Elsevier Editorial System(tm) for Science of the Total Environment

Manuscript Draft

Manuscript Number:

Title: Land use as a driver for protist community structure in soils under agricultural use across Europe

Article Type: Research Paper

Keywords: arable
grassland
protist microbiomes
land use intensity
metagenomics
V4 region of 18S rRNA

Corresponding Author: Dr. Anne Winding, Ph.D.

Corresponding Author's Institution: Aarhus University

First Author: Susana S Santos, PhD

Order of Authors: Susana S Santos, PhD; Anne Schöler, PhD; Tue K Nielsen, PhD; Lars H Hansen, PhD; Michael Schloter, PhD; Anne Winding, Ph.D.

Abstract: Soil biodiversity is threatened by intensification of land use. The consequences of different land use on belowground biodiversity remain insufficiently explored for soil protists. Alongside being abundant and extremely diverse in soil, protists provide a large range of ecosystem services, as key players in the microbial loop, turnover of organic matter and in the stimulation of plant growth-promoting rhizobacteria. However, we lack knowledge of effects of site, land use and management intensity on diversity of soil protist communities. Here we assessed protist communities in four European arable sites with contrasting land use intensities at each site: Lusignan, France; Moskanjci, Slovenia; Castro Verde, Portugal and Scheyern, Germany as well as two grassland sites: Hainich, Germany and Lancaster, UK. Each site have consistent agricultural management history of low and high land use intensities quantified in terms of land use index (LUI). We employed high-throughput sequencing of environmental DNA, targeting the V4 region of the 18S rRNA gene. By assigning the protist composition to trophic groups, we inspected for effects of management, and other biotic and abiotic variables. While overall protist richness was unaffected by LUI within sites, specific trophic groups was affected. Effects on protist biome across land uses and sites were observed. LUI sensitive taxa were taxonomically diverse in each plot, and their trophic groups responded in specific patterns to specific practices. The most abundant trophic group was phagotrophs (73%), followed by photoautotrophs (16%), plant pathogens (4%) and animal parasites (2%). Community compositions and factors affecting the structure of individual trophic groups differed between land uses and management systems. The agricultural management selected for distinct protist populations as well as specific functional traits, and the protist community and diversity were indeed affected by site, land use and management intensity, which indicates the ecological significance of protists in the soil food web.

Suggested Reviewers: Mathilde Borg
University of Greifswald
mathildeborg.dahl@uni-greifswald.de
Soil protist expert

Enrique Lara
Real Jardin Botanico CSIC
enrique.lara@rjb.csic.es
Soil protist ecologist

David Singer
University of Sao Paulo
david.singer.bio@outlook.com
Biologist specialised into protistology

Alexandre Jousset
Utrecht University, NL
A.L.C.Jousset@uu.nl
Microbial biotechnology for sustainable agriculture

Stefan Geisen
Wageningen University, NL
s.geisen@nioo.knaw.nl
Expert in soil protistology and ecosystem services

Opposed Reviewers:

Cover letter for submitted manuscript "Land use as a driver for protist community structure in soils under agricultural use across Europe"

Roskilde, December 3rd 2019.

Dear Editor,

We are submitting the manuscript entitled "Land use as a driver for protist community structure in soils under agricultural use across Europe" to be considered for publication in *Science of the Total Environment Journal*.

The work presented for publication studied the impact of different land uses, soil managements and intensities on protist community diversity and structure in soils across Europe. Protists are important components of trophic chains and nutrient cycles in terrestrial environments, where their relationships and functional roles may vary from parasitic to phagotrophic or saprotrophic, influencing the overall soil community dynamics. However, broad surveys of protists in agricultural environments are scarce.

Protist microbiomes were highly differentiated across environments and network analysis and functional profiling demonstrated that land use management and intensities have functionally consistent effects. The results of this study identified specific protist taxa and protist trophic groups that can be used to track community patterns that are linked to land use intensification. Our snapshot study of protist diversity (living or dead) provides a baseline of protist diversity in different European areas that can help to characterize and understand the agricultural protist microbiome.

The field studies summarized in our manuscript are related to the biosphere, lithosphere and anthroposphere: concerning the effects of anthropogenic activities (agricultural management) on soil structure and in consequence on protist biodiversity and general ecosystem services. Hence, we feel the manuscript covers the interest of the readers of *Science in the Total Environment Journal.*

We confirm that the manuscript describes original work and has not been published elsewhere nor is it currently under consideration for publication elsewhere. All authors approved the manuscript submission and they have no competing interests.

We hope this manuscript meets the journal requisites and is suitable for publication in *Science of the Total Environment Journal* and look forward to the reviewers comments.

Sincerely,

Anne Winding

*Title Page

Land use as a driver for protist community structure in soils under agricultural use across Europe

Susana S. Santos[1,3], Anne Schöler[2,4], Tue Kjærgaard Nielsen[1,5], Lars Hansen[1,5], Michael Schloter[2], Anne Winding[1#]

[1]Department of Environmental Science, Aarhus University, Roskilde, Denmark

[2]Research Unit for Comparative Microbiome Analysis (COMI), Helmholtz Zentrum München, Neuherberg, Germany

[3] Current address: Department of Agroecology, Aarhus University, Flakkebjerg, Denmark

[4] Current address: DKFZ German Cancer Research Center, Berlin, Germany

[5] Current address: Department of Plant and Environmental Science, Copenhagen University, Denmark

#corresponding author:

Anne Winding

Department of Environmental Science

Aarhus University

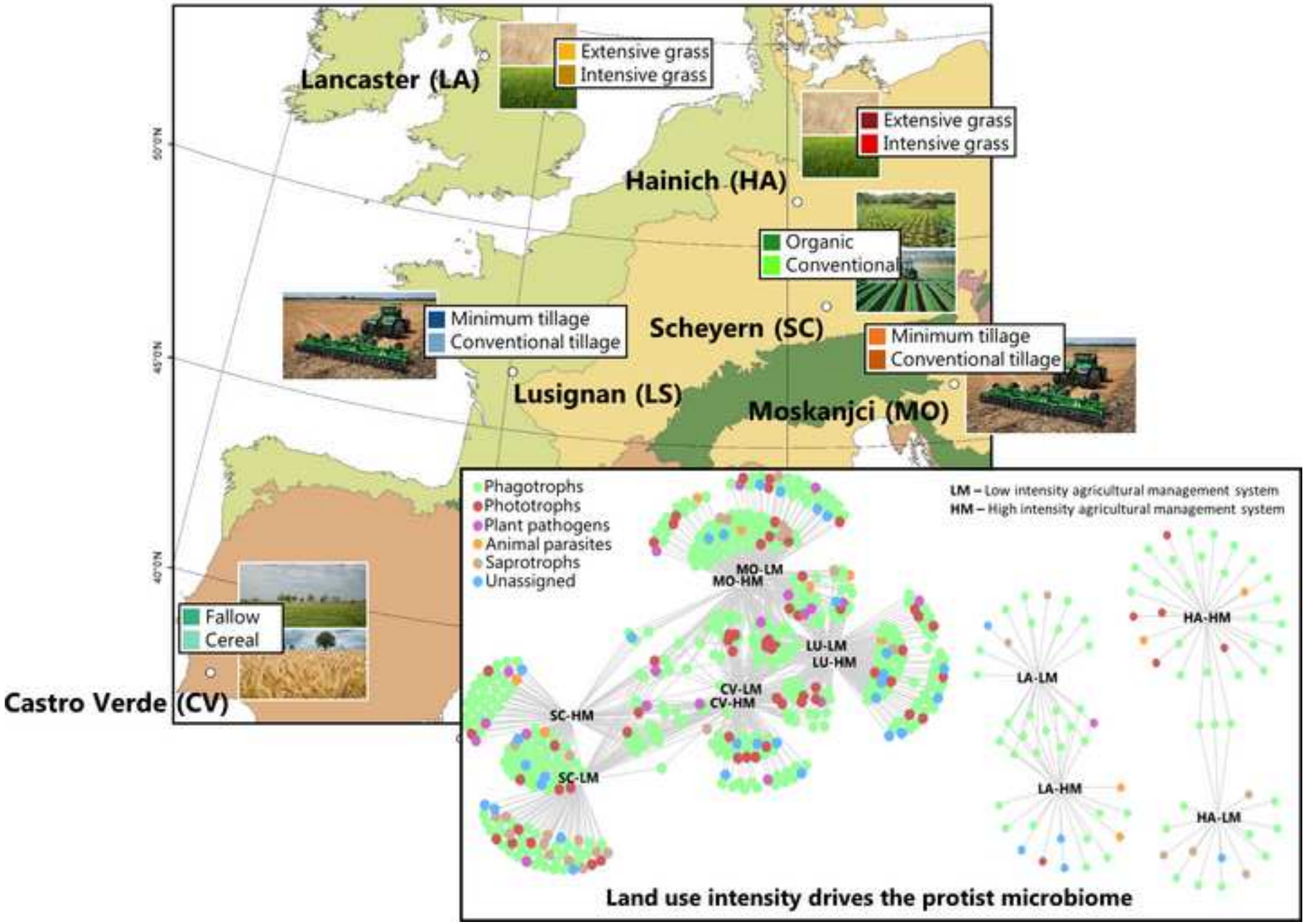
Frederiksborgvej 399,

DK-4000 Roskilde, Denmark

Phone: +45 87158615

e-mail: aw@envs.au.dk

*Graphical Abstract



Highlights

- Soil protist microbiomes across diverging land uses were significantly different.
- Land use affects the microbiome of trophic protist groups.
- Protist taxa sensitive to land use intensity were identified.
- Impact of environmental parameters on different protist groups differed.

Land use as a driver for protist community structure in soils under agricultural use across Europe

Susana S. Santos[1,3], Anne Schöler[2,4], Tue Kjærgaard Nielsen[1,5], Lars Hansen[1,5], Michael Schloter[2], Anne Winding[1#]

[1]Department of Environmental Science, Aarhus University, Roskilde, Denmark

[2]Research Unit for Comparative Microbiome Analysis (COMI), Helmholtz Zentrum München, Neuherberg, Germany

[3] Current address: Department of Agroecology, Aarhus University, Flakkebjerg, Denmark

[4] Current address: DKFZ German Cancer Research Center, Berlin, Germany

[5] Current address: Department of Plant and Environmental Science, Copenhagen University, Denmark

#corresponding author:

