soil particles (Foissner, 1999b). Soil protists spend most of the time as dormant cysts, and active community composition is therefore difficult to assess (Foissner, 1999b). Many small forms can be overlooked (Tarnawski and Lara, 2015) or their varied life stages can lead to taxonomic misinterpretations (Blandenier et al., 2017). Cryptic diversity, i.e. forms that are morphologically identical but genetically different and may have also different ecological optima, also biases observations based on morphology (Singer et al., 2015).

However, given their immense diversity, a holistic view of micro-eukaryotic communities was impossible to assess until recently because available tools (both microscopic and molecular) were insufficient to reveal their entire diversity. Indeed, cloning/sequencing approaches did not reach the sequencing depth to reveal a significant part of the communities, especially the rare biosphere (Schiaffino et al., 2016). High throughput sequencing permitted accessing these rare organisms, and continuous technological improvements brought the DNA based methods to become the golden standard in microbial communities assessments (Pawlowski et al., 2016).

2 Achievements

The main aims of this thesis were (1) to characterize soil micro-eukaryotic diversity using the recent advances in molecular techniques, and to draw conclusions on ecosystem functioning and (2) to apply and adapt new statistical tools to the study of environmental micro-eukaryotic diversity.

Soil micro-eukaryotes diversity

In this thesis, we showed that micro-eukaryotic diversity was probably higher in soils than in other better studied environments such as plankton (Chapter 5). Spatial heterogeneity is most probably responsible for this higher richness; indeed, beta diversity was higher than in both freshwater and marine plankton, and its turnover component clearly higher (Chapter 5, p. 88). Thus, diversity accumulates at large scales and ends up being larger than in the plankton. This might be due to the facts that (1) dispersal is not as straightforward as in planktonic systems and (2) local environmental conditions, which vary strongly from site to site, exert differential ecological filters that shape communities (Baselga, 2010). Micro-eukaryotic diversity does not necessarily follow above-ground patterns: croplands, which are undoubtedly a habitat with a low plant diversity, harbour in general a higher micro-eukaryotic diversity than, for instance, forests (H $\sim$ 5) (p. 63). The factors influencing their diversity have still not been fully characterized, which makes the study of micro-eukaryotic communities a specific field of ecology where the theories need to be adapted.

Characteristic soil micro-eukaryotic communities typically host many Fungi, typically Basidiomycota in forests and Ascomycota and Mucoromycota in open habitats (Chapters 4). Other well represented large eukaryotic groups are the Cercozoa (Rhizaria), represented mostly by an immense diversity of small flagellates and amoeboflagellates (Harder et al., 2016) (Chapter 2, 3, 4 and 5). Furthermore, many taxa, thought as strictly aquatic, are more and more frequently found in soils like Foraminifera (Lejzerowicz et al., 2010), MAST stramenopiles (Figure 5.4) or dinoflagellates (Bates et al., 2013). In general, micro-eukaryotic diversity (particularly protists) has been less studied in soils than in aquatic samples partly because of difficulties in isolating and cultivating soil strains (Geisen et al., 2017).

An outcome of the strong environmental filters exerted by soils is described in Chapter 1. The influence of a cadaver on the soil underneath reveals a whole array of organisms that have rarely or never been reported. These microbial eukaryotes are supposed to be transported as cysts and to develop only when suitable conditions are met. They are, therefore, excellent bioindicators which can then be reliably used to estimate the interval elapsed since death (post mortem interval estimation). Examples of bioindicators can be found in the euglyphids OTUs eugly_666 and eugly_13 which appeared after one year (Seppey et al., 2016). Sequences from similar organisms were found in a Japanese sewage (Miyaoka et al., 2017), also characterized by high nutrient amounts. An ongoing project on the same dataset (Reczuga et al., in prep.) also identified a cadaver indicator OTU related to the very nitrophilic and rarely isolated *Fonticula alba* (Worley et al., 1979). Altogether, extremely N-rich samples such as cadavers or faeces host a very specialized eukaryotic diversity comprising many uncharted clades (Bass et al., 2016).

Functional characterization of microbial eukaryotes and the variation of their communities along ecological gradients can inform on ecosystem functioning. The role of bacterivorous protists in soil ecosystems has been highlighted since the early eighties, when it was shown that their predatorial effect on bacteria liberated labile compounds which could be then taken up by plants. This model, called "microbial loop" (Clarholm, 1985), was eventually refined later when it was found that amoeba predation selected auxin-producing bacteria in the rhizosphere, thus indirectly driving root architecture (Bonkowski, 2004). Later, the significance of protist predation on other eukaryotes was evidenced. Fungal growth can be controlled by protists, which may have deleterious effects on plants if mycorrhizae are destroyed, but also positive if phytopathogenic species can be eliminated (Geisen et al., 2016; Foissner, 1999a). In any case, by

predating fungi or bacteria, protists release labile nutrients in the system. In Chapter 4 (Seppey et al., 2017), we showed another pathway for carbon to enter soil from the atmosphere through subaerial algae, that are consumed by protists which can be specialized on this type of prey. The role of phototrophic preys nutrient cycling has been widely overlooked in soils in comparison to aquatic ecosystems but are potentially a significant carbon input in soils trophic networks (Schmidt et al., 2016). The carbon fixation through micro-eukaryotic photosynthesis depends not only on the taxonomic composition of the phototrophic taxa but also on the ratio photosynthesis/phagotrophy eukaryotic algae are involved in. Indeed, this ratio can vary a lot according to light or nutrient amount as most of these organisms are potentially mixotrophs (Mitra et al., 2016). Considering the importance that mixotrophs and phototrophs can obtain in terrestrial ecosystems, the study of eukaryotic photosynthesis in soils is a promising field of research.

Use of unusual analyses

The use of statistical tools in ecology was first aimed to study macroscopic organisms. In this thesis, we applied some of them to microbial eukaryotic diversity. An example is the application of spatial distribution models to the diversity of protist groups in a landscape (Chapter 3). Initially developed to map abundance and distribution of macro-organisms like plants (Guisan and Zimmermann, 2000), these analyses have been applied here to common soil protist groups to predict their diversity in the Swiss western Alps. Most studies on soil microbial diversity focus on local edaphic conditions (Ekelund and Ronn, 1994; Eisenhauer et al., 2012) rather than topo-climatic variables (but see Lara et al., 2016). Here, we showed that in most cases, a large part of the diversity variance can be explained by both edaphic and topo-climatic variables, and that the predictive power is slightly better for topo-climatic predictors than for edaphic variables for most of the taxa studied. The application of this type of models may give precious clues about biodiversity hotspots or zones at risk of pathogen infestations, like in the case of oomycetes, a well-known group of plant pathogens. Indeed, some species like *Phytophthora cinnamomi* (a wide spectrum virulent tree pathogen) are nowadays expanding their distribution range, possibly taking advantage of climate change (Hardham, 2005). The application of spatial models would be extremely useful to predict, at regional scale, its expansion in order to take the appropriate measures (Duque-Lazo and Navarro-Cerrillo, 2017; Hudgins et al., 2017).

Assessing community diversity may be seen as a first step in the exploration of micro-eukaryotic communities. However, beyond taxonomic and functional diversity, communities are also characterized by a dense and complex network of interactions between their members. The structure of these links can be characterized with co-occurrence networks, a tool first developed in the information theory. In biology, it has been applied to genomics (Friedman and Alm, 2012) transcriptomics (Moschen et al., 2016) and, recently, to microbial ecology (Barberan et al., 2012). These networks have been shown to be more sensitive than plain community composition to monitor atmospheric pollution (Karimi et al., 2016), which is extremely useful in environmental monitoring. In Chapter 2, we combined functional characterization and network analysis to determine which OTUs may be keystone species in peatbog ecosystems, and, based on these considerations, inferred hypothesis on ecosystem functioning as a whole. In our case, first level predators and mixotrophs, which feed mostly on decomposers such as bacteria and small yeasts, and the decomposers themselves, constitute the keystone organisms. This suggests that the system is bottom-up regulated. The advantage of mixotrophs and low level predators seems logical given the characteristic lack of nutrients in peat-bogs.

3 Perspectives

Metabarcoding is undeniably a powerful tool to assess the diversity and structure of micro-organisms communities. Even though this technique is nowadays mostly used for academic research purposes, it increasingly appears as promising for applied research like bioindication in polluted (Pawlowski et al., 2016) or perturbed zones (Seppey et al., 2016) and, as suggested previously, for pathogen risk assessment (Chapter 3). The advantages of metabarcoding are that it provides a quick and reliable picture of an environments diversity at relatively low cost, which decreases year after year. This rapid screening of the micro-eukaryotic diversity allows then to target interest organisms groups, that can be studied through their morphology or by other molecular markers. High throughput approaches like Illumina permitted saturating OTU diversity in entire biomes, as it has been done for sunlit planktonic communities from the global ocean in the TARA project (de Vargas et al., 2015). This technology also multiplies the list of unknown OTUs already started with cloning/sequencing approaches in the 2000's (van Hannen et al., 1999; Lopez-Garcia et al., 2001; Diez et al., 2001).

Research on microbial eukaryotes has relied on morphospecies described on the base of light- and sometimes electron microscopy for a long time. Nowadays, it is the practice to add a taxonomically informative DNA sequence, in addition to morphometrical data to the description of species. In addition, data on diet, habitat and behaviour are of utmost importance to gather the most information on their functional ecology and thus build a reference database that allows giving most meaning to environmental DNA surveys data. In this purpose, a common, groundbreaking effort is being recently undertaken by the protistologist community to unite this existing knowledge in a single database, the UniEuk Project (Berney et al., 2017).

Nowadays, sequencing technology applied to environmental DNA allows sequencing larger and larger fragments, with sequencing depth increasing accordingly. It can be foreseen that in a near future, whole ribosomal

operons will be amplified, which will help considerably in improving the phylogenetic reconstruction of the eukaryotic tree (Marande et al., 2009). Availability of ITS sequences will also make studies at the population level possible, as this marker is several times more variable than SSU rRNA and commonly used in protist population-level studies (Pawlowski et al., 2012; Heger et al., 2011). At the same time, it will be possible to assemble partial genomes out of environmental DNA, allowing at least the retrieval of functional genes in certain cases. These genes will be linked to the ribosomal operons or other taxonomically informative DNA fragments, and organisms will therefore be characterized phylogenetically. In the meantime, scientists working on micro-eukaryotes genomics will increase the size of genetic databases, what will allow annotating these gene sequences obtained through environmental sequencing and belonging to unchartered organisms. When these databases will be built and the corresponding powerful bioinformatics tools will be developed, we will be in a good position to give the "dark matter" of soil diversity a face and a meaning.

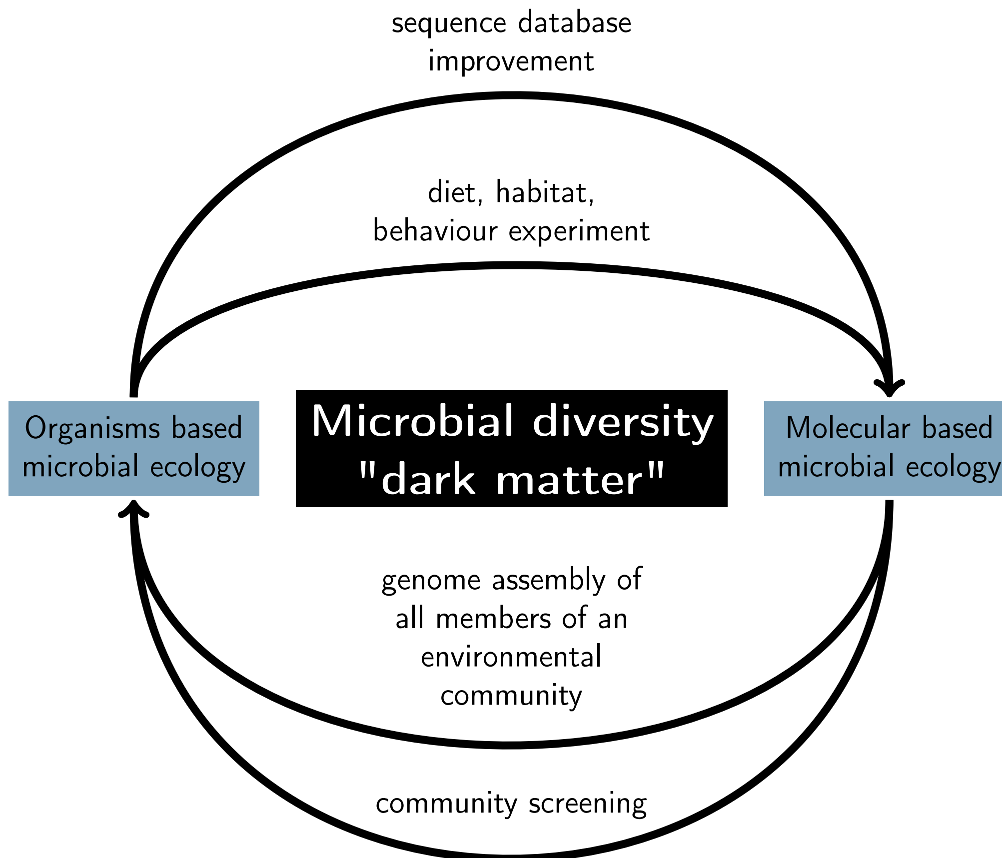


Figure S5.5: How organisms and molecular based methods can increase knowledge in reciprocals fields.

