Received: 6 March 2023 | Revised: 5 June 2023 | Accepted: 12 June 2023

DOI: 10.1111/gcb.16871

Global Change Biology WILEY

RESEARCH ARTICLE

Ecosystem type drives soil eukaryotic diversity and composition in Europe

Julia Köninger[1,2] | Cristiano Ballabio[2] | Panos Panagos[2] | Arwyn Jones[2] | Marc W. Schmid[3] | Alberto Orgiazzi[2] | Maria J. I. Briones[1]

[1]Departamento de Ecología y Biología Animal, Universidade de Vigo, Vigo, Spain

[2]European Commission, Joint Research Centre, Ispra, Italy

[3]MWSchmid GmbH, Glarus, Switzerland

Correspondence
Maria J. I. Briones, Departamento de Ecología y Biología Animal, Universidade de Vigo, Vigo 36310, Spain.
Email: mbriones@uvigo.es

Alberto Orgiazzi, European Commission, Joint Research Centre, Ispra, Italy.
Email: alberto.orgiazzi@gmail.com

Funding information
Collaborative Doctoral Partnership JRC and UVIGO, Grant/Award Number: Agreement No. 35533

Abstract

Soil eukaryotes play a crucial role in maintaining ecosystem functions and services, yet the factors driving their diversity and distribution remain poorly understood. While many studies focus on some eukaryotic groups (mostly fungi), they are limited in their spatial scale. Here, we analyzed an unprecedented amount of observational data of soil eukaryomes at continental scale (787 sites across Europe) to gain further insights into the impact of a wide range of environmental conditions (climatic and edaphic) on their community composition and structure. We found that the diversity of fungi, protists, rotifers, tardigrades, nematodes, arthropods, and annelids was predominantly shaped by ecosystem type (annual and permanent croplands, managed and unmanaged grasslands, coniferous and broadleaved woodlands), and higher diversity of fungi, protists, nematodes, arthropods, and annelids was observed in croplands than in less intensively managed systems, such as coniferous and broadleaved woodlands. Also in croplands, we found more specialized eukaryotes, while the composition between croplands was more homogeneous compared to the composition of other ecosystems. The observed high proportion of overlapping taxa between ecosystems also indicates that DNA has accumulated from previous land uses, hence mimicking the land transformations occurring in Europe in the last decades. This strong ecosystem-type influence was linked to soil properties, and particularly, soil pH was driving the richness of fungi, rotifers, and annelids, while plant-available phosphorus drove the richness of protists, tardigrades, and nematodes. Furthermore, the soil organic carbon to total nitrogen ratio crucially explained the richness of fungi, protists, nematodes, and arthropods, possibly linked to decades of agricultural inputs. Our results highlighted the importance of long-term environmental variables rather than variables measured at the time of the sampling in shaping soil eukaryotic communities, which reinforces the need to include those variables in addition to ecosystem type in future monitoring programs and conservation efforts.

KEYWORDS
environmental drivers, eukaryotes, LUCAS survey, metabarcoding, soil biodiversity, soil fauna

1 | INTRODUCTION

Understanding the baseline of soil biodiversity status is a crucial requirement for predicting and preventing its declines/losses and when designing protection and conservation management strategies. The spatial heterogeneity of soil organisms and the lack of comprehensive large-scale soil monitoring efforts hamper our current understanding of the factors driving species abundance and diversity, and their geographical patterns. Despite rapid advances in molecular methods for measuring soil biodiversity (e.g., metagenomics and metabarcoding; Orgiazzi et al., 2015), the majority of the research studies have mainly concentrated on fungal and bacterial communities. Only recently, analyses have expanded to include protists (Bates et al., 2013; Oliverio et al., 2020; Xiong et al., 2021) and soil fauna (Aslani et al., 2022; Luan et al., 2020; Schulz et al., 2019). These studies revealed that the impact of land use intensification varies depending on the organism under consideration, with microbial communities appearing to be less susceptible (George et al., 2019; Schulz et al., 2019) compared to animals (Schulz et al., 2019; Tsiafouli et al., 2015). Also, most studies have merely focused on individual eukaryotic groups, such as for example, nematodes (Treonis et al., 2018), collembolans (Bahrndorff et al., 2018), or protists (Malard et al., 2022; Singer et al., 2021). The results from the few works available that included a combination of different organisms show great heterogeneity in the influential factors shaping their communities. Thus, while there is a general consensus on land use being the main driver of α-diversity and soil pH and precipitation being the best predictors of β-diversity (Aslani et al., 2022; George et al., 2019), contrasting results have been reported in relation to soil properties depending on the eukaryotic group investigated (e.g., no effect on animals but positive effects on protists and fungi (George et al., 2019)). Crucially, most of these analyses have been limited to a few sampling sites (Ritter et al., 2018; Wu et al., 2011) or concentrated in specific regions (George et al., 2019) or habitats (Aslani et al., 2022; Fournier et al., 2020; Malard et al., 2022) or did not distinguish between different animal taxa (George et al., 2019; Schulz et al., 2019). Despite the fact that biodiversity is not restricted to natural systems, wide-scale analysis of eukaryotic diversity across different ecosystem types (including managed systems) at continental scale has never been attempted.

In this study, we sequenced the 18S-rDNA gene from soil samples collected from 787 sites across the European Union (EU) plus the United Kingdom (UK) along a land use gradient, decreasing in intensity from annual and permanent croplands, managed and unmanaged grasslands to broadleaved and coniferous woodlands (Table S1) to build a first baseline for soil eukaryotic diversity in Europe. We examined whether plant cover, topography, sampling month, historical land use information, as well as biological, edaphic and climatic properties, are potential drivers of soil eukaryotic distribution patterns in European soils (Table S2). Since this study is part of a large-scale soil assessment (LUCAS Soil Survey (Orgiazzi et al., 2018)), we were able to include less-investigated soil variables that have been previously found to be important drivers of soil biodiversity (i.e., bulk density Devigne et al., 2016) and erosion risk (Panagos et al., 2015).

We addressed the following research questions: (i) How diverse are soil eukaryotes across different ecosystem types in Europe? (ii) Which are the main drivers for the observed diversity within (α-diversity) and between (β-diversity) sites? (iii) Based on the gained knowledge, could the quality of future monitoring/conservation efforts be improved?

2 | MATERIALS AND METHODS

2.1 | Soil sampling

Soil samples were collected as part of the Statistical Office of the European Union's (EUROSTAT) Land Use Change Area Survey (LUCAS) between April and December 2018. LUCAS regularly monitors land use and cover changes in the European Union (EU) at >270,000 points chosen from an ideal 2×2 km grid cell across Europe (Orgiazzi et al., 2018). However, from 2009 onwards, soil physicochemical properties were measured in a subset of 20,000 sites, and in 2018, soil biodiversity analyses were included for the first time at 1000 sites in all EU member states (except for Malta and Luxemburg) plus the UK (Orgiazzi et al., 2022).

Each LUCAS soil sample consists of five topsoil (0–20 cm) subsamples, collected at the exact coordinate along four cardinal directions (north, east, south, and west) in a 2-m distance. Due to seasonal bias, we excluded the winter samples ($n=7$) from our analyses. Therefore, the soil samples collected between April and November 2018 were included in our analyses (see Figure 1b,c for a spatial and temporal overview). To mirror a representative subsample of the European land cover, including croplands, woodlands, grasslands, shrublands, bare land, and wetlands, conditioned Latin hypercube sampling (CLHS; Minasny & McBratney, 2006) was used to select points of the original LUCAS dataset (with some limitations related to the sample size; Orgiazzi et al., 2018). For statistical comparisons, we focused on three main ecosystem types: croplands, grasslands, and woodlands, which reduced the total number of sampled sites to 787. We distinguished between permanent croplands and annual croplands, managed grasslands and grasslands with sparse woody vegetation (trees/shrubs), and broadleaved and coniferous woodlands (Figure 1; Table S1).

The intensity gradient of land use was established according to land use classes established in LUCAS (for a detailed overview see Table S1). Among the sampled sites, more than 90% of coniferous woodlands are used for wood production, whereas within broadleaved woodlands, 27% were located in agricultural stripes, gardens or community centers, as well as in recreational or abandoned areas (Table S1). We, therefore, assumed forest management to less frequently disrupt soils (compared to recreational areas, agricultural stripes, etc.) and hence, categorized broadleaved woodlands as more intensely managed than coniferous woodlands. To account for the series of land use in the three survey years (2009, 2015, and 2018),