Anne Winding

Department of Environmental Science

Aarhus University

Frederiksborgvej 399,

DK-4000 Roskilde, Denmark

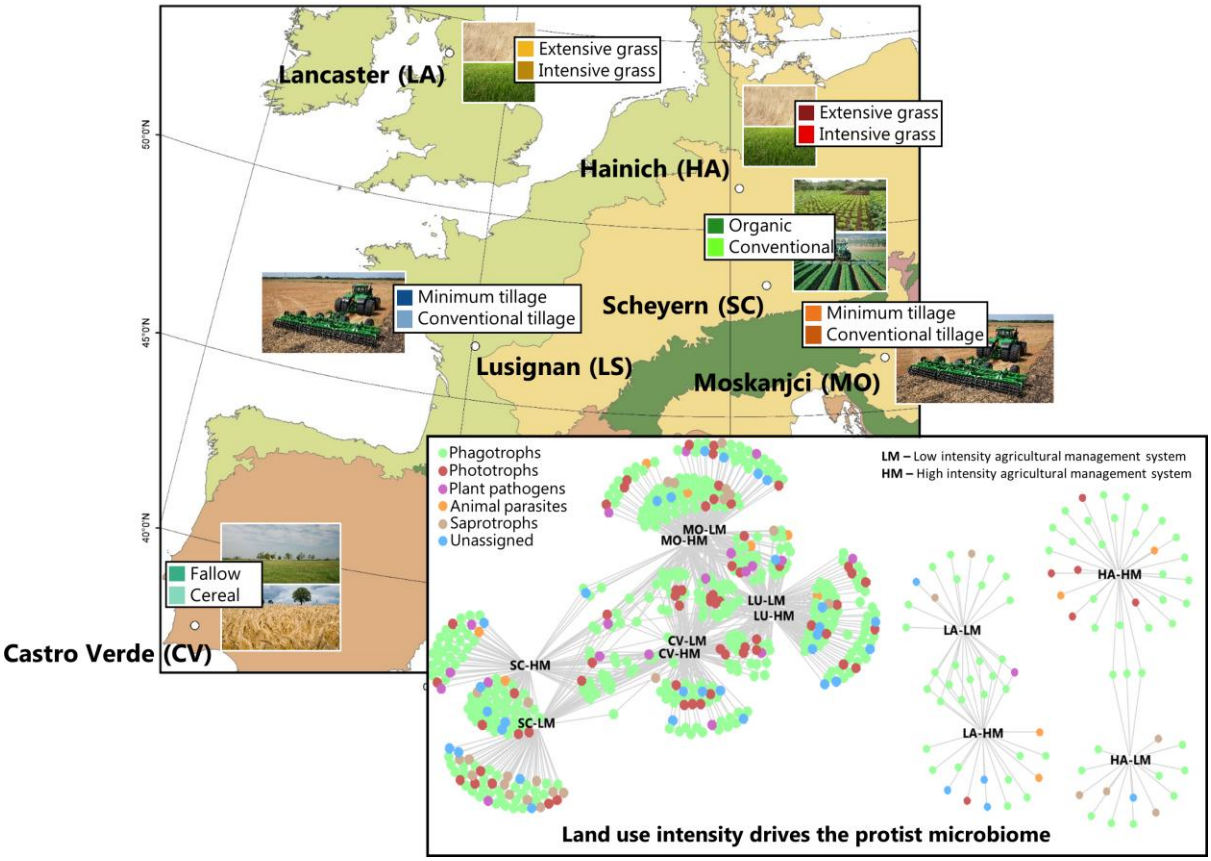
Phone: +45 87158615

e-mail: aw@envs.au.dk

Highlights

- Soil protist microbiomes across diverging land uses were significantly different.
- Land use affects the microbiome of trophic protist groups.
- Protist taxa sensitive to land use intensity were identified.
- Impact of environmental parameters on different protist groups differed.

Graphical abstract



Running title: Land use effects on protist microbiome

Abstract

Soil biodiversity is threatened by intensification of land use. The consequences of different land use on belowground biodiversity remain insufficiently explored for soil protists. Alongside being abundant and extremely diverse in soil, protists provide a large range of ecosystem services, as key players in the microbial loop, turnover of organic matter and in the stimulation of plant growth-promoting rhizobacteria. However, we lack knowledge of effects of site, land use and management intensity on diversity of soil protist communities. Here we assessed protist communities in four European arable sites with contrasting land use intensities at each site: Lusignan, France; Moskanjci, Slovenia; Castro Verde, Portugal and Scheyern, Germany as well as two grassland sites: Hainich, Germany and Lancaster, UK. Each site have consistent agricultural management history of low and high land use intensities quantified in terms of land use index (LUI). We employed high-throughput sequencing of environmental DNA, targeting the V4 region of the 18S rRNA gene. By assigning the protist composition to trophic groups, we inspected for effects of management, and other biotic and abiotic variables. While overall protist richness was unaffected by LUI within sites, specific trophic groups was affected. Effects on protist biome across land uses and sites were observed. LUI sensitive taxa were taxonomically diverse in each plot, and their trophic groups responded in specific patterns to specific practices. The most abundant trophic group was phagotrophs (73%), followed by photoautotrophs (16%), plant pathogens (4%) and animal parasites (2%). Community compositions and factors affecting the structure of individual trophic groups differed between land uses and management systems. The agricultural management selected for distinct protist populations as well as specific functional traits, and the protist community and diversity were indeed affected by site, land use and management intensity, which indicates the ecological significance of protists in the soil food web.

Keywords: Arable, grassland, protist microbiomes, land use intensity, metagenomics, V4 region of 18S rRNA

1. Introduction

Understanding the response of the soil microbiome to changes in land use and environmental drivers has become an important research focus, as microbes are important drivers of soil quality (George et al., 2019). While the dramatic implications of land use and management strategies at the landscape level are apparent and well recognized for the biodiversity of plants and animals, these patterns are less obvious for soil microbial communities. Most datasets available furthermore focus on bacteria and fungi, (Szoboszlay et al., 2017a; Malik et al., 2018; Peters et al., 2019). However, the effects of land use on microbial community composition and diversity may depend on the taxa being studied.

Soil microeukaryotes, including protists, are an extremely diverse group of unicellular organisms that underpin globally important ecosystem services (Geisen et al., 2018). The functional role of protists in soils is diversely linked to nutrient cycling and includes phagotrophy, phototrophy, symbiosis, saprotrophy, or a mix of these strategies (Geisen, 2016b; Seppey et al., 2017; Geisen et al., 2018). Soil protist communities can be shaped by a number of factors such as the presence and composition of prey, a range of biotic and abiotic factors (such as plant species) and soil properties, such as pH and humidity (Gao et al., 2019). The ecological key role of protists contrasts with the lack of understanding of their diversity and structure in different land use management strategies.

When analysing protist communities by qualitative and quantitative methods in German grasslands, Glaser et al. (2015) and Fiore-Donno et al. (2016) observed that land use was of minor importance, while site appeared to be the main factor shaping protist communities. However, Lentendu et al. (2014) showed that changes in soil pH, moisture and organic nutrient availability caused by different land management strategies were responsible for shifts in the composition of different eukaryotic communities, including protists. Additionally, a recent study demonstrated that protists are the most susceptible component of the soil microbiome to the application of nitrogen fertilizers (Zhao et al., 2019). Nevertheless, it still remains to clarify how the impacts of land use at the field scale compare to continental scale drivers of microbial diversity such as climate and soil type, affect the protist microbiomes. Therefore, identifying the factors driving the abundance and the protist microbiome structure will provide fundamental knowledge on the maintenance

of ecosystems services in agricultural systems and the prediction of their responses to environmental disturbance.

Therefore, the main goal of this study was to determine the differences in the protist diversity and community structure resulting from diverging land use and management strategies at different sites. The studied sites are a series of long term observatories (LTOs) representing arable and grassland habitats across Europe. The LTOs have a consistent agricultural management history of several years and are composed of three independent, replicated plots of two contrasting treatments, with low and high land use intensity (LUI) that was previously calculated on the basis of the LUI index by Griffiths et al. (2016). The LUI index corresponds to a specific regime of mowing, grazing and fertilization, which have been quantified and combined to create a uniform land use intensity value (Griffiths et al., 2016). This classification allows simultaneous comparisons between all experimental plots.

We used an amplicon sequencing approach to assess the community structure of protists based on soil DNA, partial amplification of the 18S rRNA gene and MiSeq sequencing of the amplicons. The obtained reads were assigned to five broad trophic groups of protists (phagotrophs, phototrophs, animal parasites, plant pathogens, and saprotrophs) (Seppey et al. 2017, Schulz et al. 2019) to compare community composition between land uses and intensities, across and within sites. Finally, we analysed the importance of potential abiotic environmental factors as drivers for the distribution of the trophic groups. We hypothesized that (i) in response to LUI, distinct protist communities will develop in arable and grassland soils; and (ii) LUI shapes the protist microbiomes, with lower protist richness under high LUI.

2. Materials and Methods

2.1 Site, soil sampling and soil properties

Soil samples from six countries belonging to four different climatic zones were collected from six long-term observatories (LTO) in collaboration with the EU FP7 EcoFINDERS (#264465) project during fall 2013 as described by Griffiths *et al.* (2016) (Figure 1A). The LTOs were four arable sites: Lusignan, France; Moskanjci, Slovenia; Castro Verde, Portugal and Scheyern, Germany; in addition two grassland sites: Hainich, Germany and Lancaster, UK were sampled. The sites covered a broad range of soil physical-chemical characteristics

to ensure a contrast in the soil microbiological properties. At each site, we sampled top soil (0-15 cm) in three replicate plots of both low and high LUI management. Samples were kept in an insulated box containing frozen 'cool blocks' until they could be stored at 4°C in a laboratory. Additional information related to the LTO's characteristics can be found in Griffiths et al. (2016) and Table 1. Soil chemical properties varied with land use type (Table 1). Grassland soils had a significantly higher moisture content, total C and N, organic matter and clay content, but lower silt content than arable soils. Within the arable soils, Castro Verde soils were more acidic and had lower moisture content, organic matter and total C and N content. In terms of soil texture, all arable soils exhibited internal significant differences. Within the grassland soils, Hainich showed the highest pH values, lowest moisture and organic matter content and highest clay percentage. Significant differences in soil parameters between low and high LUI within each site were not observed (ANOVA, $p > 0.5$). Biologically relevant climatic variables, such as mean average precipitation (MAP) and mean average maximum temperature (MAmT) for October 2013 were retrieved from the WorldClim project (Hijmans et al., 2005) for each site.

2.2 DNA extraction and quantitative PCR (qPCR) analysis

Total genomic DNA was extracted from 0.25 g (wet weight) of soil samples using the Power Lyzer™ PowerSoil®DNA Isolation Kit (MoBio laboratories Inc, USA), following the manufacturer's instructions. DNA from three soil replicates per plot were each extracted three times, and the three extraction replicates were pooled for further analysis. The quantity and quality of the extracted DNA were checked using a NanoDrop ND-2000c UV-Vis Spectrophotometer (NanoDrop Technologies, Thermo Scientific, Nærum, Denmark). The abundance of 18S rRNA genes was determined by 18S rRNA gene qPCR assay using the primers Euk345f/Euk499r, and a separate specific assay to amplify a 130 bp fragment on free-living amoeba 18S rRNA gene using the primers Amo1400F/Amo1540R, as described in Santos *et al.* (2015). qPCR analysis were performed in a 7300 Real-Time PCR System (Applied Biosystems) with a final volume of 20 µl, using the Power SYBR® Green PCR Master Mix (Applied Biosystems, UK) that contained 10 µl of SYBR® Green PCR Master Mix, 0.4 µM of each primer, 0.6 mg ml^{-1} of BSA and 2 µl of 1:50 dilutions of soil DNA as template. The final mix was added to 96-well PCR plate (4titude® Ltd, Surrey, UK) and fluorescence measurements were done at the end of cycle at the elongation step. The dissociation curve

from 60°C to 95°C was measured after the last qPCR cycle. Data were analysed using SDS v1.4.1 (Applied Biosystems). The amplification efficiency was determined using a standard curve for ten-fold serial dilutions of defined DNA amounts of *Acanthamoeba castellannii*. The initial DNA amount in each reaction was approximately 40 ng μl^{-1}. For each cycle, the C_t (cycle threshold) was calculated, the standard curves (C_t values *vs* DNA concentration) were constructed and the corresponding slope was used to calculate PCR amplification efficiency (E) according to the equation of $E = 10^{(-1/slope)} -1$. Standard curves were linear from $2.95x10^6$ to 20 copies, with R^2 values higher than 0.98 and PCR efficiencies from 86 to 105%.

