# Primer and database choice affect fungal functional but not biological diversity findings in a national soil survey

George, Paul; Creer, Simon; Griffiths, Robert I.; Emmett, Bridgett A.; Robinson, David A.; Jones, Davey L.

**Frontiers in Environmental Science**

*DOI:*
10.3389/fenvs.2019.00173

Published: 01/11/2019

Peer reviewed version

# Primer and database choice affect fungal functional but not biological diversity findings in a national soil survey

Paul B.L. George[*,1,2], Simon Creer[1], Robert I. Griffiths[2], Bridget A. Emmett[2], David A. Robinson[2], Davey L. Jones[1,3]

[1] *School of Natural Sciences, Bangor University, Deiniol Road, Bangor, Gwynedd, LL57 2UW, United Kingdom*

[2] *Centre for Ecology & Hydrology, Environment Centre Wales, Deiniol Road, Bangor, Gwynedd, LL57 2UW, United Kingdom*

[3] *School of Agriculture and Environment, University of Western Australia, Crawley, Western Australia, WA6009, Australia*

*Corresponding author. School of Natural Sciences, Bangor University, Deiniol Road, Bangor, Gwynedd, LL57 2UW, United Kingdom. *Email*: afp67e@bangor.ac.uk; pblgeorge@gmail.com

**British English formatting**

**KEYWORDS:** UNITE; SILVA; identification bias; high-throughput sequencing; Arbuscular mycorrhizal fungi; Archaeorhizomycetes

**ABSTRACT**

The internal transcribed spacer (ITS) region is the accepted DNA barcode of fungi. Its use has led to a step-change in the assessment and characterisation of fungal communities from environmental samples by precluding the need to isolate, culture, and identify individuals. However, certain functionally important groups, such as the arbuscular mycorrhizas (Glomeromycetes), are better characterised by alternative markers such as the 18S rRNA region. Previous use of an ITS primer set in a nationwide metabarcoding soil biodiversity survey revealed that fungal richness declined along a gradient of productivity and management intensity. Here, we wanted to discern whether this trend was also present in data generated from universal 18S primers. Furthermore, we wanted to extend this comparison to include measures of functional diversity and establish trends with soil types and soil organic matter (SOM) content. Over the 413 individual sites examined (arable, grassland, woodland, moorland, heathland), we found congruent trends of total fungal richness and β-diversity across land uses, SOM class and soil type with both ITS and 18S primer sets. A total of 24 fungal classes were shared between datasets, in addition to 15 unique to ITS1 and 12 unique to 18S. However, using FUNGUILD, divergent trends of functional group richness became apparent, especially for symbiotrophic fungi, likely driven by an increased detection rate of Glomeromycetes in the 18S dataset. The disparate trends were also apparent when richness and β-diversity were compared to soil properties. Additionally, we found SOM class to be a more meaningful variable than soil type biodiversity for predicting biodiversity analyses because organic matter was calculated for each sample whereas soil type was assigned from a national soil map. We advocate that a combination of fungal primers should be

used in large-scale soil biodiversity surveys to capture important groups that can be underrepresented by universal barcodes. Utilising such an approach can prevent the oversight of ubiquitous but poorly described species as well as critically important functional groups.

## INTRODUCTION

Soil fungi are the dominant eukaryotic component of soil communities and are known to perform crucial ecosystem functions (Peay et al., 2008). Characterising the diversity of fungi within the landscape and their response to anthropogenic perturbation therefore represents an important topic within ecology. High-throughput sequencing has allowed the rapid estimation and identification of fungi by overcoming historical limitations of culture isolation and classifying fruiting bodies (Tedersoo et al., 2015). Using these DNA-based approaches it has been estimated that global fungal diversity in soil ranges from 3.5 – 5 million species. Yet at the beginning of the present decade, only around one-tenth of fungal diversity was thought to have been described (Rosling et al., 2011). In terms of ecosystem function, the majority of fungi are important in organic matter turnover and nutrient recycling as they facilitate the conversion of complex organic polymers into forms more readily accessible to other organisms (Peay et al., 2008; Nguyen et al., 2016). Consequently, they play a crucial role in regulating both below- and above-ground productivity (Peay et al., 2008). Many soil fungi also form important interactions with plants. Some form mutualistic relationships, best exemplified by the wide range of mycorrhizas (Wang and Qui, 2006; Smith and Read, 2008; Nguyen et al., 2016), whereas others are pathogens, responsible for numerous plant and animal