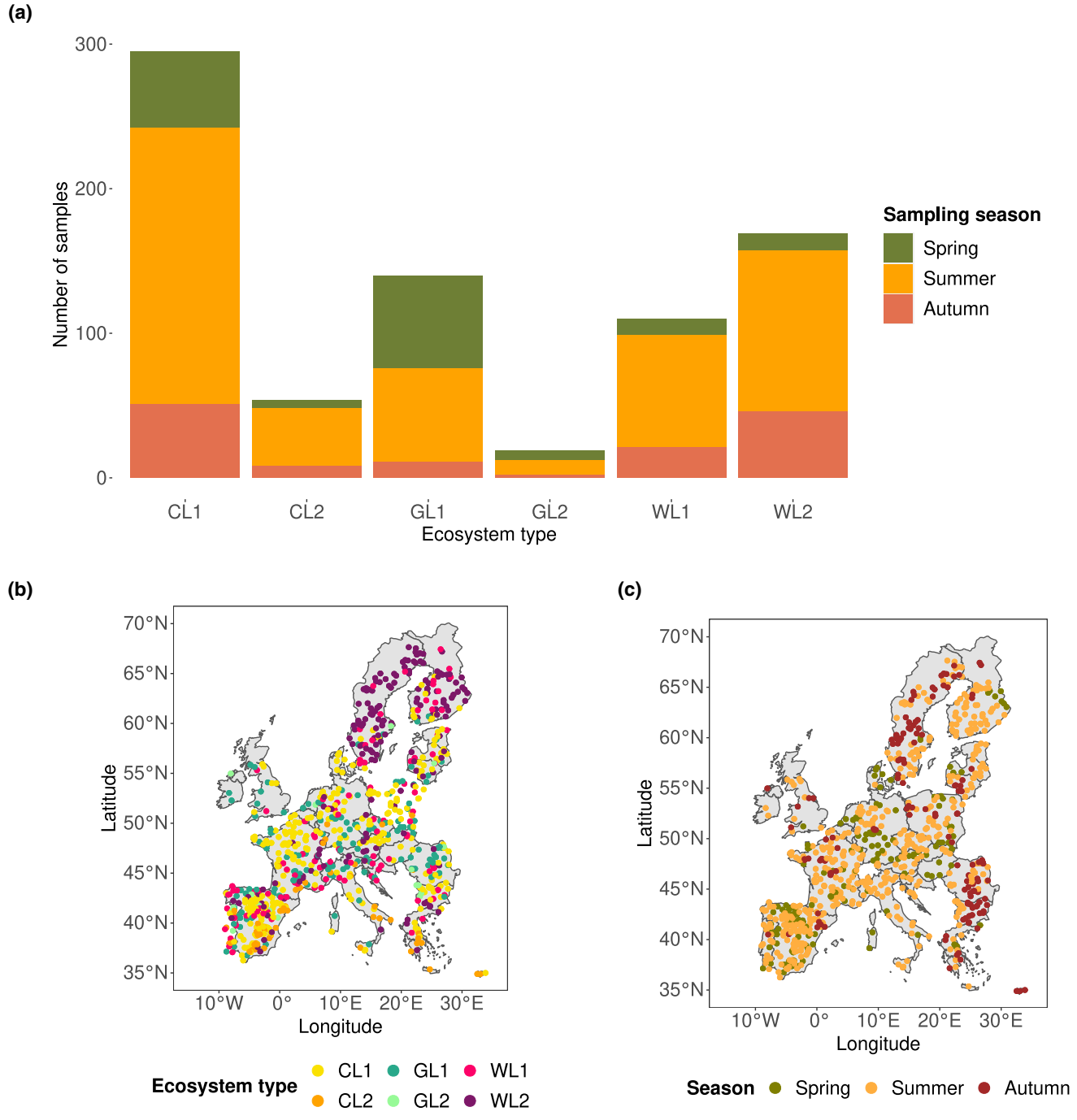


FIGURE 1 Distribution of the sampling sites along ecosystem types and sampling season. (a) Spatio-temporal overview of the sampling campaigns performed at the 787 sites included in this study. (b) Distribution of sampling sites among the different ecosystem types investigated. (c) Distribution of sampling sites according to sampling season. CL1, annual croplands; CL2, permanent croplands; GL1, managed grasslands; GL2, unmanaged grasslands; WL1, broadleaved woodlands; WL2, coniferous woodlands.

we created the character variable listing the shortcuts of the land use in the three surveys, named "intensity gradient of land use between 2009 and 2018." For plant cover, we included information on the density of grasses, crops, and understory vegetation in woodlands, which were derived from photos taken on site during the sampling campaigns (for classification, see Table S2).

2.2 | Environmental drivers

Several environmental variables were used to identify the main drivers of eukaryotic distribution including soil physical and chemical variables, topography (elevation), climate (average monthly precipitation and temperatures), and land cover (for a detailed overview, see Table S1).

Briefly, soil properties (total extractable potassium and plant-available phosphorus contents, clay [%], sand [%], coarse fragments [%], oil organic C to total N ratio [Corg:N], electric conductivity [dS/m], bulk density [g/cm^3], carbonates [g/kg^{-1}], and water content [% of fresh soil weight]) were measured following the International Organization for Standardization (ISO) methods (Orgiazzi et al., 2018). Cmic (μg Cmic g soil dw^{-1}), basal respiration (μl O_2 h^{-1} g soil dw^{-1}), respiration quotient (qO_2), and water content were obtained from a previous study performed on the same sampling sites (Smith et al., 2021). Erosion risk (tha^{-1}year^{-1}) was derived from Panagos et al. (2015). Bulk density (0–20 cm) was added as a measured attribute in LUCAS 2018 (Orgiazzi et al., 2018). Climatic variables were obtained from the WorldClim dataset as averaged values for the period 1970–2000 (Fick & Hijmans, 2017). In addition, we also included climatic variables measured in the previous 20 years as well as in the sampled year (annual average temperature and rainfall, average temperature/rainfall/aridity in the sampling month). Data on soil depth (cm) were extracted from SOILGrids (Poggio et al., 2021), and surface soil moisture (mm) and soil moisture abnormality from the NASA-USDA database (NASA-USDA, 2015). Considered variables are described in more detail in Table S2.

2.3 | DNA sequencing

Five subsamples from each soil sample were prepared for DNA extraction and sequencing. The samples were stored at −20°C upon arrival at the University of Tartu (Estonia) but stored at 4°C before analyses to allow for controlled thawing.

We analyzed the hypervariable 4 region (V4) of the eukaryotic 18S, amplified with the primers Euk575F (ASCYGYGGTAAYWC CAGC) and Euk895R (TCHNHGNATTTCACCNCT). To prevent amplification biases, a fourfold greater amount of the reverse primer for optimal performance was applied. For sequencing the eukaryotic SSU, we used Illumina MiSeq.

For DNA extraction, the Qiagen DNeasy PowerSoil HTP 96 Kit Q12955-4 was used. After extracting three 0.2 g aliquots per sample, the subsamples were pooled. We used negative and positive controls during extraction to detect potential external contamination and cross-contamination. Forward and reverse primers were pooled and visualized on TBE 1% agarose gel. PCR products were purified using UltraClean 96 PCRCleanup Kit (Qiagen). Quality check and quantification of DNA were performed with Qubit™ 1× dsDNA HS Assay Kit using Qubit 3 fluorometer on 12 replicates. Samples with <50,000 reads were re-sequenced. Detailed protocols for DNA extraction, amplification, and sequencing can be found in Orgiazzi et al. (2022).

2.4 | Bioinformatic pipeline, data processing, and quality check

Targeting the 18S gene, the Illumina amplicon data were demultiplexed using LotuS with default options (Ozkurt et al., 2021). Paired-end reads were assembled using FLASH 1.2.10 (Magoč & Salzberg, 2011) with default options (minimum overlap 10 bp). Unmerged reads were removed from the final datasets. The remaining merged reads were truncated up to the 18S primer sequences and filtered for the presence of both primer sequences with a custom Python script allowing for up to two mismatches per primer. We considered OTUs with a 100% similarity as species proxies and clustered them into amplicon sequence variants (ASVs; Oliverio et al., 2020) using DADA2 (version 1.16.0; Callahan et al., 2016). Only considering ASVs occurring more than one time over all sites (ASVs >1), we obtained 116,921 ASVs. Retained primers were clipped (Callahan et al., 2016; Oliverio et al., 2020). Because no comparable studies exist for 18S data, we followed the methods proposed in previous studies (e.g., Brandt et al. (2021)) and filtered 18S data for single sequences occurring in the dataset (singletons).

ASVs were annotated using the taxonomy data available from SILVA (version 138, SSU, NR99; Yilmaz et al., 2014). We filtered the taxonomy data to exclude those eukaryotic sequences that were not soil organisms (e.g., humans, fish). In this study, we considered those organisms that live inside the soil for at least a life stage, excluding eggs. We also excluded plants. We manually corrected the filtered results in the following cases: (i) when placed in the wrong kingdom or (ii) when not living in soil (e.g., living in marine/freshwater habitats). Taxonomic grouping was quality checked based on literature and expert knowledge. The fungal taxonomic ranking was based on Tedersoo et al. (2018).

2.5 | Statistical analyses

For the statistical analysis of the sequences, we classified the eukaryote data into fungi, protists, and animals (Jurburg et al., 2021). Due to the higher morphological heterogeneity in animals, we grouped their sequences according to the following taxonomical phyla groups: rotifers, tardigrades, arthropods, nematodes, and annelids. Platyhelminthes, mollusks, and gastrotricha were not well represented to further consider them in the analyses. To ensure reproducibility, ASV counts were normalized using the scaling with ranked subsampling (Beule & Karlovsky, 2020) through the "SRS" R package. SRS curves were drawn, and the minimum counts of the different samples (Cmin) were set for the different groups with the help of the SRS shiny app (Figure S12). After this step, the sample size for the eukaryotic groups was reduced to 787 sites for fungi, 785 sites for protists, 315 sites for rotifers, 589 sites for tardigrades, 587 sites for nematodes, 583 sites for arthropods, and 666 sites for annelids.