2.3 Amplicon library preparation

The hypervariable V4 region of the 18S rRNA gene was amplified using the general eukaryotic primers TAReukV4F/TAReukREV (5'-CCAGCASCYGCGGTAATTCC-3'/5'-ACTTTCGTTCTTGATYRA-3') following the procedure described by Stoeck *et al.* (2014). Forward and reverse primers were tagged at the 5' end with Illumina overhang adapters, following the manufacturer's instructions. PCR reactions were run in duplicate using a peqSTAR 96X Universal Thermal Cycler (PEQLAB, VWR International GmbH, Germany) with the proof-reading DNA polymerase, NEB Next High Fidelity 2X (BioLabs, Frankfurt, Germany), 0.5 µM of each primer and 10 ng μl^{-1} of template DNA in a final reaction volume of 25 µl. The PCR runs was initiated with a first denaturation phase of 98°C for 30 sec, followed by eight cycles with a denaturation temperature of 98°C for 10 seconds, an annealing temperature of 55°C for 30 seconds and an elongation temperature of 72°C for 30 seconds, and terminated by a final incubation at 72°C for 5 minutes. Amplicons were gel-size selected for the appropriate length of 450-500 bp and purified (QIAquick Gel Extraction Kit, Qiagen, Germany). PCR products were multiplexed using Nextera® XT Index adapters (Illumina) in a subsequent PCR and purified using a magnetic bead capture kit (Ampure-Agencourt, Beckman Coulter, Krefeld, Germany). Library quality was assessed using the Bioanalyzer 2100 with the Agilent DNA 1500 DNA kit (Agilent, Santa Clara, CA). Subsequently, all samples were quantified using a Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen, Schwerte, Germany) and pooled equimolarly. Sequencing was performed with an Illumina MiSeq® Reagent Kit v3 (600 cycles) for paired end sequencing on a MiSeq® sequencing machine.

188 **2.4 Sequencing processing, taxonomic assignment and phylogenetic analysis**

189 Recovered sequences were processed by using the UPARSE pipeline (Edgar 2013; 2016)
190 with USEARCH v10, as suggested by (Xiong et al., 2019). The clustering of operational
191 taxonomic units (OTUs) was based on the following workflow: (i) quality filtering of the
192 sequences using a "maxee" (i.e., maximum per sequence expected error frequency) value of
193 1; (ii) dereplication of identical sequences performed via the fastx_uniques command; (iii)
194 OTU clustering with *-cluster_otus* that combines a chimera removal step; (iv) and generating
195 OTU abundance table by mapping the total reads to representative sequence with *-otutab*
196 command. Taxonomic annotation was assigned to representative sequence of each OTU by
197 UCLUST (Edgar, 2010) in QIIME *(-assign_taxonomy.py*) against PR2 database (version 4.11.1)
198 (Guillou et al., 2013). All sequences unassigned were removed, retaining a total of 3,978,906
199 sequences for the 36 samples (average 110,525 reads per sample). Furthermore, we
200 discarded plant (Streptophyta), animal (Metazoa), and fungi sequences to generate the
201 retained and conservative protist OTU table. To avoid sampling size effects on the alpha
202 diversity analysis, the number of reads per sample was normalized by randomly
203 subsampling to the lowest number of protist reads across all samples (Figure S1,
204 supplementary information).

205 **2.5 Statistical analysis**

206 All statistical analyses were conducted in R v3.6.1 (R Core Team, 2013). Briefly, rarefaction
207 analyses were performed on the filtered OTU tables with QIIME1. Rarefaction curves were
208 built based on the observed OTUs with *rarecurve* function of the *vegan* package (v.2.5-6;
209 Oksanen et al., 2019). Rarefaction analysis showed that all samples were approaching
210 saturation with the achieved number of reads. It is, therefore, reasonable to infer that the
211 obtained sequences covered the majority of the entire protist microbiome (Figure S1). Alpha
212 diversity (Richness, and Shannon index) was calculated per sample based on randomly
213 subsampled read count data to compensate for variation in read numbers across samples,
214 using the *vegan* package (Oksanen et al., 2018). The *aov* function was used to calculate
215 pairwise ANOVA *p*-values for alpha diversity measures between samples. The effects of land
216 use and LUI were tested on alpha diversity within and among each site, respectively.

217 The relationship between protist OTUs that derived from different land uses (arable and
218 grassland sites) was first visualized using Venn diagrams (*limma* package, Smyth 2005). The

soil protist microbiome composition was ordinated by NMDS analysis using Bray-Curtis dissimilarity distance at an OTU level with the *phyloseq* package (McMurdie and Holmes, 2013). Differences among microbiome composition of the different sites and plots were calculated using permutational analysis of variance (PERMANOVA) using the functions *adonis* in the *vegan* package with 10^4 permutations (Oksanen et al., 2019). One-way analysis of variance (ANOVA corrected by Benjamini–Hochberg FDR) was used to compare mean differences between the sites and treatments using STAMP (Parks et al., 2014). For each trophic group, the top 10 genera were further represented along with the combined local contribution to beta diversity (LCBD, with default Hellinger dissimilarity distance) using the *MicrobiomeSeq* package (http://www.github.com/umerijaz/microbiomeSeq).

To explore the most discriminant OTUs between LUI systems (LUI sensitive OTUs, hereafter LUIs-OTUs), only OTUs with more than 2 sequences (avoiding single-count OTUs) and present in at least three samples (the number of replicates per treatment) were considered. For that, negative binomial linear discriminant analysis and generalized linear models were performed using the likelihood ratio test (FDR-corrected $p < 0.05$) with the R package *edgeR* (Robinson et al., 2009). However, our analysis included only OTUs that played a key role and had a high discrimination power in explaining the protist community composition at each soil plot. These were assessed previously, by indicator species analysis with the R package *indicspecies* (De Cáceres et al., 2010). The IndVal indices ("multipatt" function; *indicspecies* package) combines the species relative abundance with its relative frequency of occurrence to statistically determine species associated to one or several particular plots. The significance of the indicator value of each species was assessed by a randomization procedure with 9999 permutations (De Cáceres et al., 2010). The significant OTU associations to one or more of the different LUI systems were visualized using bipartite networks. The networks were constructed using the Fruchterman-Reingold layout with 10^4 permutations as implemented in the R package *igraph* (Csárdi and Nepusz, 2006) and described in (Hartman et al., 2018). Moreover, mean relative abundances of LUIs-OTUs were plotted using the function *heatmap.2* from *gplots* package (Warnes et al., 2019).

To estimate the source of variation for protist communities and to compare the effects of the biotic and abiotic factors we used variation partitioning ("varpart" function in *vegan* package). Thus, models were constructed containing four groups of predictors: LUI, soil

structure (clay, sand and silt content), climatic variables (MAP and MAmT) and soil properties (pH, moisture, organic matter, carbon and nitrogen). To examine the relationships among the different trophic groups (hellinger transformation) and environmental variables, redundancy analysis (RDA) was performed using the *vegan* package, and significance was tested using the "envfit" function using 999 permutations.

3. Results

3.1 Structure of the soil protist microbiome

The 18S rRNA gene abundance was in the range of $1.35x10^3$ to $2.58x10^4$ copies per g^{-1} of soil with Lusignan and Castro Verde values being significantly lower than Hainich and Sheyern (Figure 1B). The amoeba-specific 18S rRNA gene was up to 10-fold lower in abundance than total 18S rRNA gene with soil from Moskanjci having significantly higher abundances ($p < 0.05$, Tukey's test).

Sequencing and quality filtering resulted in a total of 5,140,865 high-quality 18S rRNA gene sequences for all samples. After removal of singletons, fungi, metazoa, streptophyta and the unclassified sequences, the protist dataset comprised 7,111 OTUs and 2,932,711 high quality sequences (57% of the retained eukaryotic sequences and 36% of the OTUs) with an average read length of 425 bp and an average of 82,105 ± 22,492 (Mean ± SD) reads per sample. Arable and grassland sites present different microbial habitats with specific sets of protist OTUs (Figure S2). The grassland sites shared 34% of OTUs with the arable sites (Figure S2A) with only 14%, and 26% of the OTUs being shared among the arable and grassland sites, respectively (Figure S2B-C).

Protist richness was significantly higher in arable sites compared to grassland sites (ANOVA, $p < 0.05$). The only exception was the very dry agricultural site in Castro Verde, where also low diversity pattern of protists were observed (Figure 2A). Shannon index indicated significantly higher protist diversity for Moskanjci and Lancaster sites (Figure 2A). Additionally, significant differences were observed for protist diversity between low and high LUI at Moskanjci site (Figure 2A). Non-metric multidimensional scaling (NMDS) using Bray–Curtis distances showed consistent differences in β-diversity between sites and across land uses. Plots show tight clustering of the arable sites, whereas the grassland sites form a more dispersed assemblage (Figure 2B). Within site, a separation between low and high LUI

can be observed for the Castro Verde and Hainich plots. PERMANOVA analysis of the total dataset corroborated the trend observed in the ordination plot, with site explaining 50% of the protist variation observed ($p < 0.001$), and a smaller but significant impact of LUI explaining 12% of the variation (Figure 2B).

Rhizaria was by far the protist group with the highest relative abundance (38% of total reads) at all sites, followed by Amoebozoa (31%), Alveolata (12%), Archaeplastida (10%), and Stramenopiles (4%) (Figure 2C). In addition, five rare groups remained: Opisthokonta (0.4% of the assigned protist sequences), Hacrobia (0.3%), Excavata (0.1%), Apusozoa (0.04%) and Protalveolata (0.02%) (data not shown). However differences in relative abundance pattern for the different groups between sites and in response to LUI were observed. Rhizaria OTUs were more abundant in Castro Verde and Lancaster compared to the other sites and were entirely comprised of Cercozoa (ANOVA, $p < 0.05$; Benjamini–Hochberg FDR). Among Amoebozoa, proportions of Conosa were relative constant across sites, while Lobosa was significantly less abundant in Castro Verde. Ciliates were the most common Alveolates and were most abundant in Hainich site, whereas the proportion of Apicomplexa was higher at Castro Verde. While the most abundant Archaeplastida taxa, Chloroplastida (green algae), was in relative lower abundance in Lancaster, Rhodophyta (red algae) was in higher relative abundance in Castro Verde. The proportions of Stramenopiles, Ochrophyta were reduced in Hainich site, while those of Oomycetes and Bicosoecida were relatively constant across sites (ANOVA, $p < 0.05$; Benjamini–Hochberg FDR, data not shown).

3.2 Richness and community composition of trophic groups

Based on the different trophic groups, overall the relative abundance of protists decrease in the order of phagotrophs (73% of the assigned protist sequences), phototrophs (16%), plant pathogens (4%), animal parasites (2%), saprotrophs (1%) and unassigned (4%). The relative OTU richness of phagotrophs was higher at Moskanjci and Scheyern, followed by Lusignan, whereas Castro Verde and Lancaster sites did not show significant differences in OTU richness. Hainich samples recovered the lowest richness for phagotrophs (Figure 3).