Bibliography

Adl MS, and Gupta VVSR. (2006) **Protists in soil ecology and forest nutrient cycling** Can J Forest Res 36(7):1805–1817.

Adl SM, Habura A, and Eglit Y. (2013) **Amplification primers of SSU rDNA for soil protists** Soil Biol Biochem 69:328–342.

Aitkenhead-Peterson JA, Alexander MB, Bytheway JA, Carter DO, and Wescott DJ. (2015) **Applications of Soil Chemistry in Forensic Entomology**. In Tomberlin J, and Benbow M (editors), Forensic entomology pages 283–296. CRC Press. chapter 23.

Allison SD, and Martiny BH. (2008) **Resistance, resilience, and redundancy in microbial communities** Proc Natl Acad Sci USA 105(1): 11512–11519.

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, and Lipman DJ. (1997) **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs** Nucleic Acids Res 25 (17):3389–3402.

Amaral-Zettler LA, McCliment EA, Ducklow HW, and Huse SM. (2009) **A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes** PLoS One 4(7):e6372.

Amendt J, Krettek R, and Zehner R. (2004) **Forensic entomology** Naturwissenschaften 91(2):51–56.

Anderson B, Meyer J, and Carter DO. (2013) **Dynamics of Ninhydrin-Reactive Nitrogen and pH in Gravesoil During the Extended Post-mortem Interval** J Forensic Sci 58(5):1348–1352.

Anderson JP, and Domsch KH. (1973) **Quantification of Bacterial and Fungal Contributions to Soil Respiration** Arch Microbiol 93(2):113–127.

Anderson OR. (2014) **Living Together in the Plankton: A Survey of Marine Protist Symbioses** Acta Protozool 53(1):29–38.

Anderson OR, and Cavalier-Smith T. (2012) **Ultrastructure of *Diplophrys parva*, a New Small Freshwater Species, and a Revised Analysis of Labyrinthulea (Heterokonta)** Acta Protozool 51:291–304.

Anderson TH, and Weigel HJ. (2003) **On the current debate about soil biodiversity** LANDBAUFORSCH VOLK 53(4):223–233.

Andre HM, Ducarme X, and Lebrun P. (2002) **Soil biodiversity: myth, reality or conning?** Oikos 96(1):3–24.

Appeltans W, Ahyong ST, Anderson G, Angel MV, Artois T, Bailly N, et al. (2012) **The Magnitude of Global Marine Species Diversity** Curr Biol 22(23):2189–2202.

Arrigo KR. (2005) **Marine microorganisms and global nutrient cycles** Nature 437(7057):349–355.

Azam F, Fenchel T, Field JG, Gray JS, Meyerreil LA, and Thingstad F. (1983) **The ecological role of water-column microbes in the sea** Mar Ecol Prog Ser 10(3):257–263.

Barberan A, Bates ST, Casamayor EO, and Fierer N. (2012) **Using network analysis to explore co-occurrence patterns in soil microbial communities** ISME J 6(2):343–351.

Bardgett RD. (2005)a **Soil formation**. In Crawley M, Little C, Southwood T, and Ulfstrand S (editors), The Biology of Soil, A community and ecosystem approach pages 1–6. Oxford University Press Oxford. chapter 1.2.

Bardgett RD. (2005)b **Soil-forming factors**. In Crawley M, Little C, Southwood T, and Ulfstrand S (editors), The Biology of Soil, A community and ecosystem approach pages 6–17. Oxford University Press Oxford. chapter 1.3.

Bardgett RD. (2005)c **Soil biological properties and global change**. In Crawley M, Little C, Southwood T, and Ulfstrand S (editors), The Biology of Soil, A community and ecosystem approach pages 140–182. Oxford University Press Oxford. chapter 6.

Barton PS, Cunningham SA, Lindenmayer DB, and Manning AD. (2013) **The role of carrion in maintaining biodiversity and ecological processes in terrestrial ecosystems** Oecologia 171(4):761–772.

Baselga A. (2010) **Partitioning the turnover and nestedness components of beta diversity** Global Ecol Biogeogr 19(1):134–143.

Baselga A, Orme D, Villeger D, De Bortoli J, and Leprieur F. (2013) **betapart: Partitioning beta diversity into turnover and nestedness components**. URL `http://CRAN.R-project.org/package=betapart`. R package version 1.3.

Bass D, Richards TA, Matthai L, Marsh V, and Cavalier-Smith T. (2007) **DNA evidence for global dispersal and probable endemicity of protozoa** BMC Evol Biol 7.

Bass D, Chao EEY, Nikolaev S, Yabuki A, Ishida K, Berney C, et al. (2009) **Phylogeny of Novel Naked Filose and Reticulose Cercozoa: Granofilosea cl. n. and Proteomyxidea Revised** Protist 160(1):75–109.

Bass D, Silberman JD, Brown MW, Pearce RA, Tice AK, Jousset A, et al. (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** Environ Microbiol 18(5):1604–1619.

Bates ST, Clemente JC, Flores GE, Walters WA, Parfrey LW, Knight R, and Fierer N. (2013) **Global biogeography of highly diverse protistan communities in soil** ISME J 7(3):652–659.

Batzer DP, and Baldwin AH. (2012) **Wetland Habitats of North America: Ecology and Conservation Concerns**. Stephen Bechtel Fund imprint in ecology and the environment. University of California Press.

Beakes GW, Glockling SL, and Sekimoto S. (2012) **The evolutionary phylogeny of the oomycete "fungi"** Protoplasma 249(1):3–19.

Behnke A, Engel M, Christen R, Nebel M, Klein RR, and Stoeck T. (2011) **Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions** Environ Microbiol 13:340–349.

Belnap J. (2003) **The world at your feet: desert biological soil crusts** Front Ecol Environ 1(4):181–189.

Belyea LR, and Malmer N. (2004) **Carbon sequestration in peatland: patterns and mechanisms of response to climate change** Glob Change Biol 10(7):1043–1052.

Benjamini Y, and Hochberg Y. (1995) **Controlling the False Discovery Rate - A Practical and Powerful Approach to Multiple Testing** J R Stat Soc Series B Stat Methodol 57(1):289–300.

Benninger LA, Carter DO, and Forbes SL. (2008) **The biochemical alteration of soil beneath a decomposing carcass** Forensic Sci Int 180 (2-3):70–75.

Berard A, Dorigo U, Humbert JF, and Martin-Laurent F. (2005) **Microalgae community structure analysis based on 18S rDNA amplification from DNA extracted directly from soil as a potential soil bioindicator** Agron Sustain Dev 25(2):285–291.

Berg B. (2000) **Litter decomposition and organic matter turnover in northern forest soils** Forest Ecol Manag 133(1):13–22.

Berge T, Hansen PJ, and Moestrup O. (2008) **Feeding mechanism, prey specificity and growth in light and dark of the plastidic dinoflagellate *Karlodinium armiger*** Aquat Microb Ecol 50(3):279–288.

Berner EK, and Berner RA. (2012) **Global environment: water, air, and geochemical cycles**. Princeton University Press.

Berney C, Ciuprina A, Bender S, Brodie J, Edgcomb V, Kim E, et al. (2017) **UniEuk: Time to Speak a Common Language in Protistology!** J Eukaryot Microbiol 64(3):407–411.

Beyens L, Ledeganck P, Graae BJ, and Nijs I. (2009) **Are soil biota buffered against climatic extremes? An experimental test on testate amoebae in arctic tundra (Qeqertarsuaq, West Greenland)** Polar Biol 32(3): 453–462.

Blandenier Q, Seppey CVW, Singer D, Vlimant M, Simon A, Duckert C, and Lara E. (2017) ***Mycamoeba gemmipara* nov. gen., nov. sp., the First Cultured Member of the Environmental Dermamoebidae Clade LKM74 and its Unusual Life Cycle.** J Eukaryot Microbiol 64(2): 257–265.

Bobrov AA, Charman DJ, and Warner BG. (1999) **Ecology of testate amoebae (Protozoa : Rhizopoda) on peatlands in western Russia with special attention to niche separation in closely related taxa** PROTIST 150(2):125–136.

Boddy L, Frankland J, and van West P. (2007) **Ecology of Saprotrophic Basidiomycetes**. British Mycological Society Symposia Series. Elsevier Science.

Boenigk J. (2008) **Nanoflagellates: functional groups and intraspecific variation** DENISIA 23:331–335.

Boenigk J, Pfandl K, Stadler P, and Chatzinotas A. (2005) **High diversity of the '*Spumella*-like' flagellates: an investigation based on the SSU rRNA gene sequences of isolates from habitats located in six different geographic regions** ENVIRON MICROBIOL 7(5):685–697.

Boiffin J, Papy F, and Eimberck M. (1988) **Influence of Cropping Systems on Concentrated Flow Erosion Risks .1. Analysis of the Conditions for Initiating Erosion** AGRON J 8(8):663–673.

Bongiorni L. (2012). **Thraustochytrids, a Neglected Component of Organic Matter Decomposition and Food Webs in Marine Sediments** pages 1–13. Springer Berlin, Heidelberg.

Bongiorni L, Pusceddu A, and Danovaro R. (2005) **Enzymatic activities of epiphytic and benthic thraustochytrids involved in organic matter degradation** AQUAT MICROB ECOL 41(3):299–305.

Bonkowski M. (2004) **Protozoa and plant growth: the microbial loop in soil revisited** NEW PHYTOL 162(3):617–631.

Bonkowski M, and Clarholm M. (2012) **Stimulation of Plant Growth through Interactions of Bacteria and Protozoa: Testing the Auxiliary Microbial Loop Hypothesis** ACTA PROTOZOOL 51(3):237–247.

Boomsma JJ, Jensen AB, Meyling NV, and Eilenberg J. (2014) **Evolutionary Interaction Networks of Insect Pathogenic Fungi**. In Berenbaum M (editor), ANNUAL REVIEW OF ENTOMOLOGY volume 59 pages 467–485.

Booth RK. (2001) **Ecology of testate amoebae (Protozoa) in two lake superior coastal wetlands: Implications for paleoecology and environmental monitoring** WETLANDS 21(4):564–576.

Booth RK. (2008) **Testate amoebae as proxies for mean annual water-table depth in Sphagnum-dominated peatlands of North America** J QUATERNARY SCI 23(1):43–57.

Borcard D, Legendre P, and Gillet F. (2011) **Numerical Ecology with R**. Springer New York.

Bornemissza GF. (1957) **An analysis of arthropod succession in carrion and the effect of its decomposition on the soil fauna** J ZOOL (5):1–12.

Bragazza L, Siffi C, Iacumin P, and Gerdol R. (2007) **Mass loss and nutrient release during litter decay in peatland: the role of microbial adaptability to litter chemistry** SOIL BIOL BIOCHEM 39(1):257–267.

Brandt A, Gooday AJ, Brandao SN, Brix S, Broekeland W, Cedhagen T, et al. (2007) **First insights into the biodiversity and biogeography of the Southern Ocean deep sea** NATURE 447(7142):307–311.

Brodie Jr ED, Formanowicz Jr DR, and Brodie III ED. (1991) **Predator avoidance and antipredator mechanisms: distinct pathways to survival** 3(1):73–77.

Bulit C. (2014) **Good Reasons and Guidance for Mapping Planktonic Protist Distributions** ACTA PROTOZOOL 53(1):13–27.

Bunting L, Leavitt PR, Simpson GL, Wissel B, Laird KR, Cumming BF, et al. (2016) **Increased variability and sudden ecosystem state change in Lake Winnipeg, Canada, caused by 20th century agriculture** LIMNOL OCEANOGR 61(6):2090–2107.