The Shapiro–Wilk test was used to check for normality of the data. Since not all eukaryotic groups followed a normal distribution, the nonparametric Kruskal–Wallis test was used for testing whether ecosystem features (ecosystem type, plant cover, soil depth, erosion risk, climate) significantly affected α-diversity metrics (observed richness and Shannon diversity). When significant differences were detected, the Dunn post-hoc test was conducted with adjusted

p-values using the Benjamini–Hochberg method. Several factors were subdivided into categorical classes: ecosystem type and plant cover, soil depth, pH, erosion risk, sampling season, and climatic variables (Table S1). For significantly different groups, we compared the α-diversity median value to identify the group with the highest/lowest α-diversity.

We applied random forest models to determine which environmental variables could better explain α-diversity and to account for the variation explained by the observed richness (R package *ranger*). A backward selection procedure of the standardized variables was performed by applying the recursive feature elimination (RFE) function of the *caret* R package. Multivariate analyses were used to describe patterns in the community structure of the taxa of the studied eukaryotic groups in response to ecosystem features.

Prior to the β-diversity analyses and the distance-based redundancy analysis (dbRDA), we standardized the environmental variables (*z*-score) and then filtered them to prevent overfitting and multicollinearity. This resulted in soil (silt content), topography (elevation, aspect, slope, erosion risk, and bioregions) together with temporal (sampling season) variables being excluded from the analyses. We gradually filtered autocorrelated variables until all variables' variance inflation factor (VIF) was <12. Variable selection was supported by the Akaike information criteria (AIC; Table S3).

All RDA models were tested for significance using the permutational ANOVA test of the *vegan* R package (Oksanen et al., 2020). All variables selected were found to be significant, and hence, the forward selection of the "*ordiR2step*" function of the *vegan* package was used with three stopping rules: (1) once the adjusted R^2 starts to decrease; (2) when the full model adjusted R^2 is exceeded, and (3) if the preselected permutational significance level ($p < .05$) is exceeded (Oksanen et al., 2020). Results of the ANOVA forward selection are shown in Table S4. The ordination was performed using *vegan*'s *dbrda* function and created the plots with the *ggord* package. We applied the *varpart* function to assess the partitioning of the variance explained by the environmental variables (the same variables as used for the dbRDA analysis). We assessed the spatial dependence of the data by applying distance-based Moran's Eigenvector Maps (dbMEM; *adespatial* R package). Finally, we calculated the fraction of the variation explained by spatial correlation and the other environmental variables (unique and shared effects) with adjusted R^2 values using the *varpart* function (*vegan* package).

We performed a multivariate analysis to test the homogeneity of dispersion (PERMDISP) and the effect of community partitioning on the Bray–Curtis distances using *betadisper* function (Table S5). The results were visualized using boxplots (Figure S10). ANOVA was used to test whether these similarity distance values significantly differed between eukaryotic groups (Table S6). In addition, we ran a pairwise comparison with *permutest* to see if there were significant differences within each group. The permutational analysis (PERMANOVA) was used to calculate the variance associated with β-diversity (*adonis* function, *vegan* R package). In this step, we used Bray–Curtis dissimilarity distances between sites and used *permutest* to detect significant differences between groups and their mean dispersions (Table S7).

We considered the 5% significance level in all statistical tests. Statistical analyses were performed with R version 4.2. To ensure reproducibility and inclusion into other data sets, laboratory protocols, downstream bioinformatics scripts, and study outputs have been deposited on GitHub.

3 | RESULTS

3.1 | Taxonomical assignment and proportion of shared taxa between ecosystem types

We found 97% of the 79,383 ASVs to be assigned to eukaryotes. We further classified the 97 eukaryotic phyla into three kingdoms: protists (57%), fungi (33%), and animals (10%). Within the latter group, 52.2% of reads belonged to nematodes, 32.9% to arthropods, 10.6% to rotifers, 3.4% to tardigrades, and 0.8% to annelids (Figure 2). For fungi, the dominant groups were two phyla belonging to the Dikaria subkingdom, namely Ascomycota (approximately 50%) and Basidiomycota (>35%), followed by the subphylum Mucoromycotina (>10%; Figure S1). The highest relative abundance for protists was found for two phyla, namely Cercozoa (>50%) and Ciliophora (20%), and the superphylum Heterokonta (>10%).

Missing taxonomic information was particularly high for protists and arthropods on order level (72% and 48.5% of the total number of ASVs, respectively; Figure S1). This lack of information increased further at lower taxonomic levels. For example, at the species level, 86.7% of rotifers, 85% of protists species, 79.3%

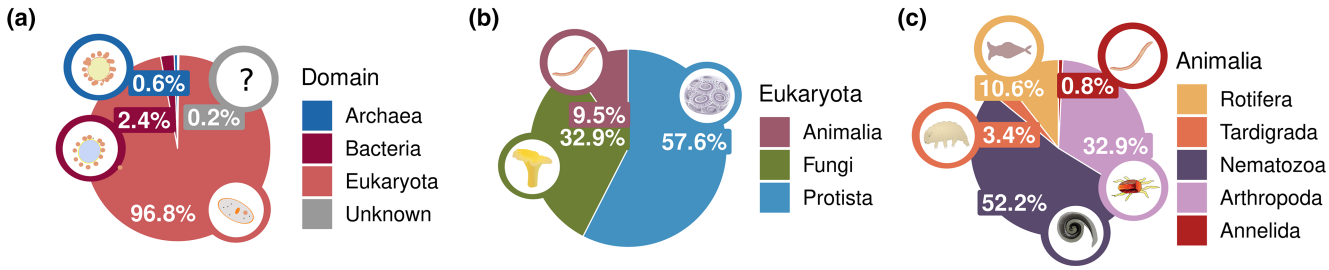


FIGURE 2 18S-rDNA sequencing read results. (a) Proportion of ASVs assigned to different domains, with the majority of sequence reads found to belong to eukaryotes. (b) Proportion of taxa assigned to the different eukaryotic kingdoms. (c) Proportion of taxa assigned to the different phyla.

of arthropod species, 87.8% of tardigrade species, and 69.9% of nematode species could not be assigned to any known sequence (unknown).

The analysis of the ecosystem specificity of taxa indicated that the greatest differences occurred between coniferous woodlands and the other ecosystem types (Figure S1). For all eukaryotic groups, except for annelids, the highest number of taxa was found in croplands (Figure 3). We also observed a high proportion of ASVs being shared by two or more ecosystems (Figure 3). For all eukaryotic groups, the highest number of taxa was shared between the three ecosystems (22.6% for fungi, 24.1% for protists, 30.6% for rotifers, 41.2% for tardigrades, 30.7% for nematodes, 35% for arthropods,

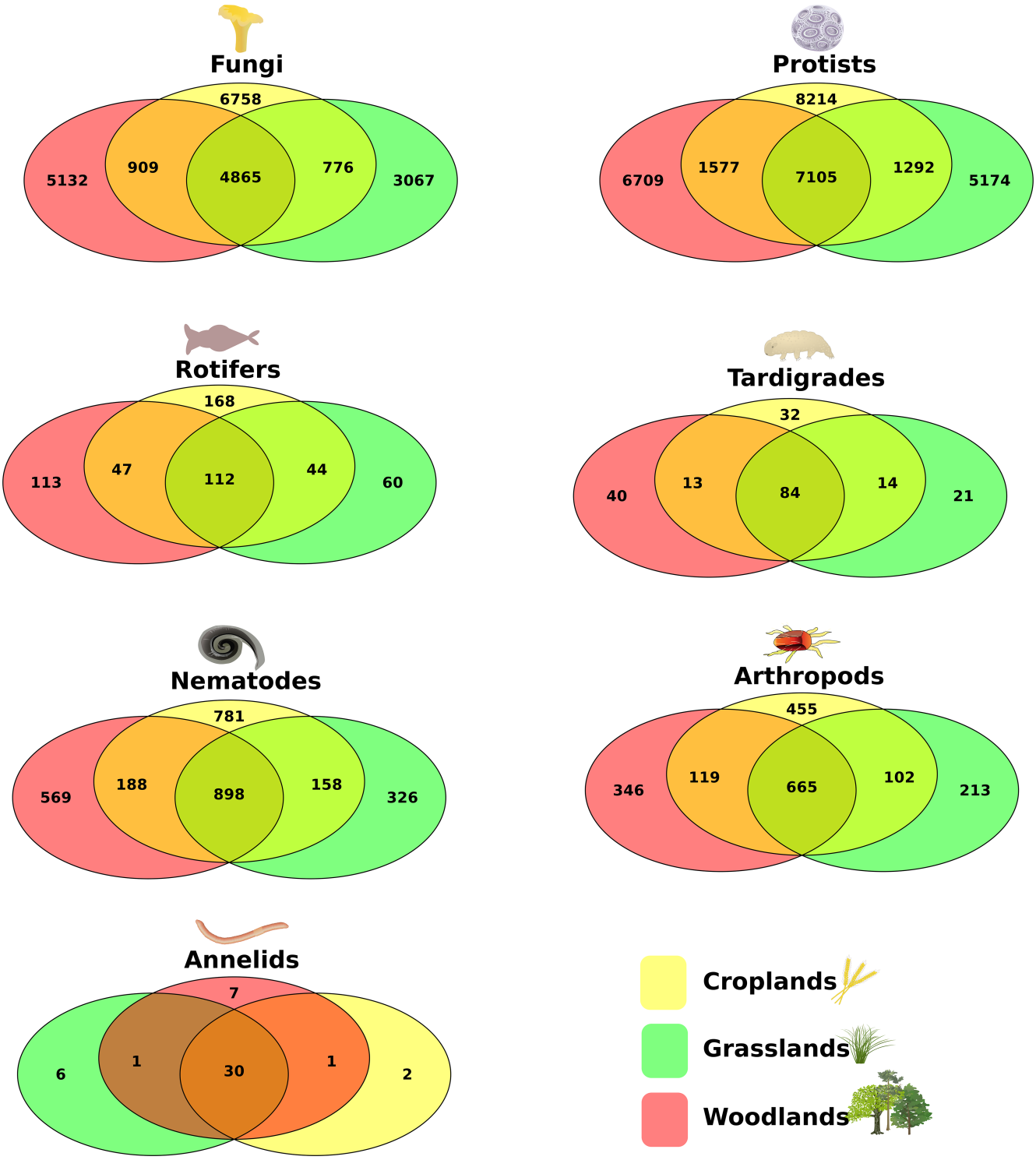


FIGURE 3 Proportion of taxa (ASVs) shared between different ecosystem types. Overlaps of taxa in croplands (yellow), grasslands (green), and woodlands (pink) ranged from 22.6% for fungi, 24.1% for protists, 30.6% for rotifers, 41.2% for tardigrades, 30.7% for nematodes, 35% for arthropods to 63.8% for annelids.

and 63.8% for annelids). For fungi, protists, rotifers, nematodes, and arthropods the highest amount of niche taxa was found in croplands, while for tardigrades and annelids, more specialized groups were found in woodlands.