Unclassified Variosea taxa (Amoebozoa, Conosa), were dominant in arable samples, with exception of Castro Verde that revealed a phagotrophic composition similar to the Lancaster site, with dominance of Rhogostoma-lineage (Rhizaria, cercozoa) (Figure S3A). Hainich also

311 showed a distinct community pattern with Oxytrichidae (Alveolata, Ciliophora) and AND16-
312 lineage (Amoebozoa, Conosa) being the dominant taxa (Figure S3A).
313 The grassland phototrophic community composition was significantly different from the
314 arable sites and the Castro Verde site showed significant differences from all the other plots
315 investigated (Figure 3). While *Cryptodesmus* (Archaeplastida, Chlorophyta) and *Muriella*
316 (Archaeplastida, Chlorophyta) were more relative abundant in the three arable sites,
317 unclassified Cyanidiales (Archaeplastida, Rhodophyta) dominated at Castro Verde and at
318 both grassland sites (Figure S3B). Additionally, Castro Verde site showed a higher relative
319 abundance of unclassified Chlamydomonadales (Archaeplastida, Chlorophyta) when
320 compared to the other sites (Figure S3B). Hainich showed a higher relative abundance of
321 unclassified Watanabea-Clade (Archaeplastida, Chlorophyta) taxa, while *Chlorochytrium*
322 (Archaeplastida, Chlorophyta) was more abundant in Lancaster.
323 Plant pathogen richness was significantly increased at Lancaster and reduced at Lusignan
324 (Figure 3). Moreover, at Moskanjci and Castro Verde, differences between high and low LUI
325 richness were observed, with low LUI showing higher richness in Moskanjci, but the
326 opposite occurring in Castro Verde (ANOVA, $p < 0.05$). The most abundant plant pathogen in
327 all sites was the cercozoan *Polymyxa* (Rhizaria, Plasmodiophorida) with exception of
328 Lusignan where oomycota taxa were the most abundant (Figure S3C). Other unclassified
329 Plasmodiophorida were higher in relative abundance in Castro Verde and Lancaster (Figure
330 S3C).
331 The relative OTU richness of animal parasites was high in grassland sites, followed by
332 Moskanjci and other arable sites, with gregarines (Alveolata, Apicomplexa) being the most
333 abundant taxa at all sites (Figure 3, Figure S3D). While Lancaster showed a higher relative
334 abundance of unclassified Apicomplexa, and Anurofeca (Opisthokonta, Mesomycetozoa),
335 the rest of the sites were rich in Colpodellidae (Alveolata, Apicomplexa) and other
336 unclassified Pseudoperkinsidae (Opisthokonta, Mesomycetozoa) (Figure S3D).
337 The lowest richness for saprotrophs protists was found at Lusignan and Scheyern (Figure
338 3). Moreover, both grassland sites showed a higher richness at low LUI treatment plots,
339 when compared to the high LUI treatments (ANOVA, $p < 0.05$) (Figure 3). The community of
340 saprotrophs was mostly constituted of Thraustochytriaceae (Stramenopiles, Labyrinthulea)
341 (Figure S3E).

In each land use system the unassigned protists were dominated by different Cercozoa, Stramenopiles and Amoebozoa that could not be further specified and showed a significant higher richness in the grassland sites and Castro Verde (Figure 3, Figure S3F).

3.3 Effect of land use intensity and other biotic and abiotic factors

Discriminant analysis was used to identify those protist OTUs that differed significantly among the different LUIs by implementing indicator species analysis and likelihood ratio tests in *edgeR*. We defined OTUs that were supported by both methods as LUIs-OTUs (Figure 4A) and found a total of 512 and 86 LUIs-OTUs in arable and grassland sites, respectively (Figure 4A and 4B). While the arable sites shared LUIs-OTUs among them, the grassland sites did not have any LUIs-OTU shared between them. There was little overlap (12 OTUs) between arable and grassland LUIs-OTUs (Figure S4). Despite the majority of the LUIs-OTUs being associated to a combination of ecosystems and treatments, OTUs with a high probability of being found only at low or high intensity LUI, within each site were observed (Figure 4). These LUIs-OTUs, drive the differences in community membership and abundance within each site (Figure 4A,B). The LUIs-OTUs covered a broad range of phyla and none of the higher taxonomic groups (e.g., supergroup, division, order) exhibited a consistent response to management (Figure S5). Nevertheless, within the arable sites, Castro Verde and Lusignan the highest proportion of LUIs-OTUs were shared (Figure 4A). Even though animal parasites and saprotrophs were in relative higher abundance in Castro Verde (Figure 3), they were not selected as indicator species, as they are not significantly different between treatments (Figure 4A,C). For Lusignan a higher relative abundance of unassigned taxa in the lower LUI plots was observed. Scheyern showed the highest difference in LUIs-OTUs composition between treatments, with lower LUI having a higher number and relative abundance of saprotrophs (Figure 4A,C). LUIs-OTUs classified as phagotrophs were most abundant in the high LUI treatments at Moskanjci, whereas for low LUI a higher abundance of plant pathogenic taxa was observed (Figure 4A,C). For both grassland sites, LUIs-OTUs identified as saprotrophs were present only at low treatment plots, while animal parasites and phototrophs were only observed at high LUI treatment plots (Figure 4B,D).

LUI explained about 19% of the variance in arable sites and 22% in the grassland sites (separate PERMANOVA; $p < 0.05$) and was the best predictor of β-diversity for both land

uses (Table 2). pH was the second most important variable, followed by soil moisture in arable sites and carbon content in grassland sites. Moreover, organic matter and mean annual precipitation also appear to affect the protist microbiome at arable sites. Redundancy analysis (RDA) indicated that LUI, soil chemical properties (soil pH, moisture and organic matter) and climatic variables were the main drivers for the trophic groups, explaining 50% of the total variation (Figure 5A). Lusignan, Moskanjci and Scheyern were separated from Castro Verde and Lancaster by the first component (RDA1: explaining 65.2% of the total variation). Hainich samples, on the other hand, did not show any particular clustering based on the protist trophic groups. Accordingly, phototrophs were mostly present in the arable sites and were related with soil moisture, organic matter and precipitation. Moreover, pH seems to affect mostly the protist trophic groups at Scheyern and Lusignan sites, while LUI influences the grassland sites and Castro Verde site and is negative correlated with pH. Similarly, variance partitioning analysis showed that the largest part of the community variation was explained by plot-related variables such as LUI, soil physicochemical properties and soil structure (13%), with climate variables explained a small but significant fraction of the variation observed (Figure 5B).

4. Discussion

Land use intensity (LUI) and agriculture management system are important factors for the composition and structure of eukaryotic microbial communities across Europe. However, identifying consistent effects of land management on the soil microbiomes is difficult as LUI and their impacts on edaphic properties like soil pH, bulk density and moisture content are often site-specific and context-dependent (Malik et al., 2018). In order to understand long-term agricultural management effects on soil protists under field conditions, we investigated soil samples from different LTOs with different LUI. We hypothesized that LUI can shape protist communities but contrasting this hypothesis, the overall OTU richness of protists was not significantly different between the different LUI within each site. Nevertheless, as seen before for bacteria and fungi, protist richness was influenced by land use (arable vs. grassland), increasing along a productivity gradient across heterogeneous ecosystems at a European scale (Tardy et al., 2015; Thomson et al., 2015; Szoboszlay et al., 2017; George et al., 2019). In contrast, Seppey et al. (2017) and Schulz et al. (2019) observed little to no differences on protist richness between different land use types, including different arable systems. The conditions promoted by the different land uses, such as lower fertility, soil moisture and organic matter, in the grasslands compared to the arable sites, determined the abundance and composition of soil protist microbiomes. Besides, consistent with studies of bacteria and fungi, the results showed that protist diversity and composition differed more among the contrasting locations across Europe than between the treatments within each location (Thomson et al., 2015; Szoboszlay et al., 2017; Schöps et al., 2018; George et al., 2019). In agreement with Grossmann et al. (2016) and Schulz et al. (2019), the community composition of protists differed significantly between land use types, with grassland communities being more distinct from the ones in the arable systems, supporting the notion that community patterns of protist are highly consistent within a given habitat type.

High-throughput amplicon sequencing of 18S rRNA genes revealed the dominance of Rhizaria, Amoebozoa and Alveolata taxa in grassland and arable environments (Geisen et al., 2015; Dupont et al., 2016; Mahé et al., 2017; de Araujo et al., 2018). The phylum Cercozoa is of high ecological importance in soils, and our results confirmed the dominance of small bacterivore flagellates in the group of Glissomonadida and in Cercomonadida in soil

(Degrune et al., 2019). Amoebozoa comprise a wide variety of amoebae and flagellates and were mostly represented by variosean and tubulean species, many of which are key inhabitants of soil (Voss et al., 2019). Contrary to our study, most environmental studies using general eukaryotic primers reported a dominance of Alveolata (ciliates and apicomplexan) in soils. This is usually assigned to an amplification artefact or to biases towards the most represented taxa in the databases (Bates et al., 2013; Mahé et al., 2017; Shen et al., 2014; Fiore-Donno et al., 2016).

The protist microbiomes differed strongly among sites with LUI significantly affecting the protist community as indicated by NMDS, PERMANOVA and differential abundance analysis (LUIs-OTUs). Due to the fact that protists comprise phylogenetically diverse groups with very different ecological traits, we considered the individual trophic groups community structure, to uncover the changes associated with land use and LUI. Despite the possibility that some groups were misrepresented due to erroneous trophic classification or PCR amplification biases, and the existing risk of inferring function based on phylogenetic position, soil protists are commonly suggested to be bacterivorous, serving together with bacterivorous nematodes as the main controllers of bacteria in soil food webs (Dupont et al., 2016a; Geisen, 2016a; Schulz et al., 2019). Indeed, phagotrophic protists were the most abundant and predominantly, they function as grazers of bacteria, fungi and microfauna including other protists, driving the elemental cycle (Trap et al., 2015; Geisen, 2016b; Geisen et al., 2016). Relative abundance of phagotrophic protists was lower at sites with higher pH (Moskanjci and Hainich), contrasting other studies focussing more on land use and bacterial communities (Schulz et al., 2019). 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) (Seppey et al., 2017). Phototrophs showed higher richness at arable sites as seen previously, where phototroph protists increased in richness and abundance in plantation systems benefitting from the more open canopy (Schulz et al., 2019). Also the distribution of animal parasites and saprotrophs follows the different land use characteristics. The deceasing abundance of arthropod parasites (Mesomycetozoa and Gregarines) from grassland to arable fields is in line with the corresponding decline in habitat complexity and diversity of ecological niches for their hosts (Bates et al., 2013; Mahé et al., 2017; Seppey et al., 2017; Geisen et al., 2018;

Schulz et al., 2019). Moreover, the pesticide use is highest in arable sites reducing the insect diversity and biomass (Seppey et al., 2017). Also, the lower richness of saprotrophs protist in more intensively managed grassland systems (fertilised and grazed; high LUI treatment), when compared with a traditionally managed system (unfertilised, extensively grazed and cut; low LUI treatment) can be explained by the soil management regime that allowed a higher abundance of dead tissue available to specialized saprotrophs. Plant pathogen richness did not vary according to the land use but it was less rich at Lusignan probably due to an incorporation of a grass ley in the arable rotation, making the two treatments at this site very similar (Griffiths et al., 2016). Nevertheless, plant pathogen richness was significantly affected by LUI, being reduced at fallow systems (low LUI treatment) and at conventional tillage plots (high treatment), when compared to the cereal crops (high LUI treatment) and minimum tillage (low LUI treatment) at Castro Verde and Moskanjci sites, respectively. This can be explained by the site's historical soil management regime and the limited range of plants grown on the fallow site at Castro Verde. Indeed, tillage was previously observed to influence protist communities when compared with untilled soil (Adl et al., 2006). Moreover, for the plant pathogens our study recovered a variety of disease agents, such the Plasmodiophorida taxa, responsible for clubroot disease in Brassicaceae; *Spongospora* spp., responsible for powdery scab in potatoes; *Polymyxa* taxa that can function as carriers of viruses in beets and cereals; and other Oomycota crop disease taxa (*Aphanomyces* and *Pythium*) (Geisen et al., 2018).