Burkholder JM, Glibert PM, and Skelton HM. (2008) **Mixotrophy, a major mode of nutrition for harmful algal species in eutrophic waters** HARMFUL ALGAE 8(1, SI):77–93.

Cairney JWG, and Chambers SM. (1999) **Ectomycorrhizal Fungi: Key Genera in Profile**. Springer-Verlag Berlin Heidelberg 1st edition.

Caron DA, and Hutchins DA. (2012) **The effects of changing climate on microzooplankton grazing and community structure: drivers, predictions and knowledge gaps** J PLANKTON RES 35(2):fbs091.

Carter DO, and Tibbett M. (2003) **Taphonomic mycota: Fungi with forensic potential** J FORENSIC SCI 48(1):168–171.

Carter DO, Yellowlees D, and Tibbett M. (2007) **Cadaver decomposition in terrestrial ecosystems** NATURWISSENSCHAFTEN 94(1):12–24.

Carter MR, and Gregorich EG. (2007) **Soil Sampling and Methods of Analysis, Second Edition**. CRC Press.

Chapin III FS, Matson PA, and Vitousek P. (2011) **Principles of terrestrial ecosystem ecology**. Springer Science & Business Media.

Charman DJ. (2001) **Biostratigraphic and palaeoenvironmental applications of testate amoebae** QUATERNARY SCI REV 20(16-17):1753–1764.

Chatelain AP, Meisterfeld R, Roussel-Delif L, and Lara E. (2013) **Sphenoderiidae (fam. nov.), a New Clade of Euglyphid Testate Amoebae Characterized by Small, Round Scales Surrounding the Aperture** PROTIST 164(6):782–792.

Chen J, Blume HP, and Beyer L. (2000) **Weathering of rocks induced by lichen colonization - a review** CATENA 39(2):121–146.

Clarholm M. (1985) **Interactions of bacteria, protozoa and plants leading to mineralization of soil-nitrogen** SOIL BIOL BIOCHEM 17(2):181–187.

Clarholm M. (2005) **Soil protozoa: an under-researched microbial group gaining momentum** SOIL BIOL BIOCHEM 37(5):811–817.

Clymo RS, and Hayward PM. (2012) **The ecology of *Sphagnum***. In Smith A (editor), BRYOPHYTE ECOLOGY pages 229–289. Chapman and Hall London and New York.

Cobanoglu G. (2015) **The Use of Lichens for Biomonitoring of Atmospheric Pollution** SIGMA J ENGINEERING NATURAL SCIENCES 33(4): 591–613.

Coleman DC, Crossley Jr DA, and Hendrix PF. (2004)a **Historical overview of soils and the fitness of the soil environment**. In Cella D, Sonnack K, and B P (editors), FUNDAMENTALS OF SOIL ECOLOGY pages 1–23. Elsevier. chapter 1.

Coleman DC, Crossley Jr DA, and Hendrix PF. (2004)b **Future developments in soil ecology**. In Cella D, Sonnack K, and B P (editors), FUNDAMENTALS OF SOIL ECOLOGY pages 271–299. Elsevier. chapter 8.

Connell JH. (1972) **On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees**. In den Boer P, and Gradwell G (editors), DYNAMICS OF POPULATIONS pages 298–392. Pudoc.

Corsaro D, Walochnik J, Venditti D, Mueller KD, Hauroeder B, and Michel R. (2014) **Rediscovery of *Nucleophaga* amoebae, a novel member of the Rozellomycota** PARASITOL RES 113(12):4491–4498.

Csardi G, and Nepusz T. (2006) **The igraph software package for complex network research** INTERJOURNAL Complex Systems:1695.

Darbyshire JF, Wheatley RE, Greaves MP, and Inkson RHE. (1974) **A rapid micromethod for estimating bacterial and protozoan populations in soil** REV ECOL BIOL SOL 11(4):465–475.

Davidson EA, and Janssens IA. (2006) **Temperature sensitivity of soil carbon decomposition and feedbacks to climate change** NATURE 440 (7081):165–173.

Davidson EA, Trumbore SE, and Amundson R. (2000) **Biogeochemistry: soil warming and organic carbon content** NATURE 408(6814):789–790.

Davidson K. (1996) **Modelling microbial food webs** MAR ECOL PROG SER 145(1-3):279–296.

de Boer W, Folman LB, Summerbell RC, and Boddy L. (2005) **Living in a fungal world: impact of fungi on soil bacterial niche development** FEMS MICROBIOL REV 29(4):795–811.

de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, et al. (2015) **Eukaryotic plankton diversity in the sunlit ocean** SCIENCE 348 (6237).

Del Campo J, Balague V, Forn I, Lekunberri I, and Massana R. (2013) **Culturing bias in marine heterotrophic flagellates analyzed through seawater enrichment incubations** MICROBIAL ECOL 66(3):489–499.

Delarue F, Buttler A, Bragazza L, Grasset L, Jassey VEJ, Gogo S, and Laggoun-Defarge F. (2015) **Experimental warming differentially affects microbial structure and activity in two contrasted moisture sites in a *Sphagnum*-dominated peatland** SCI TOTAL ENVIRON 511:576–583.

Delarze R, Gonseth Y, Eggenberg S, and Vust M. **Guide des milieux naturels de Suisse: Écologie, menaces, espèces caractéristiques**. Bussigny, Switzerland.

Dent BB, Forbes SL, and Stuart BH. (2004) **Review of human decomposition processes in soil** ENVIRON GEOL 45(4):576–585.

Dieleman CM, Branfireun BA, Mclaughlin JW, and Lindo Z. (2015) **Climate change drives a shift in peatland ecosystem plant community: Implications for ecosystem function and stability** Glob Change Biol 21(1):388–395.

Diez B, Pedros-Alio C, and Massana R. (2001) **Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing** Appl Environ Microbiol 67(7): 2932–2941.

Dorrepaal E, Toet S, van Logtestijn RSP, Swart E, van de Weg MJ, Callaghan TV, and Aerts R. (2009) **Carbon respiration from subsurface peat accelerated by climate warming in the subarctic** Nature 460(7255):616–619.

Dumack K, Baumann C, and Bonkowski M. (2016)a **A Bowl with Marbles: Revision of the Thecate Amoeba Genus *Lecythium* (Chlamydophryidae, Tectofilosida, Cercozoa, Rhizaria) Includinga Description of Four New Species and an Identification Key** Protist 167(5):440–459.

Dumack K, Mueller MEH, and Bonkowski M. (2016)b **Description of *Lecythium terrestris* sp nov (Chlamydophryidae, Cercozoa), a Soil Dwelling Protist Feeding on Fungi and Algae** Protist 167(2):93–105.

Dunne JA, Williams RJ, and Martinez ND. (2002) **Network structure and biodiversity loss in food webs: robustness increases with connectance** Ecol Lett 5(4):558–567.

Dunthorn M, Otto J, Berger SA, Stamatakis A, Mahé F, Romac S, et al. (2014) **Placing Environmental Next-Generation Sequencing Amplicons from Microbial Eukaryotes into a Phylogenetic Context** Mol Biol Evol 31(4):993–1009.

Dupont AOC, Griffiths RI, Bell T, and Bass D. (2016) **Differences in soil micro-eukaryotic communities over soil pH gradients are strongly driven by parasites and saprotrophs** Environ Microbiol 18(6):2010–2024.

Duque-Lazo J, and Navarro-Cerrillo RM. (2017) **What to save, the host or the pest? The spatial distribution of xylophage insects within the Mediterranean oak woodlands of Southwestern Spain** Forest Ecol Manag 392:90–104.

Edgar RC. (2004) **MUSCLE: a multiple sequence alignment method with reduced time and space complexity** BMC Bioinformatics 5:1–19.

Edgar RC, Haas BJ, Clemente JC, Quince C, and Knight R. (2011) **UCHIME improves sensitivity and speed of chimera detection** Bioinformatics 27(16):2194–2200.

Eisenhauer N, Cesarz S, Koller R, Worm K, and Reich PB. (2012) **Global change belowground: impacts of elevated CO_2, nitrogen, and summer drought on soil food webs and biodiversity** Glob Change Biol 18 (2):435–447.

Ekelund F. (1998) **Enumeration and abundance of mycophagous protozoa in soil, with special emphasis on heterotrophic flagellates** Soil Biol Biochem 30(10-11):1343–1347.

Ekelund F, and Ronn R. (1994) **Notes on protozoa in agricultural soil with emphasis on heterotrophic flagellates and naked amebas and their ecology** FEMS Microbiol Rev 15(4):321–353.

Ekelund F, Ronn R, and Christensen S. (2001) **Distribution with depth of protozoa, bacteria and fungi in soil profiles from three Danish forest sites** Soil Biol Biochem 33(4-5):475–481.

Elbert W, Weber B, Burrows S, Steinkamp J, Buedel B, Andreae MO, and Poeschl U. (2012) **Contribution of cryptogamic covers to the global cycles of carbon and nitrogen** Nat Geosci 5(7):459–462.

Eldridge DJ, and Greene RSB. (1994) **Microbiotic soil crusts-a review of their roles in soil and ecological processes in the rangelands of Australia** Soil Res 32(3):389–415.

Elser JJ, Stabler LB, and Hassett RP. (1995) **Nutrient limitation of bacterial growth and rates of bacterivory in lakes and oceans: a comparative study** Aquat Microb Ecol 9(2):105–110.

Ettema CH, and Wardle DA. (2002) **Spatial soil ecology** Trends Ecol Evol 17(4):177–183.

Falkowski PG. (2002) **The ocean's invisible forest - Marine phytoplankton play a critical role in regulating the earth's climate. Could they also be used to combat global warming** Sci Am 287(2):54–61.

Falkowski PG, and Raven JA. (2013) **Aquatic photosynthesis**. Princeton University Press.

Fellous S, and Salvaudon L. (2009) **How can your parasites become your allies?** Trends Parasitol 25(2):62–66.

Fernández LD, Fournier B, Rivera R, Lara E, Mitchell EAD, and Hernandez CE. (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 Ecol Biogeogr 25(10):1216–1227.

Fierer N, Lauber CL, Ramirez KS, Zaneveld J, Bradford MA, and Knight R. (2012) **Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients** ISME J 6(5):1007–1017.

Flynn KJ, Stoecker DK, Mitra A, Raven JA, Glibert PM, Hansen PJ, et al. (2013) **Misuse of the phytoplanktonzooplankton dichotomy: the need to assign organisms as mixotrophs within plankton functional types** J Plankton Res 35(1):3–11.

Foissner W. (1983) **Estimation of numbers of protista in soil: a test of the direct method** J Protozool 30(3):A49.

Foissner W. (1987) **Soil protozoa: fundamental problems, ecological significance, adaptations in ciliates and testaceans, bioindicators, and guide to the literature** Prog Protist (2):69–212.

Foissner W. (1997) **Protozoa as bioindicators in agroecosystems, with emphasis on farming practices, biocides, and biodiversity** Agr Ecosyst Environ 62(2-3):93–103.

Foissner W. (1999)a **Description of two new, mycophagous soil ciliates (Ciliophora, Colpodea): *Fungiphrya strobli* n. g., n. sp. and *Grossglockneria ovata* n. sp.** J Eukaryot Microbiol 46(1):34–42.

Foissner W. (1999)b **Soil protozoa as bioindicators: pros and cons, methods, diversity, representative examples** Agr Ecosyst Environ 74 (1):95–112.

Foissner W. (2006) **Biogeography and dispersal of micro-organisms: A review emphasizing protists** Acta Protozool 45(2):111–136.

Fox J, and Weisberg S. (2011) **An R Companion to Applied Regression**. Sage Thousand Oaks CA second edition.

Fraile I, Schulz M, Mulitza S, and Kucera M. (2008) **Predicting the global distribution of planktonic foraminifera using a dynamic ecosystem model** Biogeosciences 5(3):891–911.

Freckleton RP, and Lewis OT. (2006) **Pathogens, density dependence and the coexistence of tropical trees** Proc R Soc B-Biol Sci 273 (1604):2909–2916.

Freeman KR, Pescador MY, Reed SC, Costello EK, Robeson MS, and Schmidt SK. (2009) **Soil CO_2 flux and photoautotrophic community composition in high-elevation, 'barren' soil** Environ Microbiol 11 (3):674–686.

Frey B, Buehler L, Schmutz S, Zumsteg A, and Furrer G. (2013) **Molecular characterization of phototrophic microorganisms in the forefield of a receding glacier in the Swiss Alps** Environ Res Lett 8(1).

Friedman J, and Alm EJ. (2012) **Inferring Correlation Networks from Genomic Survey Data** PLoS Comput Biol 8(9).