3.2 | α-diversity revealing the unexpected: Eukaryotic diversity increases with the intensity gradient

The α-diversity values for fungi, protists, nematodes, arthropods, and annelids were significantly different between ecosystem types (Kruskal–Wallis $p < 0.05$, Figure 4). The highest richness (number of ASVs) was measured for fungi (495), followed by protists (484), nematodes (77), and arthropods (59) (Figure 4a,b). In contrast, the lowest values were observed for rotifers (19), annelids (15), and tardigrades (12) across all ecosystem types (Figure 4b). Furthermore, α-diversity declined toward the less intensely managed ecosystems, with the highest values being observed in annual croplands and the lowest ones in coniferous woodlands (Figure S7a). Despite both metrics (ASV richness and Shannon index) showing similar patterns, we found fewer significant differences when comparing the Shannon index between ecosystems. For example, no significant differences in the Shannon diversity of rotifers between different ecosystem types were found.

The effect of ecosystem type on eukaryotic diversity was further confirmed by analyzing the impact of other variables on the α-diversity, such as plant cover and soil properties (pH, erosion risk, soil depth; Kruskal–Wallis $p < .05$; Figure S7b–f). Thus, the highest richness and diversity of fungi, protists, tardigrades, and arthropods were measured in soils with neutral or alkaline pH values (Figure S7c), covered by crops or grasses (Figure S7c). Also, deeper soils contained more diverse and richer eukaryotic communities of fungi, protists, tardigrades, and arthropods (for the latter group only richness; Figure S7d; $p > .05$). In the case of fungi, protists, arthropods, and annelids, the highest richness was found in soils subjected to medium erosion risk of 1–5 $t\,ha^{-1}\,year^{-1}$ (Figure S7e).

Seasonal effects (sampling season) were also found to significantly impact the diversity and richness of protists as well as the diversity of arthropods and the richness of fungi (Kruskal–Wallis $p < 0.05$ – Figure S7f). Accordingly, we found higher α-diversity values of these three eukaryotic groups in spring than in autumn (Figure S7f). Regarding climatic variables, very low and low annual mean temperatures between 1970 and 2000 had a significant negative effect on fungi and protists diversity, as well as on the richness of rotifers, nematodes, and arthropods (Kruskal–Wallis $p < .05$). In contrast, very high and high temperature ranges had a significant negative impact on richness and Shannon diversity for fungi and protists, as well as on the richness of nematodes and arthropods. While the fungal Shannon diversity was low in sites with very low

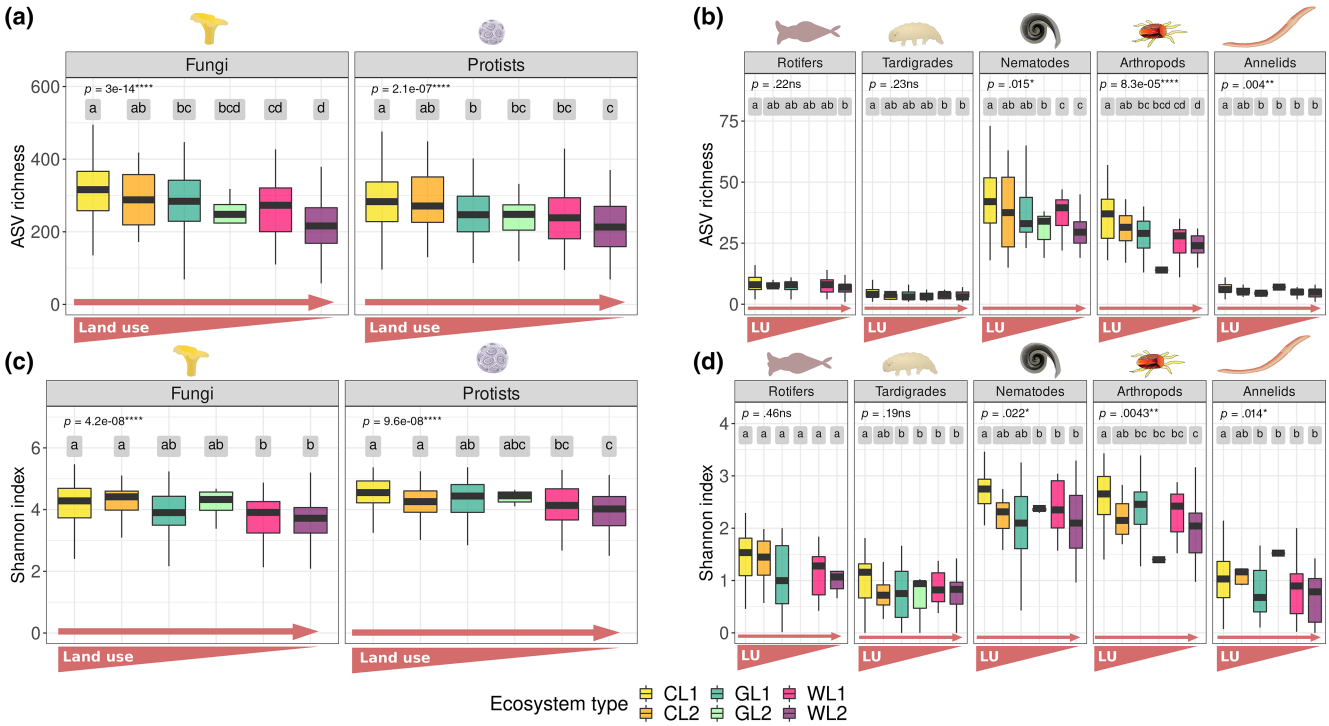


FIGURE 4 α-diversity of eukaryotic groups in different ecosystem types. Observed ASV richness for (a) fungi and protists and for (b) animals (rotifers, tardigrades, nematodes, arthropods, and annelids). Shannon index for (c) fungi and protists, and for (d) animals. The number of sampled sites was reduced after statistical filtering to 787 sites for fungi, 785 protists, 315 rotifers, 589 tardigrades, 587 nematodes, 583 arthropods, and 666 annelids. While the α-diversity varied significantly for fungi, protists, nematodes, arthropods, and annelids between different ecosystem types (Kruskal–Wallis, $p < .025$), no significant differences were observed for rotifers and tardigrades ($p > .05$). CL, croplands; GL, grasslands; WL, woodlands. Land use (LU), as indicated by the red arrow, decreases from left to right.

precipitation regimes (average mean precipitation 433 mm), the opposite effect was observed in the case of richness of fungi (i.e., significantly higher values compared to wetter sites). This finding, together with the fact that a higher diversity was found in agricultural ecosystems, could be linked to more frequent irrigation practices in agricultural fields (Figure S7i).

The random forest analysis explained 22.5% of the total variation in fungal ASV richness, followed by protists (15.4%), rotifers (2.7%), tardigrades (0.7%), nematodes (12.2%), arthropods (12.7%), and annelids (5.4%; Figure S8a), with different factors having a different influence on the observed variation in ASV richness of different eukaryotic groups. Accordingly, soil pH was the most important driver of the richness of fungi, rotifers, and annelids (Figure 5). For annelids, elevation and the respiration quotient also had an important influence on their ASV richness. The basal respiration followed by the phosphorus content were the most important factors for tardigrades. The Corg:N ratio also played an important role for fungi, protists, nematodes, and particularly for arthropods. For the latter group, it was the only variable besides land use intensity impacting their richness. Among the climatic variables, the annual temperature ranges between 1970 and 2000 played the most important role in determining the observed variation of nematodes, tardigrades, rotifers, and protists. Nevertheless, we found climatic variables to be less important in explaining the eukaryotic ASV richness compared to soil properties. For example, arthropod richness was not influenced by any climatic variable, whereas variables related to extreme heat and drought conditions in quarters or months and the mean diurnal range between 1970 and 2000 significantly impacted fungal richness. Considering the ecosystem type, the eukaryotic richness was better predicted by the land use gradient occurring over 2009, 2015, and 2018 than by only considering the land use of the sampled year (2018).