Protists essentially depend on the water layer connecting soil pores to move, feed and multiply, and hence the habitat size will increase or shrink with changes in moisture (Geisen et al., 2014). In line with this, data showed that a combination of multiple abiotic factors drive protist communities, with pH, moisture, organic matter, carbon content and annual precipitation being the principle factors (Bates et al., 2013; Dupont et al., 2016; de Araujo et al., 2018; George et al., 2019). Notably, the most influential factors driving differences in protist communities in arable and grassland soils were not the same. However, disentangling such related variables is difficult since soil carbon is often co-correlated with other soil properties including bulk density, organic matter, elevation, and mean annual precipitation (Blanco-Moure et al., 2016).

5. Conclusions

Our barcoding analyses provide detailed insights into the diversity and community structure of a broad range of protist groups, and expands our understanding of how land use intensity and edaphic properties can shape soil protist communities, across different agricultural systems in Europe. The data suggests that Rhizaria was the dominant protistan lineage in the agricultural systems. At a functional level, the protistan assemblage in agricultural systems was dominated by phagotrophs and phototrophs with animal parasites and plant pathogens constituting only 6% of the overall sequences. We observed that each management system has context-dependent effects on the diversity and structure of protist microbiomes, and that diversity studies based solely on taxonomy are not sensitive enough to depict effects of LUI in such homogeneous ecosystems. The arable plots promoted higher protist richness and diversity, and their richness can be related to the presence of certain environmental variables specific to the plots. The LUI sensitive taxa appeared dependent on local soil and climatic effects, and on the type of land use management and intensity. Moreover, within trophic groups, individual taxa generally responded in a similar way, suggesting that trophic groups of protists might reflect the general patterns in different population structures. To efficiently conserve biological diversity and maintain ecosystem functions, it is crucial to monitor diversity changes and investigate the underlying processes related to land use changes. Follow-up studies need to identify the beneficial traits of LUI sensitive protists in order to define the microbiome functions that can be manipulated through agricultural management.

Acknowledgments

Authors would like to thanks Tina Thane, Tanja Begovic and Pia Petersen for laboratory support. The authors declare no conflict of interests.

Funding Information

This work was supported by the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no.

[289949] Trainbiodiverse and eDNA Center financed by Danish Center for Energy and Environment (DCE) at Aarhus University.

References

Adl, S.M., Coleman, D.C., Read, F., 2006. Slow recovery of soil biodiversity in sandy loam soils of Georgia after 25 years of no-tillage management. Agric. Ecosyst. Environ. 114, 323–334. https://doi.org/10.1016/j.agee.2005.11.019

Bates, S.T., Clemente, J.C., Flores, G.E., Walters, W. a, Parfrey, L.W., Knight, R., Fierer, N., 2013. Global biogeography of highly diverse protistan communities in soil. Isme J. 7, 652–659. https://doi.org/Doi 10.1038/Ismej.2012.147

Blanco-Moure, N., Gracia, R., Bielsa, A.C., López, M.V., 2016. Soil organic matter fractions as affected by tillage and soil texture under semiarid Mediterranean conditions. Soil Tillage Res. 155, 381–389. https://doi.org/10.1016/j.still.2015.08.011

Csárdi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJournal Complex Syst. 1695, 1695.

de Araujo, A.S.F., Mendes, L.W., Lemos, L.N., Antunes, J.E.L., Beserra, J.E.A., de Lyra, M. do C.C.P., Figueiredo, M. do V.B., Lopes, Â.C. de A., Gomes, R.L.F., Bezerra, W.M., Melo, V.M.M., de Araujo, F.F., Geisen, S., 2018. Protist species richness and soil microbiome complexity increase towards climax vegetation in the Brazilian Cerrado. Commun. Biol. 1, 135. https://doi.org/10.1038/s42003-018-0129-0

De Cáceres, M., Legendre, P., Moretti, M., 2010. Improving indicator species analysis by combining groups of sites. Oikos 119, 1674–1684. https://doi.org/10.1111/j.1600-0706.2010.18334.x

Degrune, F., Dumack, K., Fiore-Donno, A.M., Bonkowski, M., Sosa-Hernández, M.A., Schloter, M., Kautz, T., Fischer, D., Rillig, M.C., 2019. Distinct communities of Cercozoa at different soil depths in a temperate agricultural field. FEMS Microbiol. Ecol. 95, 1–7. https://doi.org/10.1093/femsec/fiz041

Dupont, A., Griffiths, R.I., Bell, T., Bass, D., 2016. Differences in soil micro-eukaryotic communities over soil pH gradients are strongly driven by parasites and saprotrophs. Environ. Microbiol. 18, 2010–2024. https://doi.org/10.1111/1462-2920.13220

Dupont, A.Ö.C., Griffiths, R.I., Bell, T., Bass, D., 2016. Differences in soil micro-eukaryotic communities over soil pH gradients are strongly driven by parasites and saprotrophs. Environ. Microbiol. 18, 2010–2024. https://doi.org/10.1111/1462-2920.13220

Edgar, R.C., 2016. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. bioRxiv 081257. https://doi.org/10.1101/081257

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10, 996–8. https://doi.org/10.1038/nmeth.2604

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Fiore-Donno, A.M., Weinert, J., Wubet, T., Bonkowski, M., 2016. Metacommunity analysis of amoeboid protists in grassland soils. Sci. Rep. 6, 19068. https://doi.org/10.1038/srep19068

Gao, Z., Karlsson, I., Geisen, S., Kowalchuk, G., Jousset, A., 2019. Protists: Puppet Masters of the Rhizosphere Microbiome. Trends Plant Sci. 24, 165–176.

https://doi.org/10.1016/j.tplants.2018.10.011
Geisen, S., 2016a. The bacterial-fungal energy channel concept challenged by enormous functional versatility of soil protists. Soil Biol. Biochem. 102, 22–25. https://doi.org/10.1016/j.soilbio.2016.06.013
Geisen, S., 2016b. The bacterial-fungal energy channel concept challenged by enormous functional versatility of soil protists. Soil Biol. Biochem. 102, 22–25. https://doi.org/10.1016/j.soilbio.2016.06.013
Geisen, S., Koller, R., Hünninghaus, M., Dumack, K., Urich, T., Bonkowski, M., 2016. The soil food web revisited: Diverse and widespread mycophagous soil protists. Soil Biol. Biochem. 94, 10–18. https://doi.org/10.1016/j.soilbio.2015.11.010
Geisen, S., Mitchell, E.A.D., Adl, S., Bonkowski, M., Dunthorn, M., Ekelund, F., Fernández, L.D., Jousset, A., Krashevska, V., Singer, D., Spiegel, F.W., Walochnik, J., Lara, E., 2018. Soil protists: A fertile frontier in soil biology research. FEMS Microbiol. Rev. 42, 293–323. https://doi.org/10.1093/femsre/fuy006
Geisen, S., Tveit, A.T., Clark, I.M., Richter, A., Svenning, M.M., Bonkowski, M., Urich, T., 2015. Metatranscriptomic census of active protists in soils. ISME J. 9, 2178–2190. https://doi.org/10.1038/ismej.2015.30
George, P.B.L., Lallias, D., Creer, S., Seaton, F.M., Kenny, J.G., Eccles, R.M., Griffiths, R.I., Lebron, I., Emmett, B.A., Robinson, D.A., Jones, D.L., 2019. Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. Nat. Commun. 10, 1–11. https://doi.org/10.1038/s41467-019-09031-1
Glaser, K., Kuppardt, A., Boenigk, J., Harms, H., Fetzer, I., 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.
Griffiths, B.S., Römbke, J., Schmelz, R.M., Scheffczyk, A., Faber, J.H., Bloem, J., Pérès, G., Cluzeau, D., Chabbi, A., Suhadolc, M., Sousa, J.P., Martins Da Silva, P., Carvalho, F., Mendes, S., Morais, P., Francisco, R., Pereira, C., Bonkowski, M., Geisen, S., Bardgett, R.D., De Vries, F.T., Bolger, T., Dirilgen, T., Schmidt, O., Winding, A., Hendriksen, N.B., Johansen, A., Philippot, L., Plassart, P., Bru, D., Thomson, B., Griffiths, R.I., Bailey, M.J., Keith, A., Rutgers, M., Mulder, C., Hannula, S.E., Creamer, R., Stone, D., 2016. Selecting cost effective and policy-relevant biological indicators for European monitoring of soil biodiversity and ecosystem function. Ecol. Indic. 69, 213–223. https://doi.org/10.1016/j.ecolind.2016.04.023
Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., del Campo, J., Dolan, J.R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W.H.C.F., Lara, E., Le Bescot, N., Logares, R., Mahe, F., Massana, R., Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.-L., Siano, R., Stoeck, T., Vaulot, D., Zimmermann, P., Christen, R., 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41, D597–D604. https://doi.org/10.1093/nar/gks1160
Hartman, K., van der Heijden, M.G.A., Wittwer, R.A., Banerjee, S., Walser, J.C., Schlaeppi, K., 2018. Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. Microbiome 6, 74. https://doi.org/10.1186/s40168-018-0456-x
Lentendu, G., Wubet, T., Chatzinotas, A., Wilhelm, C., Buscot, F., Schlegel, M., 2014. Effects of long-term differential fertilization on eukaryotic microbial communities in an arable

soil: a multiple barcoding approach. Mol. Ecol. 23, 3341–3355. https://doi.org/10.1111/mec.12819

Mahé, F., De Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., Singer, D., Mayor, J., Bunge, J., Sernaker, S., Siemensmeyer, T., Trautmann, I., Romac, S., Berney, C., Kozlov, A., Mitchell, E.A.D., Seppey, C.V.W., Egge, E., Lentendu, G., Wirth, R., Trueba, G., Dunthorn, M., 2017. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. Nat. Ecol. Evol. 1. https://doi.org/10.1038/s41559-017-0091

Malik, A.A., Puissant, J., Buckeridge, K.M., Goodall, T., Jehmlich, N., Chowdhury, S., Gweon, H.S., Peyton, J.M., Mason, K.E., van Agtmaal, M., Blaud, A., Clark, I.M., Whitaker, J., Pywell, R.F., Ostle, N., Gleixner, G., Griffiths, R.I., 2018. Land use driven change in soil pH affects microbial carbon cycling processes. Nat. Commun. 9, 1–10. https://doi.org/10.1038/s41467-018-05980-1

McMurdie, P.J., Holmes, S., 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One 8. https://doi.org/10.1371/journal.pone.0061217

Oksanen, A.J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P.R., Hara, R.B.O., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., 2019. Package ' vegan .'