Gaston KJ. (2000) **Global patterns in biodiversity** Nature 405(6783): 220–227.

Geisen S. (2016) **The bacterial-fungal energy channel concept challenged by enormous functional versatility of soil protists** Soil Biol Biochem 102(SI):22–25.

Geisen S, Bandow C, Roembke J, and Bonkowski M. (2014)a **Soil water availability strongly alters the community composition of soil protists** Pedobiologia 57(4-6):205–213.

Geisen S, Fiore-Donno AM, Walochnik J, and Bonkowski M. (2014)b ***Acanthamoeba* everywhere: high diversity of *Acanthamoeba* in soils** Parasitol Res 113(9):3151–3158.

Geisen S, Laros I, Vizcaino A, Bonkowski M, and De Groot GA. (2015)a **Not all are free-living: high-throughput DNA metabarcoding reveals a diverse community of protists parasitizing soil metazoa** Mol Ecol 24(17):4556–4569.

Geisen S, Rosengarten J, Koller R, Mulder C, Urich T, and Bonkowski M. (2015)b **Pack hunting by a common soil amoeba on nematodes** Environ Microbiol 17(11):4538–4546.

Geisen S, Tveit AT, Clark IM, Richter A, Svenning MM, Bonkowski M, and Urich T. (2015)c **Metatranscriptomic census of active protists in soils** ISME J 9(10):2178–2190.

Geisen S, Koller R, Huenninghaus M, Dumack K, Urich T, and Bonkowski M. (2016) **The soil food web revisited: Diverse and widespread mycophagous soil protists** Soil Biol Biochem 94:10–18.

Geisen S, Mitchell EAD, Wilkinson DM, Adl S, Bonkowski M, Brown MW, et al. (2017) **Soil protistology rebooted: 30 fundamental questions to start with** Soil Biol Biochem 111:94–103.

Germond A, Hata H, Fujikawa Y, and Nakajima T. (2013) **The phylogenetic position and phenotypic changes of a *Chlorella*-like alga during 5-year microcosm culture** Eur J Phycol 48(4):485–496.

Germond A, Inouhe M, and Nakajima T. (2014) **Effects of acidic conditions on the physiology of a green alga (*Micractinium* sp.) before and after a 5-year interaction with *Tetrahymena thermophila* in an experimental microcosm** Eur J Phycol 49(4):526–537.

Gilbert D, and Mitchell EAD. (2006) **Chapter 13 Microbial diversity in Sphagnum peatlands**. In Martini I, Martínez Cortizas A, and Chesworth W (editors), Peatlands Evolution and Records of Environmental and Climate Changes volume 9 of Developments in Earth Surface Processes pages 287–318. Elsevier.

Gilbert D, Amblard C, Bourdier G, and Francez AJ. (1998)a **Short-term effect of nitrogen enrichment on the microbial communities of a peatland** Hydrobiologia 374:111–119.

Gilbert D, Amblard C, Bourdier G, and Francez AJ. (1998)b **The microbial loop at the surface of a peatland: Structure, function, and impact of nutrient input** Microbial Ecol 35(1):83–93.

Gilbert D, Amblard C, Bourdier G, Francez AJ, and Mitchell EAD. (2000) **Le régime alimentaire des Thécamoebiens (Protista, Sarcodina)** L'Annee Biologique (39):57–68.

Giner CR, Forn I, Romac S, Logares R, de Vargas C, and Massana R. (2016) **Environmental Sequencing Provides Reasonable Estimates of the Relative Abundance of Specific Picoeukaryotes** Appl Environ Microbiol 82:4757–4766.

Glaser K, Kuppardt A, Boenigk J, Harms H, Fetzer I, and Chatzinotas A. (2015) **The influence of environmental factors on protistan microorganisms in grassland soils along a land-use gradient** Sci Total Environ 537:33–42.

Gobat JM, Aragno M, and Matthey W. (2010)a **Le cycle global du carbone, de volcans en montagne calcaires, en passant par le vivant**. In Gobat J, Aragno M, and Matthey W (editors), Le sol vivant Bases de pédologie - Biologie des sols pages 565–567. Presses polytechniques et universitaires romandes Lausanne 3e edition. chapter 15.1.1.

Gobat JM, Aragno M, and Matthey W. (2010)b **L'Azote, élément caractéristique des molécules actives du vivant**. In Gobat J, Aragno M, and Matthey W (editors), Le sol vivant Bases de pédologie - Biologie des sols pages 580–590. Presses polytechniques et universitaires romandes Lausanne 3e edition. chapter 15.3.

Gobat JM, Aragno M, and Matthey W. (2010)c **Constituant organiques**. In Gobat J, Aragno M, and Matthey W (editors), Le sol vivant Bases de pédologie - Biologie des sols pages 22–33. Presses polytechniques et universitaires romandes Lausanne 3e edition. chapter 2.2.

Gobat JM, Aragno M, and Matthey W. (2010)d **Les champignons du sol, des transporteurs à longue distance**. In Gobat J, Aragno M, and Matthey W (editors), Le sol vivant Bases de pédologie - Biologie des sols pages 43–45. Presses polytechniques et universitaires romandes Lausanne 3e edition. chapter 2.5.3.

Gobat JM, Aragno M, and Matthey W. (2010)e **Le complexe argilo-humique, propriété exclusive du sol**. In Gobat J, Aragno M, and Matthey W (editors), Le sol vivant Bases de pédologie - Biologie des sols pages 74–76. Presses polytechniques et universitaires romandes Lausanne 3e edition. chapter 3.6.

Gobat JM, Aragno M, and Matthey W. (2010)f **Le rôle essential de la faune**. In Gobat J, Aragno M, and Matthey W (editors), Le sol vivant Bases de pédologie - Biologie des sols pages 152–160. Presses polytechniques et universitaires romandes Lausanne 3e edition. chapter 4.6.

Gomaa F, Mitchell EAD, and Lara E. (2013) **Amphitremida (Poche, 1913) Is a New Major, Ubiquitous Labyrinthulomycete Clade** PLoS One 8 (1).

Gomaa F, Kosakyan A, Heger TJ, Corsaro D, Mitchell EAD, and 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.

Gong Y, Patterson DJ, Li Y, Hu Z, Sommerfeld M, Chen Y, and Hu Q. (2015) ***Vernalophrys algivore* gen. nov., sp nov (Rhizaria: Cercozoa: Vampyrellida), a New Algal Predator Isolated from Outdoor Mass Culture of *Scenedesmus dimorphus*** Appl Environ Microbiol 81(12): 3900–3913.

Gonzalez JM, Sherr EB, and Sherr BF. (1990) **Size-selective grazing on bacteria by natural assemblages of estuarine flagellates and ciliates.** Appl Environ Microbiol 56(3):583–589.

Gonzalez Garraza G, Mataloni G, Fermani P, and Vinocur A. (2011) **Ecology of algal communities of different soil types from Cierva Point, Antarctic Peninsula** Polar Biol 34(3):339–351.

Gorham E. (1991) **Northern Peatlands - Role in the Carbon-cycle and Probable Responses to Climatic Warming** Ecol Aool 1(2):182–195.

Grattepanche J, Santoferrara LF, McManus GB, and Katz LA. (2014) **Diversity of diversity: conceptual and methodological differences in biodiversity estimates of eukaryotic microbes as compared to bacteria** Trends Microbiol 22(8):432–437.

Gross M. (2012) **The mysteries of the diatoms** Curr Biol 22(15):R581–R585.

Grossmann L, Jensen M, Heider D, Jost S, Glucksman E, Hartikainen H, et al. (2016) **Protistan community analysis: key findings of a large-scale molecular sampling** ISME J 10(9):2269–2279.

Guillou L, Viprey M, Chambouvet A, Welsh RM, Kirkham AR, Massana R, et al. (2008) **Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata)** Environ Microbiol 10(12):3349–3365.

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, et al. (2013) **The Protist Ribosomal Reference database (PR^2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy** Nucleic Acids Res 41(1):D597–D604.

Guisan A, and Zimmermann NE. (2000) **Predictive habitat distribution models in ecology** Ecol Model 135(2-3):147–186.

Gunde-Cimerman N, and Zalar P. (2014) **Extremely Halotolerant and Halophilic Fungi Inhabit Brine in Solar Salterns Around the Globe** Food Technol Biotech 52(2, SI):170–179.

Guzvic M. (2013) **The History of Dna Sequencing** J Med Biochem 32 (4):301–312.

Hahn MW, and Höle MG. (2001) **Grazing of protozoa and its effect on populations of aquatic bacteria** FEMS Microbiol Ecol 35(2): 113–121.

Hajek M, Horsak M, Hajkova P, and Dite D. (2006) **Habitat diversity of central European fens in relation to environmental gradients and an effort to standardise fen terminology in ecological studies** Perspect Plant Ecol 8(2):97–114.

Hamels I, Mussche H, Sabbe K, Muylaert K, and Vyverman W. (2004) **Evidence for constant and highly specific active food selection by benthic ciliates in mixed diatoms assemblages** Limnol Oceanogr 49(1): 58–68.

Hamsher SE, LeGresley MM, Martin JL, and Saunders GW. (2013) **A Comparison of Morphological and Molecular-Based Surveys to Estimate the Species Richness of Chaetoceros and Thalassiosira (Bacillariophyta), in the Bay of Fundy** PLoS One 8(10):e73521.

Handcock M, Hunter D, Butts C, Goodreau S, Krivitsky P, Bender-deMoll S, and Morris M. (2016) **statnet: Software Tools for the Statistical Analysis of Network Data**. The Statnet Project (`http://www.statnet.org`). R package version 2016.9.

Harder CB, Ronn R, Brejnrod A, Bass D, Abu Al-Soud W, and Ekelund F. (2016) **Local diversity of heathland Cercozoa explored by in-depth sequencing** ISME J 10(10):2488–2497.

Hardham AR. (2005) ***Phytophthora cinnamomi*** Mol Plant Pathol 6 (6):589–604.

Harrell Jr FE, and Dupont C. (2016) **Hmisc: Harrell Miscellaneous**. R package version 3.17-4.

Harris D, Horwath WR, and van Kessel C. (2001) **Acid fumigation of soils to remove carbonates prior to total organic carbon or carbon-13 isotopic analysis** Soil Sci Soc Am J 65(6):1853–1856.

Harrison MJ. (2005) **Signaling in the arbuscular mycorrhizal symbiosis** Annu Rev Microbiol 59:19–42.

Hawksworth DL, and Wiltshire PEJ. (2011) **Forensic mycology: the use of fungi in criminal investigations** Forensic Sci Int 206(1-3):1–11.

Heal OW. (1964) **Observations on the seasonal and spatial-distribution of testacea (protozoa, rhizopoa) in *Sphagnum*** J Anim Ecol 33(3): 395–412.

Heger TJ, Mitchell EAD, Todorov M, Golemansky V, Lara E, Leander BS, and Pawlowski J. (2010) **Molecular phylogeny of euglyphid testate amoebae (Cercozoa: Euglyphida) suggests transitions between marine supralittoral and freshwater/terrestrial environments are infrequent** Mol Phylogenet Evol 55(1):113–122.

Heger TJ, Pawlowski J, Lara E, Leander BS, Todorov M, Golemansky V, and Mitchell EAD. (2011) **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.

Henssge C, Althaus L, Bolt J, Freislederer A, Haffner HT, Henssge CA, et al. (2000) **Experiences with a compound method for estimating the time since death. I. Rectal temperature nomogram for time since death** Int J Legal Med 113(6):303–319.

Hess S, and Melkonian M. (2013) **The Mystery of Clade X: *Orciraptor* gen. nov and *Viridiraptor* gen. nov are Highly Specialised, Algivorous Amoeboflagellates (Glissomonadida, Cercozoa)** Protist 164(5):706–747.

Hess S, and Melkonian M. (2014) **Ultrastructure of the Algivorous Amoeboflagellate *Viridiraptor invadens* (Glissomonadida, Cercozoa)** Protist 165(5):605–635.

Hess S, Sausen N, and Melkonian M. (2012) **Shedding Light on Vampires: The Phylogeny of Vampyrellid Amoebae Revisited** PLoS One 7(2).

Hittinger CT. (2013) ***Saccharomyces* diversity and evolution: a budding model genus** Trends Genet 29(5):309–317.

Hollander M, and Wolfe DA. (1999) **Nonparametric statistical methods**. Wiley.

Holm S. (1979) **A simple sequentially rejective multiple test procedure** Scand J Stat 6(2):65–70.