Environmental variables explained a lower share of the total variation in Shannon diversity: 0.6% (tardigrades), 2.7% (nematodes), 3.8% (annelids), 4.9% (arthropods), 6.4% (fungi), and 11.3% (protists); see (Figure S8b). The best predictor of the fungal and protistan Shannon diversity was soil pH. For arthropods and tardigrades, the Corg:N ratio was the most important driver of the Shannon diversity, while aridity was the best variable predicting nematode diversity and plant-available phosphorus that of annelids. For rotifers, we did not find a significant impact of any environmental variable on their diversity (Figure S9).

3.3 | β-diversity of eukaryotic soil communities driven by ecosystem type

We conducted a distance-based redundancy analysis (dbRDA) to compare eukaryotic diversity between different sites (β-diversity).

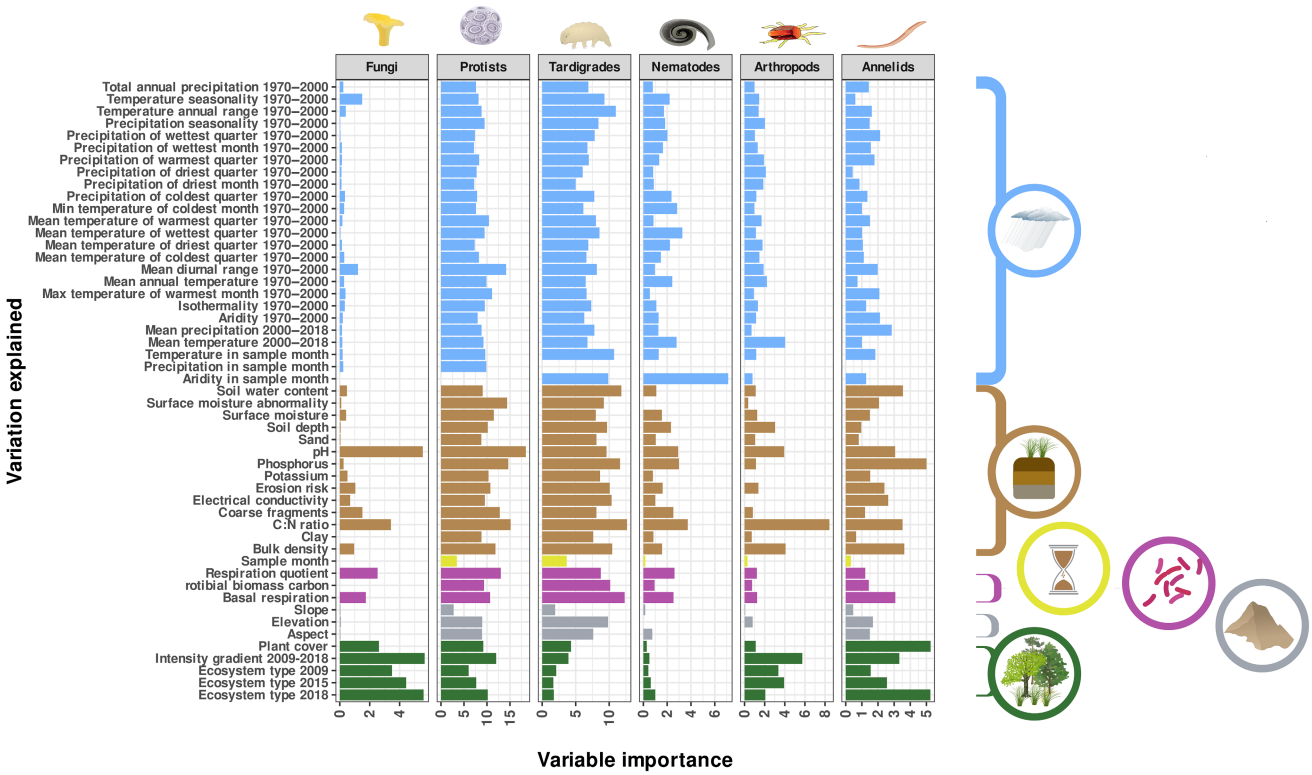


FIGURE 5 Variable importance explaining observed richness for different eukaryotic groups. Variable importance (backward selection) for the observed ASV richness of eukaryotes for fungi, protists, rotifers, tardigrades, nematodes, arthropods, and annelids (from left to right). For variable explanation, see Table S2. Colored bars indicate different groups of environmental variables: climatic properties (blue), soil properties (brown), sampling time (yellow), biological properties (purple), topography (gray), and ecosystem type and plant cover (green). Results on Shannon index are shown in Figure S9.

The results confirmed that all eukaryotic communities investigated were clustered by ecosystem type (with the highest AIC, Figure 6, ANOVA $p < 0.05$). For rotifers and tardigrades, this was less obvious, and their diversity overlapped over different ecosystems (Figure 6c,d).

The analysis also indicated that climatic variables (namely averaged temperatures for the period 1970–2000 rather than the temperature value of the sampled month), soil pH, conductivity, and soil depth better explained the eukaryotic diversity in annual croplands. In contrast, in permanent croplands, precipitation seasonality and soil carbonate content were the most important factors driving the community composition. This could be explained by the fact that the majority of the permanent croplands were olive groves and vineyards cultivated in Mediterranean carbonate-rich soils (i.e., Calcisols). In contrast, high phosphorus content was the main driver of eukaryotic community structure in managed grasslands and higher values of the Corg:N ratio and surface soil moisture of those in woodlands. Surface soil moisture was significantly correlated with eukaryotic β-diversity in broadleaved woodlands, whereas in coniferous woodlands the best correlations were observed with the Corg:N ratio and microbial biomass carbon (Cmic; Figure 6a).

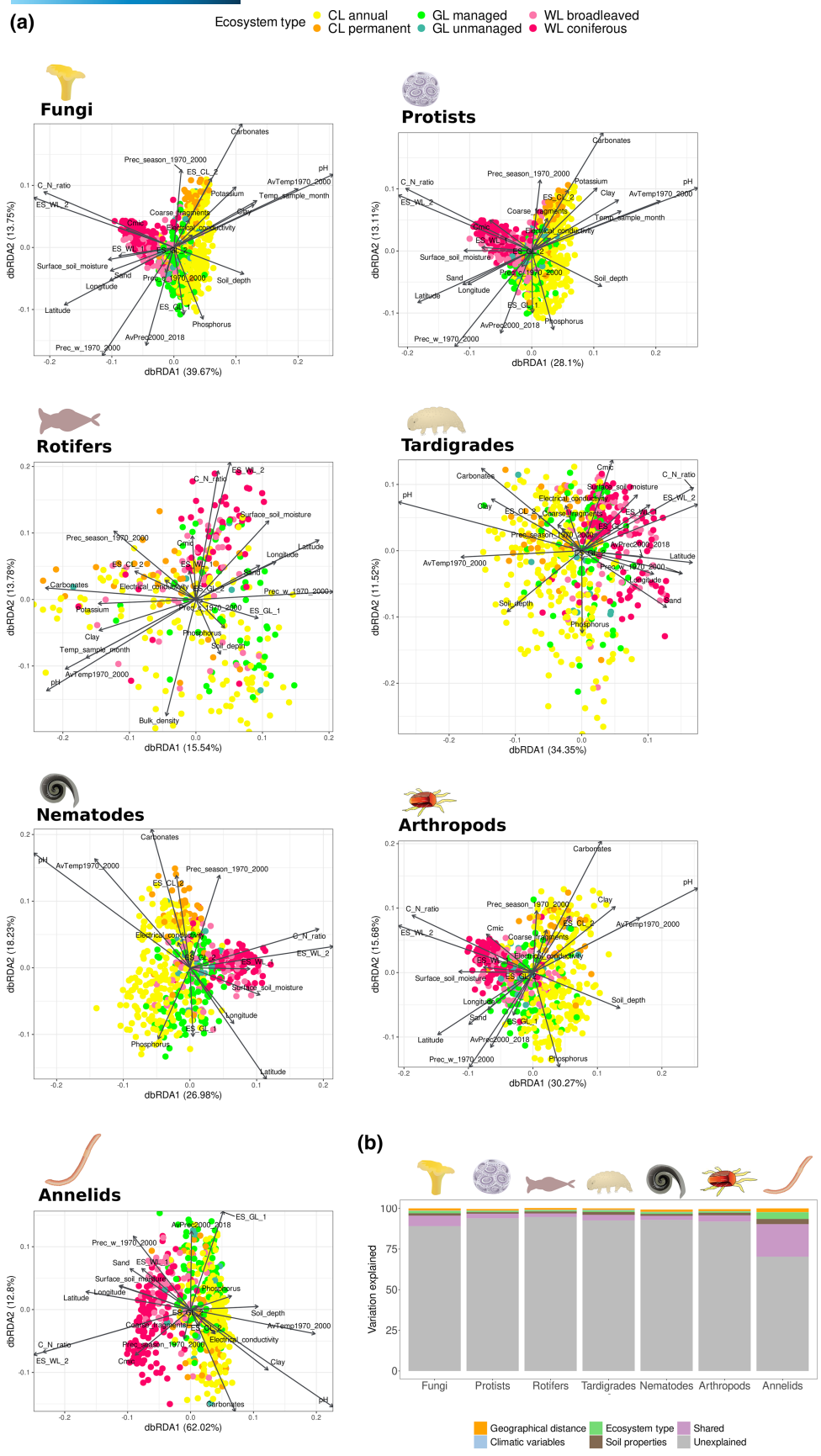
The PERMANOVA test also confirmed that the β-diversity of the studied eukaryotic groups was significantly affected by ecosystem properties (*adonis* <0.001, Table S7). We found that plant cover, pH, soil depth, erosion risk, mean annual temperature, precipitation, and annual temperature ranges (all three climatic variables between 1970 and 2000) significantly shaped eukaryotic community structure. However, we found less dispersion in β-diversity between sites located in alkaline soils, in deep soils, in sites exposed to high to very high erosion and in those with very low precipitation values (Figure S10; average mean precipitation 433 mm). Also, for all eukaryotic groups, croplands were found to be more homogenously dispersed compared to woodlands with the only exception being the dispersion of annelids which was found more homogenous in coniferous forests compared to permanent cropland (Figure S10). Particularly, homogeneous were the clusters formed by tardigrades, nematodes, and arthropods in relation to temperature and temperature ranges (for homogeneity of multivariate dispersion see Table S5, *betadisper* $p > .05$, ANOVA for temperature for tardigrades, temperature range for nematodes and both variables for arthropods; Table S6).