Parks, D.H., Tyson, G.W., Hugenholtz, P., Beiko, R.G., 2014. STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics 30, 3123–3124. https://doi.org/10.1093/bioinformatics/btu494

Peters, M.K., Hemp, A., Appelhans, T., Becker, J.N., Behler, C., Classen, A., Detsch, F., Ensslin, A., Ferger, S.W., Frederiksen, S.B., Gebert, F., Gerschlauer, F., Gütlein, A., Helbig-Bonitz, M., Hemp, C., Kindeketa, W.J., Kühnel, A., Mayr, A. V., Mwangomo, E., Ngereza, C., Njovu, H.K., Otte, I., Pabst, H., Renner, M., Röder, J., Rutten, G., Schellenberger Costa, D., Sierra-Cornejo, N., Vollstädt, M.G.R., Dulle, H.I., Eardley, C.D., Howell, K.M., Keller, A., Peters, R.S., Ssymank, A., Kakengi, V., Zhang, J., Bogner, C., Böhning-Gaese, K., Brandl, R., Hertel, D., Huwe, B., Kiese, R., Kleyer, M., Kuzyakov, Y., Nauss, T., Schleuning, M., Tschapka, M., Fischer, M., Steffan-Dewenter, I., 2019. Climate–land-use interactions shape tropical mountain biodiversity and ecosystem functions. Nature 568, 88–92. https://doi.org/10.1038/s41586-019-1048-z

R Core Team, D., 2013. R: A Language and Environment for Statistical Computing. [WWW Document]. Vienna, Austria R Found. Stat. Comput. 2013. Available http//www.r-project.org/.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2009. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. https://doi.org/10.1093/bioinformatics/btp616

Santos, S.S., Nielsen, T.K., Hansen, L., Winding, A., 2015. Comparison of three DNA extraction methods for recovery of soil protist DNA. J. Microbiol. Methods 115, 13–19. https://doi.org/10.1016/j.mimet.2015.05.011

Schöps, R., Goldmann, K., Herz, K., Lentendu, G., Schöning, I., Bruelheide, H., Wubet, T., Buscot, F., 2018. Land-use intensity rather than plant functional identity shapes bacterial and fungal rhizosphere communities. Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.02711

Schulz, G., Schneider, D., Brinkmann, N., Edy, N., Daniel, R., Polle, A., Scheu, S., Krashevska, V., 2019. Changes in Trophic Groups of Protists With Conversion of Rainforest Into Rubber and Oil Palm Plantations. Front. Microbiol. 10, 1–14.

https://doi.org/10.3389/fmicb.2019.00240
Seppey, C.V.W., Singer, D., Dumack, K., Fournier, B., Belbahri, L., Mitchell, E.A.D., Lara, E., 2017. Distribution patterns of soil microbial eukaryotes suggests widespread algivory by phagotrophic protists as an alternative pathway for nutrient cycling. Soil Biol. Biochem. 112, 68–76. https://doi.org/10.1016/j.soilbio.2017.05.002
Shen, C., Liang, W., Shi, Y., Lin, X., Zhang, H., Wu, X., Xie, G., Chain, P., Grogan, P., Chu, H., 2014. Contrasting elevational diversity patterns between eukaryotic soil microbes and plants. Ecology 95, 3190–3202. https://doi.org/10.1890/14-0310.1
Stoeck, T., Breiner, H.W., Filker, S., Ostermaier, V., Kammerlander, B., Sonntag, B., 2014. A morphogenetic survey on ciliate plankton from a mountain lake pinpoints the necessity of lineage-specific barcode markers in microbial ecology. Environ. Microbiol. 16, 430–444. https://doi.org/10.1111/1462-2920.12194
Szoboszlay, M., Dohrmann, A.B., Poeplau, C., Don, A., Tebbe, C.C., 2017. Impact of land-use change and soil organic carbon quality on microbial diversity in soils across Europe. FEMS Microbiol. Ecol. 93, 1–12. https://doi.org/10.1093/femsec/fix146
Tardy, V., Spor, A., Mathieu, O., Lévèque, J., Terrat, S., Plassart, P., Regnier, T., Bardgett, R.D., van der Putten, W.H., Roggero, P.P., Seddaiu, G., Bagella, S., Lemanceau, P., Ranjard, L., Maron, P.-A., 2015. Shifts in microbial diversity through land use intensity as drivers of carbon mineralization in soil. Soil Biol. Biochem. 90, 204–213. https://doi.org/10.1016/j.soilbio.2015.08.010
Thomson, B.C., Tisserant, E., Plassart, P., Uroz, S., Griffiths, R.I., Hannula, S.E., Buée, M., Mougel, C., Ranjard, L., Van Veen, J. a., Martin, F., Bailey, M.J., Lemanceau, P., 2015. Soil conditions and land use intensification effects on soil microbial communities across a range of European field sites. Soil Biol. Biochem. 88, 403–413. https://doi.org/10.1016/j.soilbio.2015.06.012
Trap, J., Bonkowski, M., Plassard, C., Villenave, C., Blanchart, E., 2015. Ecological importance of soil bacterivores for ecosystem functions. Plant Soil 1–24. https://doi.org/10.1007/s11104-015-2671-6
Voss, C., Fiore-Donno, A.M., Guerreiro, M.A., Peršoh, D., Bonkowski, M., 2019. Metatranscriptomics reveals unsuspected protistan diversity in leaf litter across temperate beech forests, with Amoebozoa the dominating lineage. FEMS Microbiol. Ecol. 1–9. https://doi.org/10.1093/femsec/fiz142
Warnes, A.G.R., Bolker, B., Bonebakker, L., Huber, W., Liaw, A., Lumley, T., Magnusson, A., Moeller, S., Schwartz, M., 2019. Package “gplots” Title Various R Programming Tools for Plotting Data.
Xiong, W., Jousset, A., Guo, S., Karlsson, I., Zhao, Q., Wu, H., Kowalchuk, G.A., Shen, Q., Li, R., Geisen, S., 2017. Soil protist communities form a dynamic hub in the soil microbiome. ISME J. 1–5. https://doi.org/10.1038/ismej.2017.171
Xiong, W., Li, R., Guo, S., Karlsson, I., Jiao, Z., Xun, W., Kowalchuk, G.A., Shen, Q., Geisen, S., 2019. Microbial amendments alter protist communities within the soil microbiome. Soil Biol. Biochem. 135, 379–382. https://doi.org/10.1016/j.soilbio.2019.05.025
Zhao, Z.B., He, J.Z., Geisen, S., Han, L.L., Wang, J.T., Shen, J.P., Wei, W.X., Fang, Y.T., Li, P.P., Zhang, L.M., 2019. Protist communities are more sensitive to nitrogen fertilization than other microorganisms in diverse agricultural soils. Microbiome 7, 1–16. https://doi.org/10.1186/s40168-019-0647-0

List of tables:

Table 1: Edaphic properties and climatic variables among different land uses and LTOs (median ± SE). MAP: Mean average precipitation; MAmT: Mean average maximum temperature. Moisture values are based on residual water content measurements, and all other data calculated on a dry weight basis (g kg^{-1} soil). Significant differences between arable and grassland sites (the values denoted in grey shadow) are indicated by asterisks, while the differences among sites, within each land use, are indicated by lowercase letters according to ANOVA ($p < 0.05$).

Land use / Site	LUI	MAP October (mm)	MAmT October (°C)	Moisture (mL g^{-1})	pH ($CaCl_2$)	Total N (g kg^{-1})	Total C (g kg^{-1})	Organic matter (g kg^{-1})	Clay (%)	Silt (%)	Sand (%)
Arable sites (n=24)	**15.4 ± 3.6*****			**16.9 ± 0.7**	**6.4 ± 0.1**	**0.1 ± 0.01**	**1.2 ± 0.1**	**20.6 ± 1.1**	**19.1 ± 0.9**	**47.4 ± 2.6****	**33.6 ± 3.2**
Lusignan (LU)	**Low - 12.4** Minimum tillage [a] **High - 12.4** Conventional tillage	68	16.4	19.0 ± 0.5 [a] 18.6 ± 0.1	6.5 ± 0.1 [a] 6.2 ± 0.1	0.1 ± 0.01 [a] 0.1 ± 0.02	1.1 ± 0.1 [a] 1.1 ± 0.9	19.6 ± 2.8 [a] 18.8 ± 1.5	18.0 ± 0.2 [a] 18.9 ± 1.1	64.4 ± 1.5 [a] 62.7 ± 3.7	17.5 ± 1.2 [a] 18.4 ± 2.8
Moskanjci (MO)	**Low - 14.8** Minimum tillage [b] **High - 22.1** Conventional tillage	80	15.5	18.3 ± 0.3 [a] 18.8 ± 0.4	7.0 ± 0.3 [b] 6.7 ± 0.1	0.2 ± 0.01 [b] 0.1 ± 0.01	1.6 ± 0.1 [b] 1.5 ± 0.1	28.3 ± 1.8 [b] 25.1 ± 0.9	21.2 ± 0.5 [b] 21.3 ± 0.3	41.9 ± 0.2 [b] 43.1 ± 0.2	36.9 ± 0.3 [b] 35.5 ± 0.5
Castro Verde (CV)	**Low - 0.1** Fallow [c] **High - 10.8** Cereal	63	22.9	12.3 ± 0.1 [b] 10.2 ± 1.3	5.8 ± 0.1 [c] 5.8 ± 0.1	0.1 ± 0.01 [c] 0.1 ± 0.01	0.8 ± 0.1 [c] 0.9 ± 0.1	13.5± 0.5 [c] 14.7 ± 0.4	12.4 ± 0.4 [c] 12.7 ± 0.4	30.9 ± 1.1 [c] 29.5 ± 2.0	56.7 ± 1.5 [c] 57.8 ± 2.3
Sheyern (SH)	**Low - 15.4** Organic [b] **High - 35.4** Conventional	52	12.8	19.2 ± 0.3 [a] 18.9 ± 0.3	6.7 ± 0.1 [b] 6.7 ± 0.1	0.1 ± 0.01 [b] 0.1 ± 0.01	1.4 ± 0.1 [b] 1.5 ± 0.1	25.0 ± 1.9 [b] 20.0 ± 2.9	24.4 ± 1.5 [d] 23.5 ± 1.1	51.9 ± 0.9 [d] 54.6 ± 1.4	23.7 ± 2.5 [d] 21.9 ± 1.5
Grassland sites (n=12)	**3.3 ± 1.5**			**33 ± 0.7*****	**6.4 ± 0.3**	**0.4 ± 0.03*****	**5.2 ± 0.3*****	**88.5 ± 5.5****	**42.6 ± 4.9*****	**34.7 ± 2.5**	**22.7 ± 6.0**
Hainich (HA)	**Low - 3** Extensive grass [a] **High - 7.6** Intensive grass	58	12	29.3 ± 2.5 [a] 29.0 ± 1.8	7.3 ± 0.2 [a] 7.5 ± 0.1	0.4 ± 0.06 0.5 ± 0.08	4.6 ± 0.6 5.2 ± 0.7	79.2 ± 10.9 [a] 86.7 ± 13.9	50.5 ± 5.3 [a] 62.8 ± 1.4	45.9 ± 5.4 32.8 ± 1.5	3.6 ± 0.1 [a] 4.5 ± 0.4
Lancaster (LA)	**Low - 0.4** Extensive grass [b] **High - 2** Intensive grass	114	13.3	36.1 ± 2.1 [b] 37.6 ± 1.7	5.1 ± 0.3 [b] 5.8 ± 0.1	0.4 ± 0.10 0.5 ± 0.05	5.0 ± 0.8 5.9 ± 0.1	85.6 ± 14.3 [b] 102.6 ± 3.4	26.6 ± 4.3 [b] 30.6 ± 3.1	26.3 ± 2.4 33.9 ± 1.4	47.1 ± 6.7 [b] 35.5 ± 4.2

699

Table 2: The most influential environmental factors affecting protist OTU community composition in arable and grassland sites, as determined by permutational multivariate analysis of variance (PERMANOVA). Significant factors are shown in bold ($p < 0.05$).