Hopkins DW, Wiltshire PEJ, and Turner BD. (2000) **Microbial characteristics of soils from graves: arm investigation at the interface of soil microbiology and forensic science** Appl Soil Ecol 14(3):283–288.

Horswell J, Cordiner SJ, Maas EW, Martin TM, Sutherland BW, Speir TW, et al. (2002) **Forensic comparison of soils by bacterial community DNA profiling** J Forensic Sci 47(2):350–353.

Howard GT, Duos B, and Watson-Horzelski EJ. (2010) **Characterization of the soil microbial community associated with the decomposition of a swine carcass** Int Biodeter Biodegr 64(4):300–304.

Howarth RW, and Marino R. (2006) **Nitrogen as the limiting nutrient for eutrophication in coastal marine ecosystems: Evolving views over three decades** Limnol Oceanogr 51(1):364–376.

Howe AT, Bass D, Vickerman K, Chao EE, and Cavalier-Smith T. (2009) **Phylogeny, Taxonomy, and Astounding Genetic Diversity of Glissomonadida ord. nov., The Dominant Gliding Zooflagellates in Soil (Protozoa: Cercozoa)** Protist 160(2):159–189.

Howe AT, Bass D, Scoble JM, Lewis R, Vickerman K, Arndt H, and Cavalier-Smith T. (2011) **Novel Cultured Protists Identify Deep-branching Environmental DNA Clades of Cercozoa: New Genera *Tremula*, *Micrometopion*, *Minimassisteria*, *Nudifila*, *Peregrinia*** Protist 162(2):332–372.

Hudgins EJ, Liebhold AM, and Leung B. (2017) **Predicting the spread of all invasive forest pests in the United States** Ecol Lett 20(4): 426–435.

Hungate RE. (2013) **The rumen and its microbes**. Elsevier.

Hutchinson GE. (1961) **The Paradox of the Plankton** Am Nat 95(882): 137–145.

IUSS ISRIC, and others . (1999) **World reference base for soil resources**.

Izaguirre I, O'farrell I, Unrein F, Sinistro R, dos Santos Afonso M, and Tell G. (2004) **Algal assemblages across a wetland, from a shallow lake to relictual oxbow lakes (Lower Paraná River, South America)** Hydrobiologia 511(1):25–36.

Jaatinen K, Fritze H, Laine J, and Laiho R. (2007) **Effects of short- and long-term water-level drawdown on the populations and activity of aerobic decomposers in a boreal peatland** Glob Change Biol 13 (2):491–510.

Janzen DH. (1970) **Herbivores and the Number of Tree Species in Tropical Forests** Am Nat 104(940):501+.

Jardillier L, Zubkov MV, Pearman J, and Scanlan DJ. (2010) **Significant CO_2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean** 4(9):1180–1192.

Jassey VEJ, Shimano S, Dupuy C, Toussaint ML, and 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 C-13 and N-15 isotopic analyses** Protist 163(3):451–464.

Jassey VEJ, Signarbieux C, Haettenschwiler S, Bragazza L, Buttler A, Delarue F, et al. (2015) **An unexpected role for mixotrophs in the response of peatland carbon cycling to climate warming** Sci Rep-UK 5.

Jeger MJ, and Pautasso M. (2008) **Comparative epidemiology of zoosporic plant pathogens** Eur J Plant Pathol 122(1):111–126.

Jennings DH. (1983) **Some Aspects of the Physiology and Biochemistry of Marine Fungi** Biol Rev Camb Philos Soc 58(3):423–459.

Jiang L, and Pu Z. (2009) **Different effects of species diversity on temporal stability in single-trophic and multitrophic communities** Am Nat 174(5):651–659.

Johnke J, Cohen Y, de Leeuw M, Kushmaro A, Jurkevitch E, and Chatzinotas A. (2014) **Multiple micro-predators controlling bacterial communities in the environment** Curr Opin Biotechnol 27:185–190.

Johnson MD. (2011)a **The acquisition of phototrophy: adaptive strategies of hosting endosymbionts and organelles** Photosynth Res 107 (1):117–132.

Johnson MD. (2011)b **Acquired Phototrophy in Ciliates: A Review of Cellular Interactions and Structural Adaptations** J Eukaryot Microbiol 58(3):185–195.

Jones MDM, Forn I, Gadelha C, Egan MJ, Bass D, Massana R, and Richards TA. (2011) **Discovery of novel intermediate forms redefines the fungal tree of life** Nature 474(7350):200–U234.

Jurgens K, and Matz C. (2002) **Predation as a shaping force for the phenotypic and genotypic composition of planktonic bacteria** Anton Leeuw Int J G 81(1-4):413–434.

Karimi B, Meyer C, Gilbert D, and Bernard N. (2016) **Air pollution below WHO levels decreases by 40 % the links of terrestrial microbial networks** Environ Chem Lett 14(4):467–475.

Karimi B, Maron PA, Chemidlin-Prevost Boure N, Bernard N, Gilbert D, and Ranjard L. (2017) **Microbial diversity and ecological networks as indicators of environmental quality** Environ Chem Lett.

Katz LA. (2012) **Origin and Diversification of Eukaryotes**. In Gottesman S, Harwood C, and Schneewind O (editors), Annual Review of Microbiology volume 66 pages 411–427.

Kivinen E, and Pakarinen P. (1981) **Geographical distribution of peat resources and major peatland complex types in the world**. Suomalainen tiedeakatemia.

Koch TA, and Ekelund F. (2005) **Strains of the heterotrophic flagellate *Bodo designis* from different environments vary considerably with respect to salinity preference and SSU rRNA gene composition** Protist 156(1):97–112.

Lachat T, Pauli D, Gonseth Y, Klaus G, Scheidegger C, and Vittoz P. (2011) **Evolution de la biodiversité en Suisse depuis 1900: Avons nous touché le fond?** Haupt Verlag.

Lamers LPM, Bobbink R, and Roelofs JGM. (2000) **Natural nitrogen filter fails in polluted raised bogs** Glob Change Biol 6(5):583–586.

Landsberg JJ, and Sands P. (2010) **Physiological ecology of forest production: principles, processes and models** volume 4. Academic Press.

Langer MR, Weinmann AE, Loetters S, Bernhard JM, and Roedder D. (2013) **Climate-Driven Range Extension of Amphistegina (Protista, Foraminiferida): Models of Current and Predicted Future Ranges** PLoS One 8(2).

Lara E, and Acosta-Mercado D. (2012) **A molecular perspective on ciliates as soil bioindicators** Eur J Soil Biol 49(SI):107–111.

Lara E, and Belbahri L. (2011) **SSU rRNA reveals major trends in oomycete evolution** Fungal Divers 49(1):93–100.

Lara E, and Gomaa F. (2017) **Symbiosis between Testate Amoebae and Photosynthetic Organisms**. In Grube M, Seckbach J, and Miggia L (editors), Algal and cyanobacteria symbioses pages 399–417. Sorle Scientific Europe.

Lara E, Berney C, Ekelund F, Harms H, and Chatzinotas A. (2007)a **Molecular comparison of cultivable protozoa from a pristine and a polycyclic aromatic hydrocarbon polluted site** Soil Biol Biochem 39 (1):139–148.

Lara E, Heger TJ, Mitchell EAD, Meisterfeld R, and Ekelund F. (2007)b **SSU rRNA reveals a sequential increase in shell complexity among the euglyphid testate amoebae (Rhizaria : Euglyphida)** Protist 158(2): 229–237.

Lara E, Moreira D, and Lopez-Garcia P. (2010) **The Environmental Clade LKM11 and *Rozella* Form the Deepest Branching Clade of Fungi** Protist 161(1):116–121.

Lara E, Seppey CVW, Gonzalez Garraza G, Singer D, Quiroga MV, and Mataloni G. (2015) **Planktonic eukaryote molecular diversity: discrimination of minerotrophic and ombrotrophic peatland pools in Tierra del Fuego (Argentina)** J Plankton Res 37(3):645–655.

Lara E, Roussel-Delif L, Fournier B, Wilkinson DM, and Mitchell EAD. (2016) **Soil microorganisms behave like macroscopic organisms: patterns in the global distribution of soil euglyphid testate amoebae** J Biogeogr 43(3):520–532.

Legendre P, and Gallagher ED. (2001) **Ecologically meaningful transformations for ordination of species data** Oecologia 129(2):271–280.

Lejzerowicz F, Pawlowski J, Fraissinet-Tachet L, and Marmeisse R. (2010) **Molecular evidence for widespread occurrence of Foraminifera in soils** Environ Microbiol 12(9):2518–2525.

Lenarczyk J. (2015) ***Pediastrum* MEYEN sensu lato (Chlorophyceae) in the phytoplankton of lowland and upland water bodies of Central Europe (Poland)** Fottea 15(2):165–177.

Lentendu G, Wubet T, Chatzinotas A, Wilhelm C, Buscot F, and Schlegel M. (2014) **Effects of long-term differential fertilization on eukaryotic microbial communities in an arable soil: a multiple barcoding approach** Mol Ecol 23(13):3341–3355.

Lesaulnier C, Papamichail D, McCorkle S, Ollivier B, Skiena S, Taghavi S, et al. (2008) **Elevated atmospheric CO_2 affects soil microbial diversity associated with trembling aspen** Environ Microbiol 10:926–941.

Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, et al. (2015) **Determinants of community structure in the global plankton interactome** Science 348(6237).

Little AF, Van Oppen MJH, and Willis BL. (2004) **Flexibility in algal endosymbioses shapes growth in reef corals** Science 304(5676):1492–1494.

Lopez-Garcia P, Rodriguez-Valera F, Pedros-Alio C, and Moreira D. (2001) **Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton** Nature 409(6820):603–607.

Luftenegger G, Petz W, Foissner W, and Adam H. (1988) **The efficiency of a direct counting method in estimating the numbers of microscopic soil organisms** Pedobiologia 31(1-2):95–101.

Magoc T, and Salzberg SL. (2011) **FLASH: fast length adjustment of short reads to improve genome assemblies** Bioinformatics 27(21): 2957–2963.

Mahé F, Rognes T, Quince C, de Vargas C, and Dunthorn M. (2014) **Swarm: robust and fast clustering method for amplicon-based studies.** PeerJ 2:e539.

Mahé F, Rognes T, Quince C, de Vargas C, and 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, et al. (2017) **Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests** Nat Ecol Evol.

Marande W, Lopez-Garcia P, and Moreira D. (2009) **Eukaryotic diversity and phylogeny using small- and large-subunit ribosomal RNA genes from environmental samples** Environ Microbiol 11(12):3179–3188.

Masinova T, Bahnmann BD, Vetrovsky T, Tomsovsky M, Merunkova K, and Baldrian P. (2017) **Drivers of yeast community composition in the litter and soil of a temperate forest** FEMS Microbiol Ecol 93(2).

Massana R, del Campo J, Sieracki ME, Audic S, and Logares R. (2014) **Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles** ISME J 8(4):854–866.

May RM. (1972) **Will a Large Complex System be Stable** Nature 238 (5364):413–&.

McFadden GI, and Melkonian M. (1986) **Use of hepes buffer for microalgal culture media and fixation for electron-microscopy** Phycologia 25(4):551–557.

Meisterfeld R. (2000)a **Order Arcellinida Kent, 1880**. In Lee J, Leedale G, and Bradbury P (editors), An Illustrated Guide to the Protozoa pages 827–859. Society of Protozoologists Lawrence, Kansas.

Meisterfeld R. (2000)b **Testate amoebae with filopodia**. In Lee J, Leedale G, and Bradbury P (editors), An Illustrated Guide to the Protozoa pages 1054–1084. Society of Protozoologists Lawrence, Kansas.

Melis C, Selva N, Teurlings I, Skarpe C, Linnell JDC, and Andersen R. (2007) **Soil and vegetation nutrient response to bison carcasses in Bialeowieza Primeval Forest, Poland** Ecol Res 22(5):807–813.

Mendoza L, Taylor JW, and Ajello L. (2002) **The class Mesomycetozoea: A group of microorganisms at the animal-fungal boundary** Annu Rev Microbiol 56:315–344.

Middelburg JJ, Vlug T, Jaco F, and Van der Nat WA. (1993) **Organic matter mineralization in marine systems** Glob Planet Change 8 (1-2):47–58.

Mills A, and Adl MS. (2006) **The effects of land use intensification on soil biodiversity in the pasture** Can J Plant Sci 86(5, SI):1339–1343.

Mitchell EAD. (2004) **Response of testate amoebae (protozoa) to N and P fertilization in an arctic wet sedge tundra** Arct Antarct Alp Res 36(1):78–83.