Ecosystem characteristics explained 3% of the total β-diversity for rotifers, 6.1% for protists, 7.4% for tardigrades, 7.1% for nematodes, 8.1% for arthropods, 10.1% for fungi, and 29.7% for annelids (Figure 6b; for a more detailed overview, see Venn diagram in Figure S11). However, past and present climate played a smaller role in explaining the β-diversity of soil eukaryotes (0%–0.2%). In the case of fungi, soil properties played the most important role (1.5%), followed by the spatial distance between sites (1.4%) and ecosystem type (1.3%). Similarly, soil properties and geographical distance were also the most important factors explaining the variation in the β-diversity of protists (1%), followed by ecosystem type (0.7%). The geographical distance was the most influential factor for nematodes (1.2%), and soil properties for rotifers and tardigrades. In the case of arthropods, soil properties explained 1.6%, ecosystem type 0.8% and the geographical distance between sites 1.2%. The β-diversity of annelids was mainly explained by the ecosystem type (4.1%), followed by soil properties (3.2%) and spatial distances (2.3%).

4 | DISCUSSION

Our study provides the first European baseline of eukaryotic soil biodiversity, and the higher number of environmental variables considered compared to previous studies enhances our current understanding of the belowground eukaryotic diversity at large spatial scales. Although our results confirm the importance of ecosystem type for α- (Bahram et al., 2018; George et al., 2019) and β-diversity (Ritter et al., 2019; Shi et al., 2015), we also provide new insights into how local environmental factors shape eukaryotic genetic diversity patterns. We also highlight limitations of currently applied methodologies and how they might be overcome in future studies. Importantly, we observed less eukaryotic species richness and diversity in woodlands, in shallow acidic soils with high erosion risk, and in areas with low annual mean temperature and high annual temperature variations. Drivers of α-diversity differed between eukaryotic groups in agreement with Aslani et al. (2022), and our results confirmed the strong influence of soil properties on eukaryotic community composition, with the Corg:N ratio and soil bulk density being the main drivers for protistan (George et al., 2019) and fungal diversity (Bahram et al., 2018), followed by soil pH for fungi diversity (George et al., 2019; Wutkowska et al., 2019). However, unlike previous findings (George et al., 2019), soil properties were also the most important driver for studied eukaryotic groups: Corg:N ratio was driving the richness of nematodes and arthropods, soil pH the richness of rotifers and annelids, and plant-available phosphorus the richness of tardigrades and protists.

Furthermore, our results also revealed that agricultural lands sustain a higher α-diversity of protists and fungi than woodlands (George et al., 2019). This finding is in line with another molecular study conducted in Indonesia (Schulz et al., 2019) that also found that protist diversity increases with land use intensification. Compared to woodlands, (croplands also contained significantly higher diversity and richness of tardigrades, nematodes, arthropods, and annelids) and higher richness of rotifers. Although this finding contrasts with previous observations of consistent declines in soil biodiversity with increasing land use intensity (Schulz et al., 2019; Tsiafouli et al., 2015), it is in line with recent analyses of microbial communities from the same LUCAS soil samples (16S bacteria and ITS fungi; Labouyrie et al., 2023) indicating that ecosystem type is the main driver shaping the whole soil food web. β-diversity was also strongly shaped by ecosystem type, in agreement with previous studies that compared grasslands and forests in the Amazonas (Ritter et al., 2019) and those investigating the effects of vegetation on bacterial and microeukaryotic communities in the arctic tundra (Shi et al., 2015). However, these are only a few examples, with



most large-scale studies not including ecosystem type in their analyses and crucially, omitting croplands. In addition, previous works did not consider the complexity of soil fauna communities, which obscure the responses of individual groups (e.g., ecosystem type was not a dominant environmental driver of tardigrade and rotifer communities).

FIGURE 6 β-diversity of different eukaryotic groups explained by ecosystem properties. (a) dbRDA analysis showing the influence of ecosystem type and environmental variables on eukaryotic community structure. (b) Variation partitioning showing the explained variance of unique and shared effects of ecosystem properties on β-diversity at the sampled sites (787 sites for fungi, 785 protists, 315 rotifers, 589 tardigrades, 587 nematodes, 583 arthropods, and 666 annelids). AvPrec_1970_2000, average temperature 1970–2000; AvTemp_1970_2000, average temperature 1970–2000; ES_CL1 = annual croplands; ES_CL2, permanent croplands; ES_GL1, managed grasslands; ES_GL2, unmanaged GL; ES_WL1, broadleaved woodlands; ES_WL2, coniferous woodlands; Prec_c_1970_2000, precipitation of coldest quarter 1970–2000; Prec_season_1970_2000, precipitation seasonality 1970–2000; Prec_w_1970_2000, precipitation of warmest quarter 1970–2000.

Despite the fact that two recent global studies (one on soil eukaryotic community assemblages (Aslani et al., 2022) and another on protists composition (Oliverio et al., 2020)) found that soil eukaryotic community structures are mainly explained by annual precipitation, in our study, precipitation, temperature, and temperature ranges were the least informative variables of the spatial variation in β-diversity. Therefore, it is likely that the effects of climate on soil eukaryotic diversity might have been overestimated in global studies (George et al., 2019) as a consequence of being based on a single time and not including a fair number of soil properties. Indeed, only pH and soil carbon were considered in the study by Oliverio et al. (2020) and pH, soil carbon, and moisture in that by Aslani et al. (2022). However, we found that the Corg:N ratio, plant-available phosphorus, soil water content, surface moisture abnormality, and long-term climatic variables played a crucial role and need to be considered in future surveys and monitoring programs. This highlights the need for more comprehensive assessments when analyzing soil biodiversity patterns at large spatial scales by including a better physical and chemical characterization of the investigated soils. However, also the heterogeneity of soil organisms might impact results which were found both on global as on European scale. In order to account for the effect of the heterogeneity of soil organisms, more large-scale studies on less heterogeneous soils are needed.

Another potential explanation for the observed higher α-diversity of fungi, protists, rotifers, tardigrades, nematodes, arthropods, and annelids with increased land use intensity could be related to soil management. It is known that some sustainable agricultural practices (e.g., the use of organic fertilizers) promote soil biodiversity (Lentendu et al., 2014), particularly in the long term (Ros et al., 2006). Unfortunately, large-scale monitoring surveys (incl. LUCAS) often do not gather soil management information, and therefore, a direct relationship between those practices and higher α-diversity could not be confirmed. Future sampling campaigns would benefit from including more information on land management, such as tillage, pesticide application, and fertilization regimes.

Surprisingly aboveground diversity (i.e., plant cover) and soil properties only explained 0.7%–22.5% of the α-diversity and 3%–29.7% of the β-diversity was unexplained (values). This confirms that the complexity of the soil habitat is still very difficult to quantify and is in agreement with previous studies (Aslani et al., 2022; Bahram et al., 2018; George et al., 2019; Oliverio et al., 2020), highlighting the need for future assessments of soil biota variability at microscale (microsites) within larger studies (Thakur et al., 2020).

Nevertheless, our study showed that annelids had the highest share of explaining variables for β-diversity variance, which could be due to the low number of taxonomical entities being identified (49 ASVs) and hence, reflecting either a low variation throughout the EU or a limited portion of the actual diversity was captured. This might be related to the limited amount of soil volume considered in this study. To better account for nematodes and annelids at least 10 g of soil should be sequenced (Kageyama & Toju, 2022). However, volumes also depend on the kit used (Lilja et al., 2023). Indeed, morphological identification, coupled with DNA barcoding of larger soil volumes, would be the best way to assess soil fauna diversity in the future.

Interestingly, we found that the three main ecosystem types (cropland, grassland, and woodland) shared a high number of taxa. This could be linked to the presence of DNA/organisms inherited from former land uses. Across the EU, many croplands have been converted from former grasslands (Schils et al., 2022) or woodlands (Kaplan et al., 2009), and therefore, the higher soil biodiversity observed in croplands could be a legacy effect from relic DNA of dormant or dead organisms (necromass). Previous studies have found that manure (Ye et al., 2019) and straw (Xia et al., 2021) can increase the amounts of soil necromass and that relic DNA is an important component of overall soil biodiversity. For example, in the case of fungi, between 31% and 96% (Carini et al., 2016), while for other eukaryotes, the total amount remains unknown.