Arable			Grassland		
	PERMANOVA			PERMANOVA	
	R^2	*p*		R^2	*p*
LUI	**0.19**	**0.0001**	**LUI**	**0.22**	**0.0002**
pH	**0.16**	**0.0001**	**pH**	**0.15**	**0.0034**
Moisture	**0.11**	**0.0006**	**Carbon**	**0.10**	**0.0256**
Organic matter	**0.05**	**0.0375**	Moisture	0.09	0.0552
MAP	**0.05**	**0.0404**	Organic matter	0.09	0.0514
Silt	0.04	0.0515	Silt	0.09	0.0605
Clay	0.04	0.0697	Clay	0.07	0.1886
MAmT	0.03	0.1463	MAP	0.06	0.2691
Nitrogen	0.03	0.1872	Nitrogen	0.04	0.5523
Carbon	0.02	0.3662	MAmT	0.00	0.6572
Residual	0.28		Residual	0.10	

Figures with legends

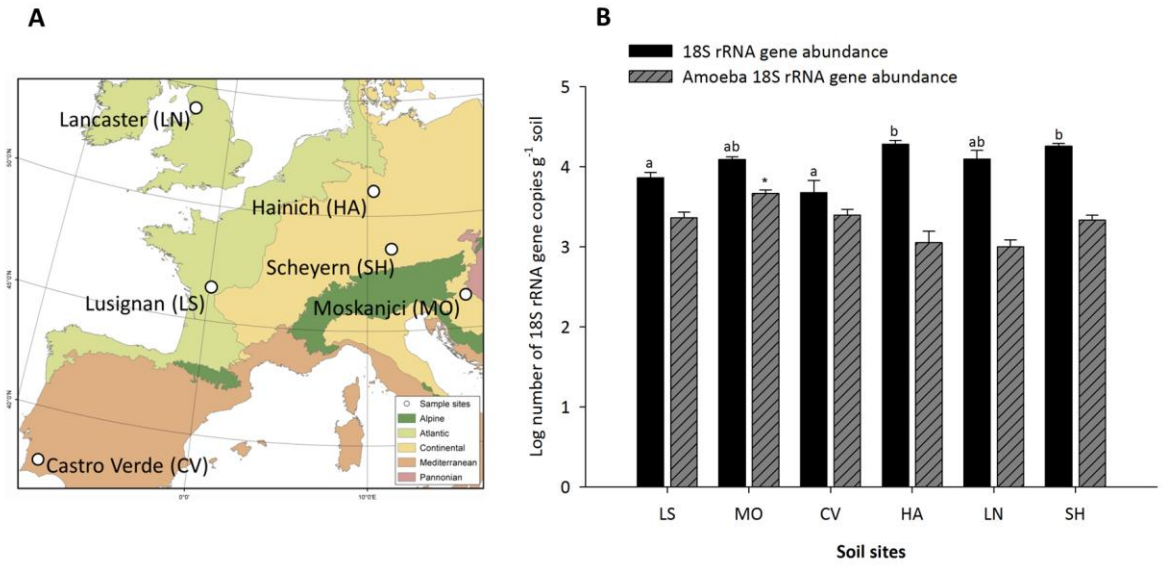


Figure 1: A) European sites used in this study. (B) Quantification of eukaryotic 18S rRNA genes in the different soils. Bars correspond to averages of three replicates ± SE. Bars topped by the same letter are not significantly different at $p < 0.05$ (Tukey). Asterisk indicates significant differences in 18S rRNA gene for free-living amoeba qPCR data.

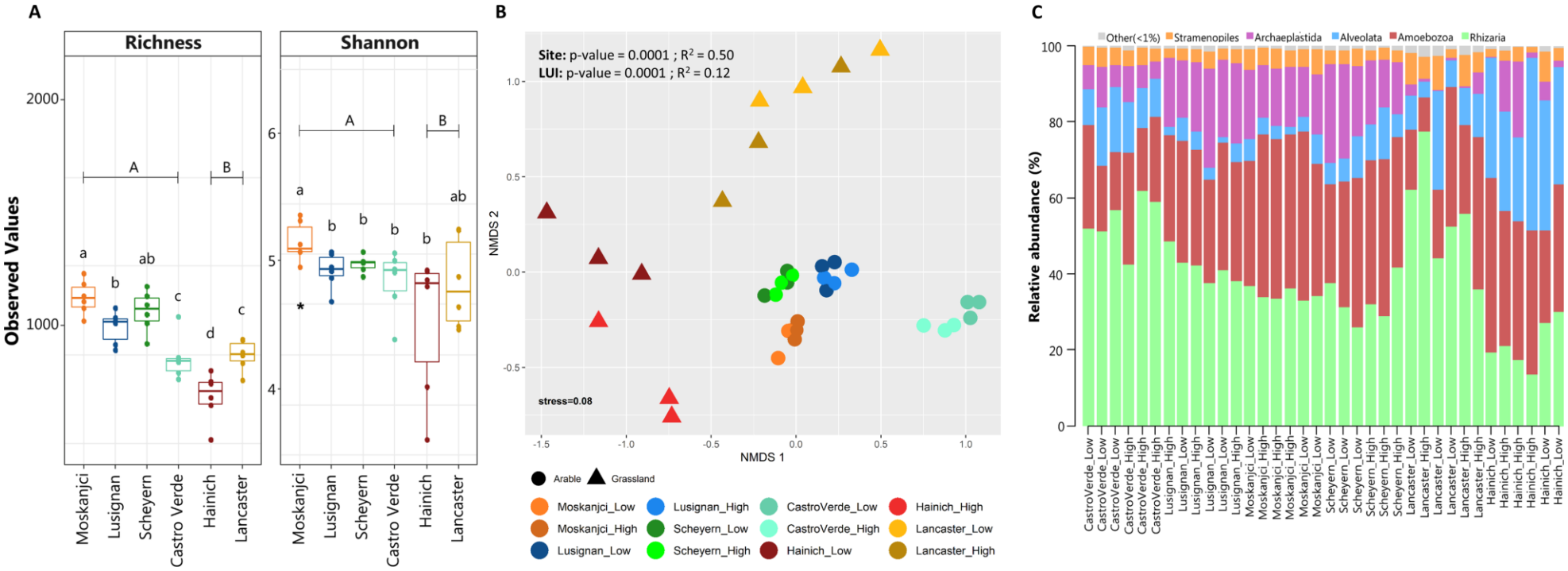
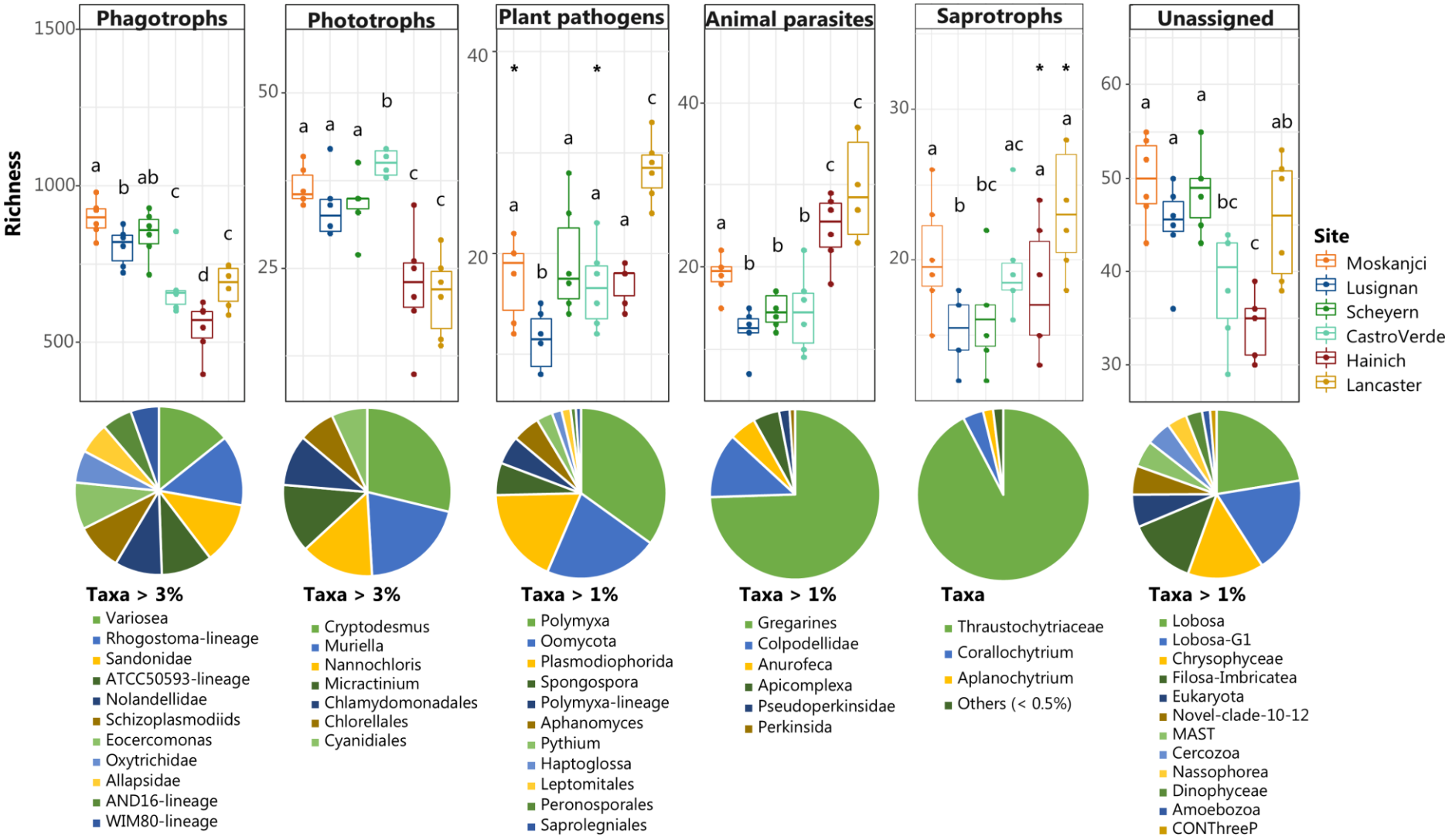