Mitchell EAD, Buttler AJ, Warner BG, and Gobat JM. (1999) **Ecology of testate amoebae (Protozoa : Rhizopoda) in Sphagnum peatlands in the Jura mountains, Switzerland and France** Ecoscience 6(4):565–576.

Mitchell EAD, Borcard D, Buttler AJ, Grosvernier P, Gilbert D, and Gobat JM. (2000) **Horizontal distribution patterns of testate amoebae (Protozoa) in a *Sphagnum magellanicum* carpet** Microbial Ecol 39 (4):290–300.

Mitchell EAD, Charman DJ, and Warner BG. (2008) **Testate amoebae analysis in ecological and paleoecological studies of wetlands: past, present and future** Biodivers Conserv 17(9):2115–2137.

Mitra A, Flynn KJ, Tillmann U, Raven JA, Caron D, Stoecker DK, et al. (2016) **Defining Planktonic Protist Functional Groups on Mechanisms for Energy and Nutrient Acquisition: Incorporation of Diverse Mixotrophic Strategies** Protist 167(2):106–120.

Miyaoka Y, Hatamoto M, Yamaguchi T, and Syutsubo K. (2017) **Eukaryotic Community Shift in Response to Organic Loading Rate of an Aerobic Trickling Filter (Down-Flow Hanging Sponge Reactor) Treating Domestic Sewage** Microbial Ecol 73(4):801–814.

Mora C, Tittensor DP, Adl S, Simpson AGB, and Worm B. (2011) **How Many Species Are There on Earth and in the Ocean?** PLos Biol 9 (8).

Moreira D, and Lopez-Garcia P. (2002) **The molecular ecology of microbial eukaryotes unveils a hidden world** Trends Microbiol 10(1): 31–38.

Moreno LI, Mills D, Fetscher J, John-Williams K, Meadows-Jantz L, and McCord B. (2011) **The application of amplicon length heterogeneity PCR (LH-PCR) for monitoring the dynamics of soil microbial communities associated with cadaver decomposition** J Microbiol Meth 84 (3):388–393.

Moschen S, Higgins J, Di Rienzo JA, Heinz RA, Paniego N, and Fernandez P. (2016) **Network and biosignature analysis for the integration of transcriptomic and metabolomic data to characterize leaf senescence process in sunflower** BMC Bioinformatics 17(5).

Mulot M, Marcisz K, Grandgirard L, Lara E, Kosakyan A, Robroek BJM, et al. (2017) **Genetic Determinism vs. Phenotypic Plasticity in Protist Morphology** J Eukaryot Microbiol.

Murase J, Noll M, and Frenzel P. (2006) **Impact of protists on the activity and structure of the bacterial community in a rice field soil** Appl Environ Microbiol 72(8):5436–5444.

Nelson DM, Treguer P, Brzezinski MA, Leynaert A, and Queguiner B. (1995) **Production and Dissolution of Biogenic Silica in the Ocean - Revised Global Estimates, Comparison with Regional Data and Relationship to Biogenic Sedimentation** Global Biogeochem Cycles 9 (3):359–372.

Neuhauser S, Kirchmair M, Bulman S, and Bass D. (2014) **Cross-kingdom host shifts of phytomyxid parasites** BMC Evol Biol 14.

Nielsen LT, Hallegraeff GM, Wright SW, and Hansen PJ. (2011) **Effects of experimental seawater acidification on an estuarine plankton community** Aquat Microb Ecol 65(3):271–285.

Novarino G, and Lucas I. (1995) **Zoological classification system of cryptomonads** Acta Protozool 34:173–173.

Noyce GL, Fulthorpe R, Gorgolewski A, Hazlett P, Honghi T, and Basiliko N. (2016) **Soil microbial responses to wood ash addition and forest fire in managed Ontario forests** Appl Soil Ecol 107:368–380.

Oksanen J, Blanchet GF, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. (2013) **vegan: Community Ecology Package**. URL http://CRAN.R-project.org/package=vegan. R package version 2.0-10.

Pajares S, and Bohannan BJM. (2016) **Ecology of Nitrogen Fixing, Nitrifying, and Denitrifying Microorganisms in Tropical Forest Soils** Front Microbiol 7.

Pan J, del Campo J, and Keeling PJ. (2016) **Reference Tree and Environmental Sequence Diversity of Labyrinthulomycetes** J Eukaryot Microbiol 64(1):88–96.

Passarini MRZ, Miqueletto PB, de Oliveira VM, and Sette LD. (2015) **Molecular diversity of fungal and bacterial communities in the marine sponge Dragmacidon reticulatum** J Basic Microbiol 55(2):207–220.

Pawlowski J. (2014). **Protist Evolution and Phylogeny**. John Wiley & Sons, Ltd.

Pawlowski J, Bolivar I, Fahrni J, de Vargas C, and Bowser SS. (1999) **Naked foraminiferans revealed** Nature 399(6731):27.

Pawlowski J, Christen R, Lecroq B, Bachar D, Shahbazkia HR, Amaral-Zettler L, and Guillou L. (2011) **Eukaryotic Richness in the Abyss: Insights from Pyrotag Sequencing** PLoS One 6(4).

Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, et al. (2012) **CBOL Protist Working Group: Barcoding Eukaryotic Richness beyond the Animal, Plant, and Fungal Kingdoms** PLos Biol 10(11).

Pawlowski J, Lejzerowicz F, Apotheloz-Perret-Gentil L, Visco J, and Esling P. (2016) **Protist metabarcoding and environmental biomonitoring: Time for change** Eur J Protistol 55(A, SI):12–25.

Payne JA. (1965) **A summer carrion study of the baby pig *Sus scrofa* Linnaeus** Ecology 46(5):592–602.

Payne R, Gauci V, and Charman DJ. (2010) **The Impact of Simulated Sulfate Deposition on Peatland Testate Amoebae** Microbial Ecol 59(1):76–83.

Payne RJ, and Mitchell EAD. (2009) **How many is enough? Determining optimal count totals for ecological and palaeoecological studies of testate amoebae** J Paleolimnol 42(4):483–495.

Payne RJ, Dise NB, Stevens CJ, Gowing DJ, and Partners B. (2013) **Impact of nitrogen deposition at the species level** Proc Natl Acad Sci USA 110(3):984–987.

Pearson WR. (2000) **Flexible Sequence Similarity Searching with the FASTA3 Program Package**. In Misener S, and Krawetz S (editors), Bioinformatics Methods and Protocols pages 185–218. Humana Press.

Petz W, and Foissner W. (1989) **The effect of Mancozeb and Lindane on the soil microfauna of a spruce forest - a field-study using a completely randomized block design** Biol Fert Soils 7(3):225–231.

Petz W, Foissner W, and Adam H. (1985) **Culture, food selection and growth-rate in the mycophagous ciliate *Grossglockneria-acuta* Foissner, 1980 - 1st evidence of autochthonous soil ciliates** Soil Biol Biochem 17(6):871–875.

Pichrtova M, and Nemcova Y. (2011) **Effect of temperature on size and shape of silica scales in *Synura petersenii* and *Mallomonas tonsurata* (Stramenopiles)** Hydrobiologia 673(1):1–11.

Pohlert T. (2014) **The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR)**. URL http://CRAN.R-project.org/package=PMCMR. R package version 4.1.

Pollierer MM, Langel R, Scheu S, and Maraun M. (2009) **Compartmentalization of the soil animal food web as indicated by dual analysis of stable isotope ratios (N-15/N-14 and C-13/C-12)** Soil Biol Biochem 41(6):1221–1226.

Posch T, Eugster B, Pomati F, Pernthaler J, Pitsch G, and Eckert EM. (2015) **Network of interactions between ciliates and phytoplankton during spring** Front Microbiol 6.

Prangnell J, and McGowan G. (2009) **Soil temperature calculation for burial site analysis** Forensic Sci Int 191(1-3):104–109.

Princiotta SD, Smith BT, and Sanders RW. (2016) **Temperature-Dependent Phagotrophy and Phototrophy in a Mixotrophic Chrysophyte** J Phycol 52(3):432–440.

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, and Gloeckner FO. (2007) **SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB** Nucleic Acids Res 35(21):7188–7196.

Pushkareva E, Johansen JR, and Elster J. (2016) **A review of the ecology, ecophysiology and biodiversity of microalgae in Arctic soil crusts** Polar Biol 39(12):2227–2240.

Rabajante JF, Tubay JM, Ito H, Uehara T, Kakishima S, Morita S, et al. (2016) **Host-parasite Red Queen dynamics with phase-locked rare genotypes** Sci Adv 2(3):e1501548.

Raberg L, Alacid E, Garces E, and Figueroa Rosa. (2014) **The potential for arms race and Red Queen coevolution in a protist host-parasite system** 4(24):4775–4785.

Ramirez KS, Lauber CL, Knight R, Bradford MA, and Fierer N. (2010) **Consistent effects of nitrogen fertilization on soil bacterial communities in contrasting systems** Ecology 91(12):3463–3470.

R_Core_Team . (2013) **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing Vienna, Austria. URL http://www.R-project.org/.

R_Core_Team . (2014) **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing Vienna, Austria. URL http://www.R-project.org/.

Reczuga MK, Seppey CVW, Szelecz I, Fournier B, Singer D, Lara E, et al. (in prep.) **Temporal patterns of soil micro-eukaryotic diversity beneath decomposing pig cadavers as assessed by high throughput sequencing**.

Richards TA, Vepritskiy AA, Gouliamova DE, and Nierzwicki-Bauer SA. (2005) **The molecular diversity of freshwater picoeukaryotes from an oligotrophic lake reveals diverse, distinctive and globally dispersed lineages** Environ Microbiol 7(9):1413–1425.

Rodriguez WC, and Bass WM. (1985) **Decomposition of buried bodies and methods that may aid in their location** J Forensic Sci 30(3): 836–852.

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

Rose JM, and Caron DA. (2007) **Does low temperature constrain the growth rates of heterotrophic protists? Evidence and implications for algal blooms in cold waters** Limnol Oceanogr 52(2):886–895.

Rydin H, and Jeglum JK. (2013) **The Biology of Peatlands**. Biology of Habitats Series. OUP Oxford 2nd edition.

Saad JF, Unrein F, Tribelli PM, Lopez N, and Izaguirre I. (2016) **Influence of lake trophic conditions on the dominant mixotrophic algal assemblages** J Plankton Res 38(4):818–829.

Sachs JL, Essenberg CJ, and Turcotte MM. (2011) **New paradigms for the evolution of beneficial infections** Trends Ecol Evol 26(4):202–209.

Santibanez A PA, Kohshima S, Scheihing A RA, Silva R R, Jaramillo M JI, Labarca P PJ, and Casassa R G. (2011) **First Record of Testate Amoebae on Glaciers and Description of a New Species *Puytoracia jenswendti* nov sp (Rhizaria, Euglyphida)** Acta Protozool 50(1): 1–14.

Santoyo G, Hernandez-Pacheco C, Hernandez-Salmeron J, and Hernandez-Leon R. (2017) **The role of abiotic factors modulating the plant-microbe-soil interactions: toward sustainable agriculture. A review** Span J Agric Res 15(1).

Sapp M, Schwaderer AS, Wiltshire KH, Hoppe H, Gerdts G, and Wichels A. (2007) **Species-specific bacterial communities in the phycosphere of microalgae?** Microbial Ecol 53(4):683–699.

Schiaffino MR, Lara E, Fernandez LD, Balague V, Singer D, Seppey CVW, et al. (2016) **Microbial eukaryote communities exhibit robust biogeographical patterns along a gradient of Patagonian and Antarctic lakes** Environ Microbiol 18(12):5249–5264.

Schlegel M. (1994) **Molecular Phylogeny of Eukaryotes** Trends Ecol Evol 9(9):330–335.

Schmidt O, Dyckmans J, and Schrader S. (2016) **Photoautotrophic microorganisms as a carbon source for temperate soil invertebrates** Biol Lett 12(1).

Scholz B, and Liebezeit G. (2012)a **Growth responses of 25 benthic marine Wadden Sea diatoms isolated from the Solthorn tidal flat (southern North Sea) in relation to varying culture conditions** Diatom Res 27 (1):65–73.

Scholz B, and Liebezeit G. (2012)b **Compatible solutes in three marine intertidal microphytobenthic Wadden Sea diatoms exposed to different salinities** Eur J Phycol 47(4):393–407.

Seppey CVW, Fournier B, Szelecz I, Singer D, Mitchell EAD, and Lara E. (2016) **Response of soil euglyphid testate amoebae (Rhizaria: Cercozoa) to pig cadavers assessed by high-throughput sequencing** Int J Legal Med 130(2):551–562.