This might be related to our results showing that changes in land use (i.e., the ecosystem present in the previous years) and in management practices (i.e., land use intensity), rather than the ecosystem type present at the time of the sampling, better explained the α-diversity of eukaryotes in European soils. Furthermore, long-term temperature data (from the period of 1970–2000) was found to be more relevant in determining eukaryotic community composition compared, for instance, to the temperature of the sampled month. The relevance of long- versus short-term environmental effects could also be behind the observed little impact of the sampling season on eukaryotic genetic diversity (only a significant effect was observed for α-diversity in the case of protists).

Taken together, our findings call for a more holistic consideration on long-term information of land use and climatic factors when interpreting soil biodiversity data. Furthermore, our results clearly demonstrate the key role of ecosystem type on the diversity and structure of soil eukaryotic assemblages. It should be noted that in this study, we targeted the V4 region which is prone to amplification biases, for

example, neglecting certain organism groups such as the protistan Amoebozoa and Heterolobosea (Discoba; Vaulot et al., 2022) or grouping rare species with commonly abundant species (Lara et al., 2022). While the V4 of the 18S gene allows a general overview of soil eukaryotes and provides more phylogenic information compared to other primers, the limitations of the study need to be considered when trying to identify indicators for planning conservation activities.

Finally, it is also noteworthy to mention that the large variation in size of eukaryotes (from micro- to macrofauna) requires the adoption of different soil sampling strategies for more accurate estimates of their genetic diversity (Wiesel et al., 2015) as well as a higher taxonomic resolution, for example, by using more variable genetic markers such as the ITS for fungi. In order to prioritize soil biodiversity hotspots, more information is also needed on which species are active, dormant, or dead. For example, the latter (relic DNA) could be accounted for when assessing soil (eukaryotic) biodiversity by extracting extracellular DNA (washing the DNA after the soil sampling; Nagler et al., 2022). Also, to better interpret results, more information on the function of organisms is needed and whether the preservation/enhancement of organisms could be beneficial or detrimental (e.g., pathogenic and harmful) to other organisms and to the overall functioning of the soil. Together with more data on management, this information might contribute to defining soil biodiversity indicators, which are particularly important in light of future legislative initiatives (e.g., EU Soil Health Law (European Commission, 2023)).

ACKNOWLEDGMENTS

The LUCAS Survey is coordinated by Unit E4 of the Statistical Office of the European Union (EUROSTAT). The LUCAS Soil sample collection is supported by the Directorate-General Environment (DG-ENV), Directorate-General Agriculture and Rural Development (DG-AGRI), Directorate-General Climate Action (DG-CLIMA), and Directorate-General Eurostat (DG-ESTAT) of the European Commission. This work was realized in the framework of the Collaborative Doctoral Partnership JRC and UVIGO under the Agreement No. 35533. We thank Leonidas Liakos and Francis Matthews for advice concerning available databases and Elise van Eynde and Maëva Labouyrie for their support with the statistical analysis. Authors also acknowledge funding for open access charge from Universidade de Vigo/CISUG.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study is publically available, i.e., all the protocols, metadata and raw data are accessible on the European Soil Data Centre (https://esdac.jrc.ec.europa.eu/content/soil-biodiversity-dna-eukaryotes). Additionally, the raw data (DNA sequences) generated in this study have been deposited in the Sequence Read Archive (SRA) database under BioProject ID PRJNA985135.

ORCID

Julia Köninger https://orcid.org/0000-0002-0413-7931
Cristiano Ballabio https://orcid.org/0000-0001-7452-9271
Panos Panagos https://orcid.org/0000-0003-1484-2738
Arwyn Jones https://orcid.org/0000-0003-2781-8905
Marc W. Schmid https://orcid.org/0000-0001-9554-5318
Alberto Orgiazzi https://orcid.org/0000-0002-8390-6654
Maria J. I. Briones https://orcid.org/0000-0002-4051-8117

REFERENCES

Aslani, F., Geisen, S., Ning, D., Tedersoo, L., & Bahram, M. (2022). Towards revealing the global diversity and community assembly of soil eukaryotes. *Ecology Letters*, *25*(1), 65–76. https://doi.org/10.1111/ele.13904

Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M., Bengtsson-Palme, J., Anslan, S., Coelho, L. P., & Harend, H. (2018). Structure and function of the global topsoil microbiome. *Nature*, *560*(7717), 233–237. https://doi.org/10.1038/s41586-018-0386-6

Bahrndorff, S., de Jonge, N., Hansen, J. K., Lauritzen, J. M. S., Spanggaard, L. H., Sørensen, M. H., Yde, M., & Nielsen, J. L. (2018). Diversity and metabolic potential of the microbiota associated with a soil arthropod. *Scientific Reports*, *8*(1), 1–8. https://doi.org/10.1038/s41598-018-20967-0

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. *The ISME Journal*, *7*(3), 652–659. https://doi.org/10.1038/ismej.2012.147

Beule, L., & Karlovsky, P. (2020). Improved normalization of species count data in ecology by scaling with ranked subsampling (SRS): Application to microbial communities. *PeerJ*, *8*, e9593. https://doi.org/10.7717/peerj.9593

Brandt, M. I., Trouche, B., Quintric, L., Günther, B., Wincker, P., Poulain, J., & Arnaud-Haond, S. (2021). Bioinformatic pipelines combining denoising and clustering tools allow for more comprehensive prokaryotic and eukaryotic metabarcoding. *Molecular Ecology Resources*, *21*(6), 1904–1921. https://doi.org/10.1111/1755-0998.13398

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Carini, P., Marsden, P. J., Leff, J. W., Morgan, E. E., Strickland, M. S., & Fierer, N. (2016). Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nature Microbiology*, *2*(3), 1–6. https://doi.org/10.1038/nmicrobiol.2016.242

Devigne, C., Mouchon, P., & Vanhee, B. (2016). Impact of soil compaction on soil biodiversity–Does it matter in urban context? *Urban Ecosystem*, *19*, 1163–1178. https://doi.org/10.1007/s11252-016-0547-z

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fournier, B., Samaritani, E., Frey, B., Seppey, C. V., Lara, E., Heger, T. J., & Mitchell, E. A. (2020). Higher spatial than seasonal variation in floodplain soil eukaryotic microbial communities. *Soil Biology and Biochemistry*, *147*, 107842. https://doi.org/10.1016/j.soilbio.2020.107842

George, P. B., Lallias, D., Creer, S., Seaton, F. M., Kenny, J. G., Eccles, R. M., Griffiths, R. I., Lebron, I., Emmett, B. A., & Robinson, D. A. (2019). Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-09031-1

Jurburg, S. D., Keil, P., Singh, B. K., & Chase, J. M. (2021). All together now: Limitations and recommendations for the simultaneous analysis of all eukaryotic soil sequences. *Molecular Ecology Resources*, *21*(6), 1759–1771. https://doi.org/10.1111/1755-0998.13401

Kageyama, T., & Toju, H. (2022). Effects of source sample amount on biodiversity surveys of bacteria, fungi, and nematodes in soil ecosystems. *bioRxiv*. https://doi.org/10.1101/2022.05.26.493541

Kaplan, J. O., Krumhardt, K. M., & Zimmermann, N. (2009). The prehistoric and preindustrial deforestation of Europe. *Quaternary Science Reviews*, *28*(27–28), 3016–3034. https://doi.org/10.1016/j.quascirev.2009.09.028

Labouyrie, M., Ballabio, C., Romero, F., Panagos, P., Jones, A., Schmid, M. W., Mikryukov, V., Dulya, O., Tedersoo, L., Bahram, M., Lugato, E., van der Heijden, M. G.-A., & Orgiazzi, A. (2023). Patterns in soil microbial diversity across Europe. *Nature Communications*, *14*, 3311. https://doi.org/10.1038/s41467-023-37937-4

Lara, E., Singer, D., & Geisen, S. J. E. M. (2022). Discrepancies between prokaryotes and eukaryotes need to be considered in soil DNA-based studies. *Environmental Microbiology*, *24*, 3829–3839. https://doi.org/10.1111/1462-2920.16019

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

Lilja, M. A., Buivydaitė, Ž., Zervas, A., Krogh, P. H., Hansen, B. W., Winding, A., & Sapkota, R. (2023). Comparing earthworm biodiversity estimated by DNA metabarcoding and morphology-based approaches. *Applied Soil Ecology*, *185*, 104798. https://doi.org/10.1016/j.apsoil.2022.104798

Luan, L., Jiang, Y., Cheng, M., Dini-Andreote, F., Sui, Y., Xu, Q., Geisen, S., & Sun, B. (2020). Organism body size structures the soil microbial and nematode community assembly at a continental and global scale. *Nature Communications*, *11*(1), 6406. https://doi.org/10.1038/s41467-020-20271-4

Magoč, T., & Salzberg, S. L. J. B. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, *27*(21), 2957–2963. https://doi.org/10.1093/bioinformatics/btr507

Malard, L. A., Mod, H. K., Guex, N., Broennimann, O., Yashiro, E., Lara, E., Mitchell, E. A., Niculita-Hirzel, H., & Guisan, A. (2022). Comparative analysis of diversity and environmental niches of soil bacterial, archaeal, fungal and protist communities reveal niche divergences along environmental gradients in the Alps. *Soil Biology and Biochemistry*, *169*, 108674. https://doi.org/10.1016/j.soilbio.2022.108674

Minasny, B., & McBratney, A. B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences*, *32*(9), 1378–1388. https://doi.org/10.1016/j.cageo.2005.12.009

Nagler, M., Podmirseg, S. M., Ascher-Jenull, J., Sint, D., & Traugott, M. (2022). Why eDNA fractions need consideration in biomonitoring. *Molecular Ecology Resources*, *22*(7), 2458–2470. https://doi.org/10.1111/1755-0998.13658

NASA-USDA. (2015). *Global soil moisture data*. https://earth.gsfc.nasa.gov/hydro/data/nasa-usda-global-soil-moisture-data

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R., Simpson, G., & Solymos, P. (2020). Vegan: Community ecology package. R Package Version 2.5-6. 2019.