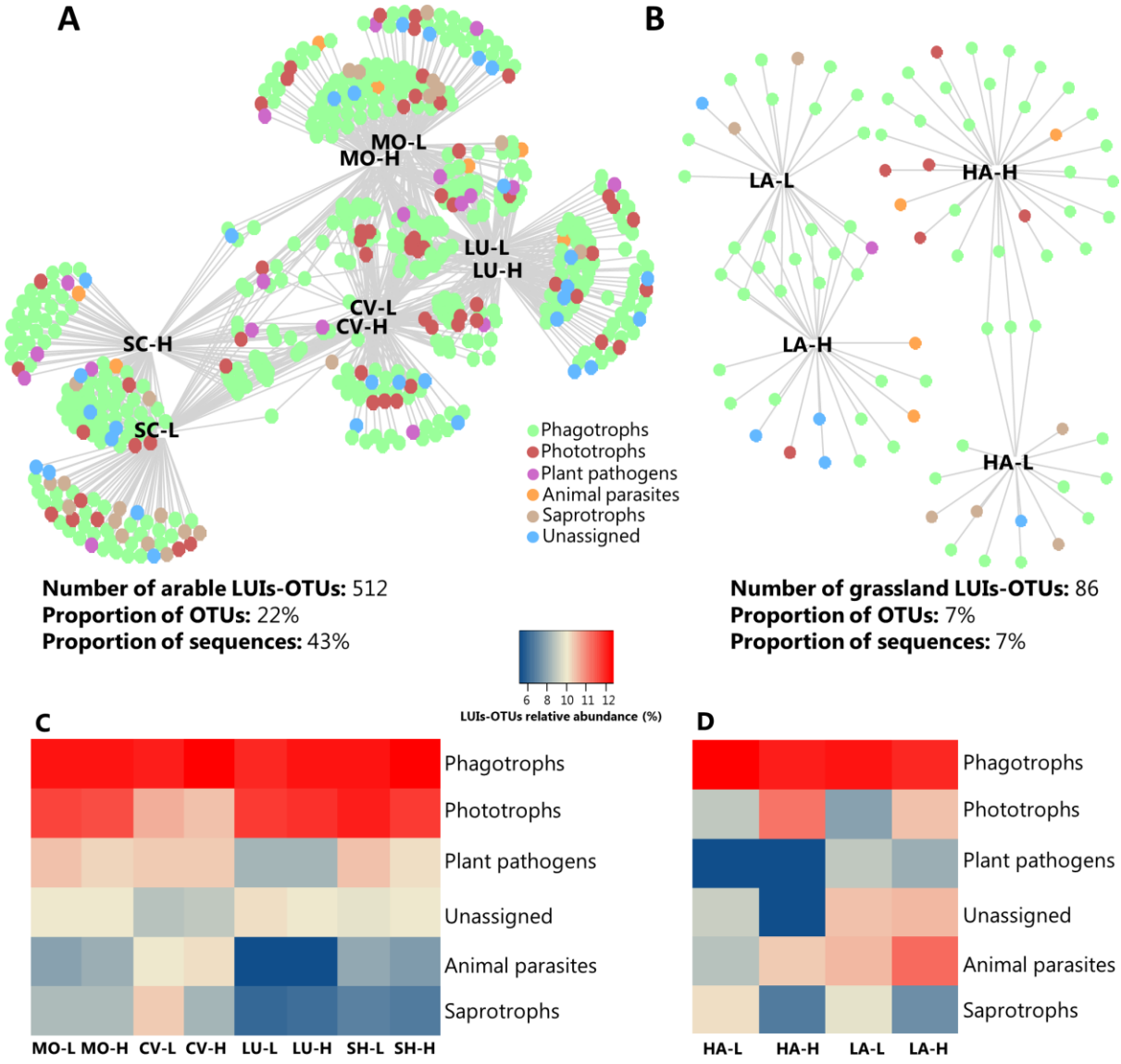
Figure 2: A) Protist OTU richness and Shannon diversity of rarefied protist dataset (the thick horizontal line is the median sample value while the box shows the interquartile range). Statistically significant differences among sites by pairwise ANOVA ($p < 0.05$) are shown above the plot with small case letters. Capital letters show differences among arable and grassland land uses. For Moskanjci site, the asterisk below the box indicates significant differences between high and low agricultural management systems (LUI). B) Beta diversity analysed by non-metric multidimensional scaling (NMDS) based on Bray-Curtis dissimilarities at OTU level and displayed in scatter diagram. The variances were assessed by permutational multivariate analysis of variance (PERMANOVA) based on a distance matrix created using Bray-Curtis method. C) Taxonomic profiles of soil protist communities at super-group level. Protist groups with relative abundances lower than 1% were summarized in the category 'other'. The x-axis sample order reflects a clustering by Bray-Curtis dissimilarities using the *hclust* function in R with method "average".

721


Phagotrophs
Phototrophs
Plant pathogens
Animal parasites
Saprotrophs
Unassigned
Richness
Site
Moskanjci
Lusignan
Scheyern
CastroVerde
Hainich
Lancaster
Taxa > 3%
Variosea
Rhogostoma-lineage
Sandonidae
ATCC50593-lineage
Nolandellidae
Schizoplasmodiids
Eocercomonas
Oxytrichidae
Allapsidae
AND16-lineage
WIM80-lineage
Taxa > 3%
Cryptodesmus
Muriella
Nannochloris
Micractinium
Chlamydomonadales
Chlorellales
Cyanidiales
Taxa > 1%
Polymyxa
Oomycota
Plasmodiophorida
Spongospora
Polymyxa-lineage
Aphanomyces
Pythium
Haptoglossa
Leptomitales
Peronosporales
Saprolegniales
Taxa > 1%
Gregarines
Colpodellidae
Anurofeca
Apicomplexa
Pseudoperkinsidae
Perkinsida
Taxa
Thraustochytriaceae
Corallochytrium
Aplanochytrium
Others (< 0.5%)
Taxa > 1%
Lobosa
Lobosa-G1
Chrysophyceae
Filosa-Imbricatea
Eukaryota
Novel-clade-10-12
MAST
Cercozoa
Nassophorea
Dinophyceae
Amoebozoa
CONThreeP

722

723 **Figure 3:** Relative OTU richness of trophic groups of soil protists per site. Box-plots sharing the same letter do not differ significantly (ANOVA, $p < 0.05$). Asterisk indicates significant differences between high and low LUI treatments within site (ANOVA, $p < 0.05$). Pie-charts show the most
725 abundant taxa per trophic group.


A
MO-L
MO-H
LU-L
LU-H
CV-L
CV-H
SC-H
SC-L
Phagotrophs
Phototrophs
Plant pathogens
Animal parasites
Saprotrophs
Unassigned
Number of arable LUIs-OTUs: 512
Proportion of OTUs: 22%
Proportion of sequences: 43%
B
LA-L
HA-H
LA-H
HA-L
Number of grassland LUIs-OTUs: 86
Proportion of OTUs: 7%
Proportion of sequences: 7%
6 8 10 11 12
LUIs-OTUs relative abundance (%)
C
Phagotrophs
Phototrophs
Plant pathogens
Unassigned
Animal parasites
Saprotrophs
MO-L MO-H CV-L CV-H LU-L LU-H SH-L SH-H
D
Phagotrophs
Phototrophs
Plant pathogens
Unassigned
Animal parasites
Saprotrophs
HA-L HA-H LA-L LA-H

726

Figure 4: Networks display protist LUIs-OTUs per site in the arable (A) and grassland (B) soils (determined using indicator species analysis and likelihood ratio tests with *edgeR*). Dots represent OTUs positively and significantly associated ($p < 0.05$). Dot colour represent the different trophic groups. Connections are given by connecting lines. Heatmaps showing the mean relative abundances of LUIS-OTUs for arable sites (C) and grassland sites (D). OTUs are labelled by their trophic group.

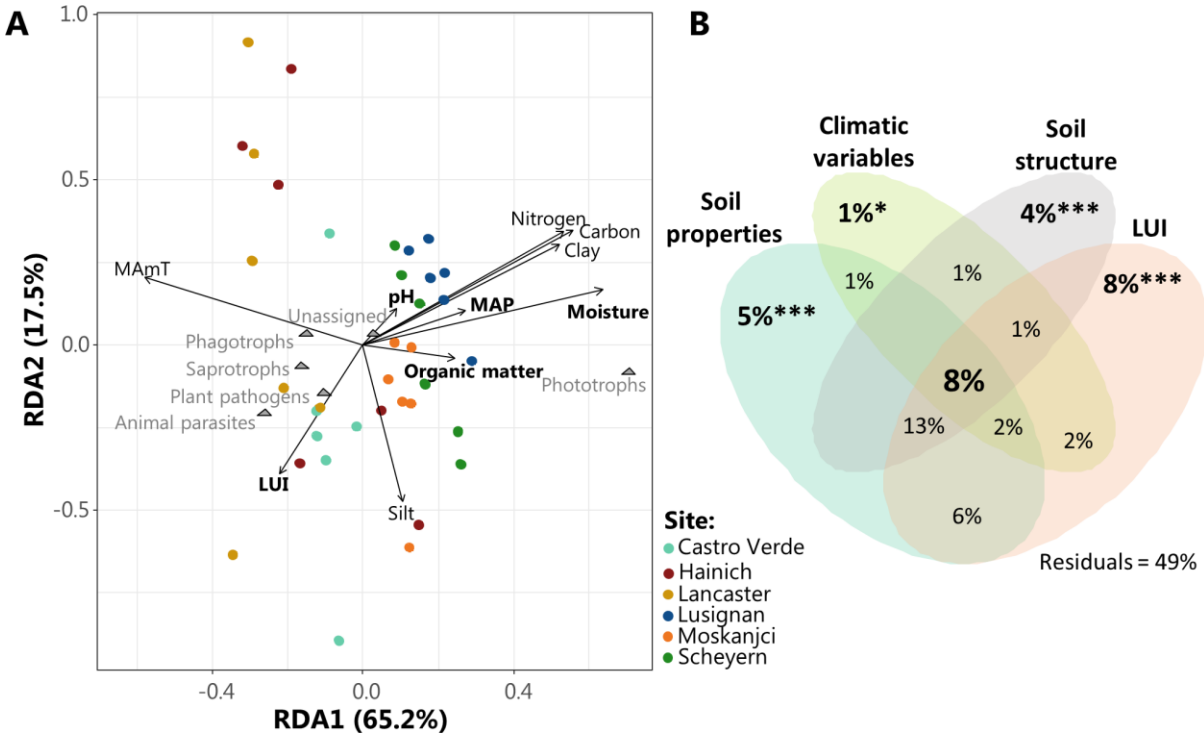


Figure 5: Redundancy analysis (RDA) on Hellinger-transformed data (A) with variance partitioning of the protist functional groups with soil and climatic variables for individual samples (B). The proportion of variance explained by the environmental factors shown is ≈ 50%. The position of trophic groups represents their centroid. The variance partitioning p-values are from permutation tests (1000 permutations) assessing if the partial effect of a factor is significantly different from zero (ANOVA, $p < 0.05$). Values show the fraction of variation explained by each parameter, as well as the shared contribution of each of the parameters.

738

Supplementary material for on-line publication only

Click here to download Supplementary material for on-line publication only: SupplementaryMaterial_SANTOSetal2019_Protists.d

Conflict of interest for submitted manuscript “Land use as a driver for protist community structure in soils under agricultural use across Europe”

Roskilde, December 3rd 2019.

Dear Editor,

We declare that all authors approved the manuscript submission and we have no conflict of interest.

Sincerely,

Anne Winding