Seppey CVW, Singer D, Dumack K, Fournier B, Belbahri L, Mitchell EAD, and Lara E. (2017) **Distribution patterns of soil microbial eukaryotes suggests widespread algivory by phagotrophic protists as an alternative pathway for nutrient cycling** Soil Biol Biochem.

Shendure J, and Ji H. (2008) **Next-generation DNA sequencing** Nat Biotechnol 26(10):1135–1145.

Simpson AGB, Stevens JR, and Lukeš J. (2006) **The evolution and diversity of kinetoplastid flagellates** Trends Parasitol 22(4):168–174.

Singer D, Kosakyan A, Pillonel A, Mitchell EAD, and Lara E. (2015) **Eight species in the *Nebela collaris* complex: *Nebela gimlii* (Arcellinida, Hyalospheniidae), a new species described from a Swiss raised bog** Eur J Protistol 51(1):79–85.

Singer D, Lara E, Steciow MM, Seppey CVW, Paredes N, Pillonel A, et al. (2016) **High-throughput sequencing reveals diverse oomycete communities in oligotrophic peat bog micro-habitat** Fungal Ecol 23:42–47.

Singh BN. (1946) **A method of estimating numbers of soil protozoa, especially amoebae, based on their differential feeding on bacteria** Ann Appl Biol 33(1):112.

Singh BN. (1955) **Culturing soil protozoa and estimating their numbers in soil**. In Kevan (editor), Soil Zoology pages 403–411. Butterworths Scientific Publications.

Smetacek V. (1999) **Diatoms and the ocean carbon cycle** Protist 150 (1):25–32.

Smith HG, Bobrov A, and Lara E. (2008) **Diversity and biogeography of testate amoebae** Biodivers Conserv 17(2):329–343.

Stamatakis A. (2006) **RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models** Bioinformatics 22(21):2688–2690.

Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner H, and Richards TA. (2010) **Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water** Mol Ecol 19(s1):21–31.

Stoecker DK. (1998) **Conceptual models of mixotrophy in planktonic protists and some ecological and evolutionary implications** Eur J Protistol 34(3):281–290.

Stokes KL, Forbes SL, and Tibbett M. (2013) **Human Versus Animal: Contrasting Decomposition Dynamics of Mammalian Analogues in Experimental Taphonomy** J Forensic Sci 58(3):583–591.

Swanson JA. (2013) **The noodle defense** J Cell Biol 203(6):871–873.

Swindles GT, and Ruffell A. (2009) **A preliminary investigation into the use of testate amoebae for the discrimination of forensic soil samples** Sci Justice 49(3):182–190.

Szelecz I, Fournier B, Seppey CVW, Amendt J, and Mitchell EAD. (2014) **Can soil testate amoebae be used for estimating the time since death? A field experiment in a deciduous forest** Forensic Sci Int 236:90–98.

Taberlet P, Coissac E, Pompanon F, Brochmann C, and Willerslev E. (2012) **Towards next-generation biodiversity assessment using DNA metabarcoding** Mol Ecol 21(8):2045–2050.

Tarnawski SE, and Lara E. (2015) **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.

Tice AK, Silberman JD, Walthall AC, Le KN, Spiegel FW, and Brown MW. (2016) ***Sorodiplophrys stercorea*: Another Novel Lineage of Sorocarpic Multicellularity.** J Eukaryot Microbiol 63(5):623–628.

Tittensor DP, Mora C, Jetz W, Lotze HK, Ricard D, Vanden Berghe E, and Worm B. (2010) **Global patterns and predictors of marine biodiversity across taxa** Nature 466(7310).

Todorov M, and Golemansky V. (1992) **Effect of pesticides Fundasol, Fuzamicin and Lavendotricin on the growth of laboratory cultures of Protozoa** Acta Zool Bulg 45:20–25.

Towne EG. (2000) **Prairie vegetation and soil nutrient responses to ungulate carcasses** Oecologia 122(2):232–239.

Tsui CKM, Marshall W, Yokoyama R, Honda D, Lippmeier JC, Craven KD, et al. (2009) **Labyrinthulomycetes phylogeny and its implications for the evolutionary loss of chloroplasts and gain of ectoplasmic gliding** Mol Phylogenet Evol 50(1):129–140.

Tsyganov AN, Nijs I, and 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(2):237–248.

Turner JT. (2004) **The importance of small planktonic copepods and their roles in pelagic marine food webs** Zool Stud 43(2):255–266.

Tylianakis JM, Laliberte E, Nielsen A, and Bascompte J. (2010) **Conservation of species interaction networks** Biol Conserv 143(10):2270–2279.

Unrein F, Gasol JM, Not F, Forn I, and Massana R. (2014) **Mixotrophic haptophytes are key bacterial grazers in oligotrophic coastal waters** ISME J 8(1):164–176.

Upreti DK, Divakar PK, Shukla V, and Bajpai R. (2015) **Recent Advances in Lichenology: Modern Methods and Approaches in Lichen Systematics and Culture Techniques**. Number vol. 2. Springer India.

Valentini A, Pompanon F, and Taberlet P. (2009) **DNA barcoding for ecologists** Trends Ecol Evol 24(2):110–117.

van den Brink PJ, and Ter Braak CJF. (1999) **Principal response curves: Analysis of time-dependent multivariate responses of biological community to stress** Environ Toxicol Chem 18(2):138–148.

van der Heijden M, Klironomos J, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T, et al. (1998) **Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity** Nature 396 (6706):69–72.

van der Wal A, Geydan TD, Kuyper TW, and de Boer W. (2013) **A thready affair: linking fungal diversity and community dynamics to terrestrial decomposition processes** FEMS Microbiol Rev 37(4): 477–494.

van Hannen EJ, Mooij WM, van Agterveld MP, Gons HJ, and Laanbroek HJ. (1999) **Detritus-dependent development of the microbial community in an experimental system: Qualitative analysis by denaturing gradient gel electrophoresis** Appl Environ Microbiol 65(6):2478–2484.

van Herk CM. (2001) **Bark pH and susceptibility to toxic air pollutants as independent causes of changes in epiphytic lichen composition in space and time** Lichenologist 33(5):419–441.

Vass AA, Bass WM, Wolt JD, Foss JE, and Ammons JT. (1992) **Time Since Death Determinations of Human Cadavers Using Soil Solution** J Forensic Sci 37(5):1236–1253.

Vaulot D, Eikrem W, Viprey M, and Moreau H. (2008) **The diversity of small eukaryotic phytoplankton (<= 3 μm) in marine ecosystems** FEMS Microbiol Rev 32(5):795–820.

Verni F, and Gualtieri P. (1997) **Feeding behaviour in ciliated protists** Micron 28(6):487–504.

Votýpka J, Modrý D, Oborník M, Šlapeta J, and Lukeš Julius. (2017). **Apicomplexa** pages 1–58. Springer International Publishing Cham.

Wall DH, Ducarme X, and Lebrun P. (2001) **Soil biodiversity**. In Chapin F, Sala O, and Huber-Sannwald E (editors), Global Biodiversity in a Changing Environment: Scenarios for the 21st Century Ecological Studies. Springer New York.

Wallace AR. (1876) **The Geographical Distribution of Animals**. MacMillan and Co. London.

Wanner M, and Elmer M. (2009) **"Hot spots" on a New Soil Surface - How Do Testate Amoebae Settle Down?** Acta Protozool 48(3): 281–289.

Wanner M, Elmer M, Kazda M, and Xylander WER. (2008) **Community assembly of terrestrial testate amoebae: How is the very first beginning characterized?** Microbial Ecol 56(1):43–54.

Watanabe H, Noda H, Tokuda G, and Lo N. (1998) **A cellulase gene of termite origin** Nature 394(6691):330–331.

Weisse T, Moser M, Scheffel U, Stadler P, Berendonk T, Weithoff G, and Berger H. (2013) **Systematics and species-specific response to pH of *Oxytricha acidotolerans* sp nov and *Urosomoida* sp (Ciliophora, Hypotricha) from acid mining lakes** Eur J Protistol 49(2):255–271.

Whittaker RH. (1969) **New Concepts of Kingdoms of Organisms** Science 163(3863):150+.

Worley AC, Raper KB, and Hohl M. (1979) ***Fonticula alba*** **- New Cellular Slime-mold (Acrasiomycetes)** Mycologia 71(4):746–760.

Wu QL, Boenigk J, and Hahn MW. (2004) **Successful predation of filamentous bacteria by a nanoflagellate challenges current models of flagellate bacterivory** Appl Environ Microbiol 70(1):332–339.

Wyss C, and Cherix D. (2006) **Traité d'entomologie forensique. Les insectes sur la scène de crime**. Presses polytechniques et universitaires romandes Lausanne.

Xinyao L, Miao S, Yonghong L, Yin G, Zhongkai Z, Donghui W, et al. (2006) **Feeding characteristics of an amoeba (Lobosea : *Naegleria*) grazing upon cyanobacteria: Food selection, ingestion and digestion progress** Microbial Ecol 51(3):315–325.

Xu H, Song W, Warren A, Al-Rasheid KAS, Al-Farraj SA, Gong J, and Hu X. (2008) **Planktonic protist communities in a semi-enclosed mariculture pond: structural variation and correlation with environmental conditions** J Mar Biol Assoc UK 88(07):1353–1362.

Yashiro E, Pinto-Figueroa E, Buri A, Spangenberg JE, Adatte T, Niculita-Hirzel H, et al. (2016) **Local Environmental Factors Drive Divergent Grassland Soil Bacterial Communities in the Western Swiss Alps** Appl Environ Microbiol 82(21):6303–6316.

Zaric S, Schulz M, and Mulitza S. (2006) **Global prediction of planktic foraminiferal fluxes from hydrographic and productivity data** Biogeosciences 3(2):187–207.

Zhang T, Jia R, and Yu LY. (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 Ecol 23:156–163.

Zimmermann NE, and Kienast F. (1999) **Predictive mapping of alpine grasslands in Switzerland: Species versus community approach** J Veg Sci 10(4):469–482.

Zimmermann NE, Edwards TC, Moisen GG, Frescino TS, and Blackard JA. (2007) **Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah** J Appl Ecol 44(5):1057–1067.

Rue Emile-Argand 11
2000 Neuchâtel
tel +41-77-448-9445
+41-32-718-3108
e-mail christophe.seppey@unine.ch

Christophe Seppey

Education

08.2013 - 08.2017 **PhD in Biology**
(Switzerland) University of Neuchâtel: Laboratory of Soil Biodiversity *Study of soil micro-eukaryotes by metabarcoding*
Supervisors:
- Prof. Edward A.D. Mitchell
- Dr. Enrique Lara

09.2011 - 07.2013 **M.Sc. in Biogeosciences**
(Switzerland) University of Neuchâtel: Laboratory of Soil Biodiversity *Response of forest soil euglyphid testate amoebae (Rhizaria: Cercozoa) to pig cadavers assessed by high-throughput sequencing (Published in Int J Legal Med)*
Supervisors:
- Prof. Edward A.D. Mitchell
- Dr. Enrique Lara

09.2007 - 07.2011 **B.Sc. in Biology**
(Switzerland) University of Lausanne

Professional experience

08.2013 - 08.2017 **PhD assistant**
(Switzerland) University of Neuchâtel: Laboratory of Soil Biodiversity
- Organisation and teaching of practicals in soil organisms biology for master students
- Redaction of the abstract booklet of the 35th meeting of the German Society of Protozoology (2016)
- Management of the laboratory web page

04.2013 - 10.2013 **Scientific collaborator (CADAVER: SNF project no. 31003A-141188/1)**
(Switzerland) University of Neuchâtel: Laboratory of Soil Biodiversity
- Field work (field preparation, sampling)
- Laboratory (NH^{+}, NO_3^{2-} concentration analyses)
- Statistical analysis

04.2010 - 07.2011 **Field work assistant**
(Switzerland) University of Lausanne
- Rodents and owls handling and measurments

08, 10.2010 **Biodiversity inventory**
(Namibia) BRinK Organisation (www.kuzikus-namibia.de)
- Rodents and reptiles handling and measurment

Competencies

Languages

mother tongue **French** — scholar **Spanish**

fluent **English** — scholar **German**

Informatic

programing **Perl, C++, Bash** — statistics **R**

redaction **LaTeX** — office **Libreoffice, Microsoft office**

text edition **Vim** — other **cluster computing**

Teaching

- Basic microscopy
- Protist dilution culture
- Technique to assess soil quality based on arthropods population
- Basic bioinformatics (pairwise and multiple sequence alignments, metabarcoding)
- Invertebrate morphology (mussel, shrimp, squid)
- Statistics (e.g. basic, numerical ecology)