Oliverio, A. M., Geisen, S., Delgado-Baquerizo, M., Maestre, F. T., Turner, B. L., & Fierer, N. (2020). The global-scale distributions of soil protists and their contributions to belowground systems. *Science Advances*, *6*(4), eaax8787. https://doi.org/10.1126/sciadv.aax8787

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., & Fernández-Ugalde, O. (2018). LUCAS soil, the largest expandable soil dataset for Europe: A review. *European Journal of Soil Sciences*, *69*(1), 140–153. https://doi.org/10.1111/ejss.12499

Orgiazzi, A., Dunbar, M. B., Panagos, P., de Groot, G. A., & Lemanceau, P. (2015). Soil biodiversity and DNA barcodes: Opportunities and challenges. *Soil Biology and Biochemistry*, *80*, 244–250. https://doi.org/10.1016/j.soilbio.2014.10.014

Orgiazzi, A., Panagos, P., Fernández-Ugalde, O., Wojda, P., Labouyrie, M., Ballabio, C., Franco, A., Pistocchi, A., Montanarella, L., & Jones, A. (2022). LUCAS soil biodiversity and LUCAS soil pesticides, new tools for research and policy development. *European Journal of Soil Science*, *72*(5), 1–14. https://doi.org/10.1111/ejss.13299

Ozkurt, E., Fritscher, J., Soranzo, N., Ng, D. Y., Davey, R. P., Bahram, M., & Hildebrand, F. (2021). LotuS2: An ultrafast and highly accurate tool for amplicon sequencing analysis. *bioRxiv*. https://doi.org/10.1101/2021.12.24.474111

Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusburger, K., Montanarella, L., & Alewell, C. (2015). The new assessment of soil loss by water erosion in Europe. *Environmental Science & Policy*, *54*, 438–447. https://doi.org/10.1016/j.envsci.2016.02.010

Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *Soil*, *7*(1), 217–240. https://doi.org/10.5194/soil-7-217-2021

Ritter, C. D., Zizka, A., Barnes, C., Nilsson, R. H., Roger, F., & Antonelli, A. (2019). Locality or habitat? Exploring predictors of biodiversity in Amazonia. *Ecography*, *42*(2), 321–333. https://doi.org/10.1111/ecog.03833

Ritter, C. D., Zizka, A., Roger, F., Tuomisto, H., Barnes, C., Nilsson, R. H., & Antonelli, A. (2018). High-throughput metabarcoding reveals the effect of physicochemical soil properties on soil and litter biodiversity and community turnover across Amazonia. *PeerJ*, *6*, e5661. https://doi.org/10.7717/peerj.5661

Ros, M., Klammer, S., Knapp, B., Aichberger, K., & Insam, H. (2006). Long-term effects of compost amendment of soil on functional and structural diversity and microbial activity. *Soil Use and Management*, *22*(2), 209–218. https://doi.org/10.1111/j.1475-2743.2006.00027.x

Schils, R. L., Bufe, C., Rhymer, C. M., Francksen, R. M., Klaus, V., Abdalla, M., Milazzo, F., Lellei-Kovács, E., Ten Berge, H., & Bertora, C. (2022). Delivery of ecosystem services from permanent grasslands in Europe: A systematic review. *Grassland at the Heart of Circular Sustainable Food Systems*, *27*, 274–276.

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. *Frontiers in Microbiology*, *10*, 240. https://doi.org/10.3389/fmicb.2019.00240

Shi, Y., Xiang, X., Shen, C., Chu, H., Neufeld, J. D., Walker, V. K., & Grogan, P. (2015). Vegetation-associated impacts on arctic tundra bacterial and microeukaryotic communities. *Applied Environmental Microbiology*, *81*(2), 492–501. https://doi.org/10.1128/AEM.03229-14

Singer, D., Seppey, C. V., Lentendu, G., Dunthorn, M., Bass, D., Belbahri, L., Blandenier, Q., Debroas, D., de Groot, G. A., & De Vargas, C. (2021). Protist taxonomic and functional diversity in soil, freshwater and marine ecosystems. *Environment International*, *146*, 106262. https://doi.org/10.1016/j.envint.2020.106262

Smith, L. C., Orgiazzi, A., Eisenhauer, N., Cesarz, S., Lochner, A., Jones, A., Bastida, F., Patoine, G., Reitz, T., & Buscot, F. (2021). Large-scale drivers of relationships between soil microbial properties and organic carbon across Europe. *Global Ecology Biogeography*, *30*(10), 2070–2083. https://doi.org/10.1111/geb.13371

Soil Health Law. (2023). https://www.eesc.europa.eu/it/our-work/opinions-information-reports/opinions/soil-health-law

Tedersoo, L., Sánchez-Ramírez, S., Kõljalg, U., Bahram, M., Döring, M., Schigel, D., May, T., Ryberg, M., & Abarenkov, K. (2018). High-level classification of the fungi and a tool for evolutionary ecological

analyses. *Fungal Diversity*, *90*(1), 135–159. https://doi.org/10.1007/s13225-018-0401-0

Thakur, M. P., Phillips, H. R., Brose, U., De Vries, F. T., Lavelle, P., Loreau, M., Mathieu, J., Mulder, C., Van der Putten, W. H., & Rillig, M. C. (2020). Towards an integrative understanding of soil biodiversity. *Biological Reviews*, *95*(2), 350–364. https://doi.org/10.1111/brv.12567

Treonis, A. M., Unangst, S. K., Kepler, R. M., Buyer, J. S., Cavigelli, M. A., Mirsky, S. B., & Maul, J. E. (2018). Characterization of soil nematode communities in three cropping systems through morphological and DNA metabarcoding approaches. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-20366-5

Tsiafouli, M. A., Thébault, E., Sgardelis, S. P., De Ruiter, P. C., Van Der Putten, W. H., Birkhofer, K., Hemerik, L., De Vries, F. T., Bardgett, R. D., & Brady, M. V. (2015). Intensive agriculture reduces soil biodiversity across Europe. *Global Change Biology*, *21*(2), 973–985. https://doi.org/10.1111/gcb.12752

Vaulot, D., Sim, C. W. H., Ong, D., Teo, B., Biwer, C., Jamy, M., & Lopes dos Santos, A. (2022). metaPR(2): A database of eukaryotic 18S rRNA metabarcodes with an emphasis on protists. *Molecular Ecology Resources*, *22*(8), 3188–3201. https://doi.org/10.1111/1755-0998.13674

Wiesel, L., Daniell, T. J., King, D., & Neilson, R. (2015). Determination of the optimal soil sample size to accurately characterise nematode communities in soil. *Soil Biology and Biochemistry*, *80*, 89–91. https://doi.org/10.1016/j.soilbio.2014.09.026

Wu, T., Ayres, E., Bardgett, R. D., Wall, D. H., & Garey, J. R. (2011). Molecular study of worldwide distribution and diversity of soil animals. *Proceedings of the National Academy of Sciences*, *108*(43), 17720–17725. https://doi.org/10.1073/pnas.1103824108

Wutkowska, M., Vader, A., Mundra, S., Cooper, E. J., & Eidesen, P. B. (2019). Dead or alive; or does it really matter? Level of congruency between trophic modes in total and active fungal communities in high arctic soil. *Frontiers in Microbiology*, *9*, 3243. https://doi.org/10.3389/fmicb.2018.03243

Xia, Y., Chen, X., Zheng, S., Gunina, A., Ning, Z., Hu, Y., Tang, H., Rui, Y., Zhang, Z., He, H., & Huang, D. (2021). Manure application accumulates more nitrogen in paddy soils than rice straw but less from fungal necromass. *Agriculture, Ecosystems & Environment*, *319*, 107575. https://doi.org/10.1016/j.agee.2021.107575

Xiong, W., Jousset, A., Li, R., Delgado-Baquerizo, M., Bahram, M., Logares, R., Wilden, B., de Groot, G. A., Amacker, N., & Kowalchuk, G. A. (2021). A global overview of the trophic structure within microbiomes across ecosystems. *Environment International*, *151*, 106438. https://doi.org/10.1016/j.envint.2021.106438

Ye, G., Lin, Y., Kuzyakov, Y., Liu, D., Luo, J., Lindsey, S., Wang, W., Fan, J., & Ding, W. (2019). Manure over crop residues increases soil organic matter but decreases microbial necromass relative contribution in upland ultisols: Results of a 27-year field experiment. *Soil Biology and Biochemistry*, *134*, 15–24. https://doi.org/10.1016/j.soilbio.2019.03.018

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., & Glöckner, F. O. (2014). The SILVA and "all-species living tree project (LTP)" taxonomic frameworks. *Nucleic Acids Research*, *42*(D1), D643–D648. https://doi.org/10.1093/nar/gkt1209

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Köninger, J., Ballabio, C., Panagos, P., Jones, A., Schmid, M. W., Orgiazzi, A., & Briones, M. J. I. (2023). Ecosystem type drives soil eukaryotic diversity and composition in Europe. *Global Change Biology*, *00*, 1–14. https://doi.org/10.1111/gcb.16871