Publication list

Published

2017 Seppey CVW, Singer D, Dumack K, Belbahri L, Mitchell EAD & Lara E. **Distribution patterns of soil microbial eukaryotes suggests widespread algivory by phagotrophic protists as an alternative pathway for nutrient cycling**. *Soil Biol Biochem*

2017 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-Parra 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. **Soil protistology rebooted: 30 fundamental questions to start with**. *Soil Biol Biochem*

2017 Mahe F, de Vargas C, Bass D, Czech L, Stamatakis A, Lara E, Singer D, Mayor J, Bunge J, Sernaker S, Siemensmeyer T, Trautmann I, Romac S, Berney C, Kozlov A, Mitchell EAD, Seppey CVW, Egge E, Lentendu G, Wirth R, Trueba G, & Dunthorn M. **Soil Protists in Three Neotropical Rainforests are Hyperdiverse and Dominated by Parasites**. *Nat Ecol Evol*

2017 Blandenier Q, Seppey CVW, Singer D, Mitchell EAD, Vlimant M, Simon A & Lara E. ***Mycamoeba gemmipara* nov. gen., nov. sp., the First Cultured Member of the Environmental Dermamoebidae Clade LKM74 and its Unusual Life Cycle**. *J Eukaryot Microbiol*

2016 Schiaffino MR, Lara E, Fernández LD, Balagué V, Singer D, Seppey CVW, Massana R & Izaguirre I. **Microbial eukaryote communities from Patagonian-Antarctic gradient of lakes evidence robust biogeographical patterns**. *Environ Microbiol*

2016 Szelecz I, Sorge F, Seppey CVW, Mulot M, Steel H, Neilson R, Griffiths BS, Amendt J & Mitchell EAD. **Effects of decomposing cadavers on soil nematode communities over a one-year period**. *Soil Biol Biochem*: 103

2016 Seppeyc CVW, Fournier B, Szelecz I, Singer D, Mitchell EAD & Lara E. **Response of forest soil euglyphid testate amoebae (Rhizaria: Cercozoa) to pig cadavers assessed by high-throughput sequencing**. *Int J Legal Med*: 103, 2

2016 Singer D, Lara E, Steciow MM, Seppey CVW, Paredes N, Pillonel A, Oszako T & Belbahri L. **High-throughput sequencing reveals diverse oomycete communities in oligotrophic peat bog micro-habitat**. *Fungal Ecol*: 23

2015 Lara E, Seppey CVW, Gonzalez Garraza G, Singer D, Quiroga MV & Mataloni G. **Planktonic eukaryote molecular diversity: discrimination of minerotrophic and ombrotrophic peatland pools in Tierra del Fuego (Argentina)**. *J Plankton Res*: 37, 3

2014 Szelecz I, Fournier B, Seppey C, Amendt J, & Mitchell E. **Can soil testate amoebae be used for estimating the time since death? A field experiment in a deciduous forest**. *Forensic Sci Int*: 236

Accepted or submited

Submitted Singer D, Kosakyan A, Seppey CVW, Pillonel A, Fernández LD, Fontaneto D, Mitchell EAD & Lara E. **Distribution patterns of microeukaryotes cryptic species communities strongly influenced by environmental filter: the case of Nebela collaris complex**. *Mol Ecol*

In preparation

Singer D, Seppey CVW, Karimi B, Mitchell EAD & Lara E. **Co-occurrence networks of micro-eukaryotic communities according to micro-habitats in the Jura peat bogs**.

Seppey CVW, Singer D, Geisen S, et al., Mitchell EAD & Lara E. **Comparison between micro-eukaryotes communities from soil and two aquatic environments through metabarcoding**.

Seppey CVW, Broennimann O, Buri A, Yashiro E, Pinto E, Singer D, Blandenier Q, Mitchell EAD, Niculita Hirzel H, Lara E & Guisan A. **Protists biodiversity hotspot through predictive modeling of metabarcoding data**.

Lara E, Schiaffino MR, Singer D, Fernández L, Seppey CVW & Izaguirre I. **Environmental DNA survey reveals that freshwater planktonic ciliate communities are richer in Antarctic Peninsula than in neighbouring South America**.

Szelecz I, Lösch S, Seppey CVW, Lara E, Singer D, Sorge F, Tschui J, Perottie MA & Mitchell EAD. **Comparative analysis of bones, mites, soil chemistry, nematodes and soil micro-eukaryotic communities of a suspected homicide to estimate a long post-mortem interval**.

Szelecz I, Feddern N, Seppey CVW, Amendt J & Mitchell EAD. **Is *Saprinus semistriatus* (Coleoptera; Histeridae) a useful taxon for PMI (post-mortem interval) estimation?**.

Fernández LD, Seppeyc CVW, Singer D, Fournier B, Tatti D, Mitchell EAD & Lara E. **Elevational diversity patterns in free-living soil unicellular eukaryotes are driven by evolutionary constraints to water and energy availability**.

Reczuga MK, Seppey CVW, Szelecz I, Fournier B, Singer D, Lara E, Mulot M & Mitchell EAD. **Temporal patterns of soil micro-eukaryotic diversity beneath decomposing pig cadavers as assessed by high throughput sequencing**.

Reczuga MK, Seppey CVW, Mulot M, Jassey V, Buttler A, Słowinska S, Słowinski M, Amandine Pillonel, Lara E, Lamentowicz M, & Mitchell EAD. **Assessing the Responses of Peatland Micro-Eukaryotes to Climate Change Using Next Generation Sequencing**.

Contribution to conferences

Oral

2013 Seppey C, Fournier B, Mulot M, Mitchell EAD, Szelecz I & Lara E. **Euglyphida communities under pig cadavers assessed by high throughput sequencing: A new indicator of the post mortem interval (PMI) ?**. *712th Meeting of the Swiss Society for Microbiology, Murten, Switzerland*

2014 Seppey C, Fournier B, Mulot M, Mitchell EAD, Szelecz I & Lara E. **Euglyphida communities under pig cadavers assessed by high throughput sequencing: A new indicator of the post mortem interval (PMI) ?**. *Invited lecturer at the Institute of Forensic Medicine, Frankfurt am Main, Germany*

2014 Seppey C, Fournier B, Mulot M, Mitchell EAD, Szelecz I & Lara E. **Euglyphida communities under pig cadavers assessed by high throughput sequencing: A new indicator of the post mortem interval (PMI) ?**. *Eukaryotic -Omics: Exploring and testing witha next-generation sequencing, Geneva, Switzerland*

2014 Seppey C, Fournier B, Mulot M, Mitchell EAD, Szelecz I & Lara E. **Euglyphids communities in Cadaver Decomposition Island by a metabarcoding approach**. *Eukaryotic -Omics: Exploring and testing witha next-generation sequencing, Geneva, Switzerland*

2016 Seppey C, Singer D, Belbahri L, Mitchell EAD, Lara E. **Contrasting patterns of soil micro-eukaryotic taxonomic and functional diversity among forest, grassland and croplands in Switzerland**. *35th Meeting of the German Society of Protozoology, Saignelégier, Switzerland*

Poster

2013 Seppey C, Fournier B, Mulot M, Mitchell EAD, Szelecz I & Lara E. **Euglyphida communities under pig cadavers assessed by high throughput sequencing: A new indicator of the post mortem interval (PMI) ?**. *32th Meeting of the German Society of Protozoology, Ittigen, Switzerland*

2014 Seppey C. **Micro-eukaryotic biodiversity and ecology assessed by metabarcoding** . *Annual Ph.D. students meeting, Neuchâtel, Switzerland*

2014 Seppey C, Hasler K, Paul, C, Quartier M, Belbhari L, Szelecz I, Fournier B, Tarnawski SE, Gonsález Garraza G, Quiroga MV, Schiaffino MR, Singer D, Izaguirre I, Mataloni G, Mitchell EAD, Lara E. **Metabarcoding of micro-eukaryotes – multiple applications from biodiversity assessment to applied ecological research**. *The first global soil biodiversity conference, Dijon, France*

2015 Seppey C, Hasler K, Paul, C, Quartier M, Belbhari L, Szelecz I, Fournier B, Tarnawski SE, Gonsález Garraza G, Quiroga MV, Schiaffino MR, Singer D, Izaguirre I, Mataloni G, Mitchell EAD, Lara E. **Metabarcoding of micro-eukaryotes – multiple applications from biodiversity assessment to applied ecological research**. *European congress of Protistology, Sevilla, Spain*

2016 Seppey C, Singer D, Dumack K, Belbahri L, Mitchell EAD, Lara E. **Co-occurrence of phagotrophic and phototrophic microbial eukaryotes suggest alternative pathways for nutrient cycling in temperate European soils**. *Meeting of the Swiss Soil Monitoring Network, Reckenholz, Switzerland*

2017 Seppey C, Broennimann O, Buri A, Yashiro E, Pinto E, singer D, Blandenier Q, van der Merr JR, Mitchell EAD, Niculita Hirzel H, Guisan A, Lara E. **Spatial modeling of soil protist diversity in the Swiss western Alps**. *37th Meeting of the German Society of Protozoology, Meissen, Germany*

Remerciements / Acknowledgements

Comme vous vous en doutez, cette thèse n'aurait pas été possible sans une pléthore de gens qui m'ont aidé et soutenu durant ces 4 années. J'aimerais tout d'abords remercier mes superviseur et co-superviseur Edward et Enrique. Edward pour m'avoir laisser la liberté de développer des compétences nouvelles et d'avoir rassemblé une si belle équipe dans une si belle atmosphère. Enrique pour tout le temps investi et les réponses à mes questions: tu mérites bien ta réputation de bibliothèque sur pattes! The three external thesis experts: David, Micah and Stefan, who take the time to read and comment this work: gentlemen, you have my thanks and greetings. J'aimerais en troisième lieux remercier ma mère Joëlle, mon père Angelin et mon frère Mathieu qui m'ont donné le goût d'apprendre, de faire de mon mieux et de finir ce que je commence, et surtout de m'avoir laisser partir.

Merci aussi à tout mes collègues du Laboratoire de Biodiversité du Sol pour tout le bon temps et l'aide dont vous m'avez fait profiter. À Bertrand, mon prédécesseur et premier de la dynastie des valaisans, pour tout les coups de main en statistique lors de ma maîtrise. To Ildikò, faithful officemate who gave me the chance to work on a unusual project I will remember after my death, who had read and corrected this thesis (twice), and with whom I shared all these late evening discussions and meals: Egészségedre! À David de m'avoir permis d'utiliser une partie non négligeable de ses données, pour toute les collaborations sans qui ma thèse ne serait pas ce qu'elle est et pour toute l'aide dans les derniers moments. À Matthieu, notre Mulot préféré pour les bons rires et les discussions de geeks. Aux anciens masterants devenus grands, Quentin B. et Clément mon successeur, valaisan 3e du nom: maîtres, que la force soit avec vous. À Amandine pour son aide indispensable au laboratoire, Nicolas D. et son panier de légumes et les 2 chefs Lassaâd et Alexandre A. A special thanks also to the colleagues outside Switzerland, Monika and Léonardo which gave my the possibility to integrate their projects and idea. Bref, merci pour tout le bon temps passé avec vous.

Un gros merci aussi à mes amis à l'extérieur du laboratoire. Aux gens du Laboratoire Sol et Végétation / Écologie Fonctionnel. À Dylan mon fidèle chauffeur botaniste. Au chaux-de-fonniers Anaele, Claire, Lidia, Boris; je vous souhaite toute la neige nécessaire! Aux licornes Stella, Vinciane et Noémie: votre bonne humeur est contagieuse! Aux parasitologues Jonas D., Daniel, Alexandre V., Giacomo et les autres; barbus, roux, ou pas. À BikeAttitude pour m'avoir gardé en forme et permit de continuer à pratiquer mon maniement de la pelle. À Christiane et Agnès de la cafétéria pour m'avoir fourni la clef de la satiété. À Quick Pizza pour leur inébranlable rapidité et de m'avoir permis d'allonger tant de soirée de travail. Merci à mes amis du Canada, Marc, Jesse, Florent, Nicolas M. chez qui le retour est toujours aussi joyeux. Et au lausanois Benjamin, Tristan, Vincent, Jonas C., Nicolas R., Selma, Laura, Balazs et tout les autres pour les pétanques, vélos, randonnées, hébergements, bref, qui avez rendu mes 4 premières années en Suisse si agréable. Merci à toute la Suisse, à laquelle je peux dire maintenant appartenir.