diseases within agriculture and forestry (Fisher et al., 2012; Nguyen et al., 2016). Depending on environmental conditions or life stage, fungi are capable of taking on some or all of these roles (i.e. saprotroph, symbiotroph, pathotroph) (Fisher et al., 2012). Despite the recognition that fungi are extremely important in soil ecosystems, characterising fungal communities has remained a challenge, exemplified by the numerous studies on soil bacteria in comparison to fungi.

Fungal barcode sequences are found within the ubiquitous, multicopy ribosomal RNA gene. Within this, the internal transcribed spacer (ITS) region has been accepted as a universal barcode for fungi (Schoch et al., 2012). Recent development of ITS-based databases such as UNITE (Kõjlalg et al., 2013) and Warcup (Deshpande et al., 2016) have overcome limitations in collecting and assigning taxonomic identities to unknown sequences, though database selection may introduce bias into results (Tedersoo et al., 2015; Xue et al., 2019). Yet ITS barcodes exhibit some limitations when dealing with unknown or environmental samples. Generally, the ITS region cannot be aligned above the family-level (Cavender-Bares et al., 2009), making phylogenies based on ITS sequence data unreliable. Importantly, the ITS region has proven unreliable at distinguishing certain fungal groups at the species-level, such as Glomeromycetes (Stockinger et al., 2010). Such inconsistencies mean that ITS primers may not accurately detect target organisms. For instance, Berruti et al. (2017), found that ITS primers underestimated Glomeromycetes in bulk soil. Such uncertainty may confound experimental results and lead to erroneous conclusions.

Despite the widespread use of ITS barcodes, other markers may better capture the diversity of some fungal taxa. Primers targeting the small and large subunits as well as

93 the ITS regions of the rRNA gene have all been applied to fungi (Tedersoo et al., 2015;
94 Xue et al., 2019). For example, early diverging lineages such as Chytridiomycota
95 (Schoch et al., 2012; Tedersoo et al., 2015) and Glomeromycetes (Tedersoo, et al., 2015)
96 are poorly represented in ITS sequencing. Additionally, advancements in classification
97 have highlighted the shortcomings of environmental DNA barcoding. For example, the
98 Archaeorhizomycetes are a poorly understood but ubiquitous class of soil fungi and their
99 previously unidentifiable sequences have been major components of past soil biodiversity
100 assessments (Anderson et al., 2003; Rosling, et al., 2011). Overlooking these lineages
101 may potentially lead to erroneous assumptions of biological and functional diversity in
102 soils.

103 Underrepresentation of Glomeromycetes in particular exemplifies this issue.
104 Arbuscular mycorrhizal fungi (AMF) form symbiotic relationships with more than 80%
105 of vascular plant families and have been categorised into the monophyletic
106 Glomeromycetes (Schüβler et al., 2001). Unlike most fungi, the ITS region has
107 consistently demonstrated poor resolution in some closely related AMF species
108 (Stockinger et al., 2010) as it is too hyper-variable (Thiéry et al., 2016). As mentioned
109 previously, the ITS region underestimates Glomeromycetes in bulk soil (Berruti et al.,
110 2017). Instead, the 18S region is more commonly used for barcoding AMF, especially in
111 ecological studies (Öpik et al., 2014). Therefore it is important to recognise biases
112 inherent even in supposedly universal barcodes.

113 We previously undertook a nation-wide assessment of soil biodiversity across
114 Wales, representing a breadth of heterogeneous land uses, which included agricultural
115 land, grasslands, woodlands, and upland bogs. In this case, fungal richness and β-

diversity were assessed using soil environmental DNA, utilising ITS1 primers (George et al., 2019). Yet, from the earliest stages of experimental design, we were cognisant that the ITS1 universal primer choice may not account for numerous functionally important fungal groups, particularly AMF. Thus, the primary objective of the present study was to assess whether observed fungal biodiversity (richness and β-diversity) across contrasting land uses from the ITS1 dataset would differ when compared to a dataset derived from an alternative choice of primer and database. We therefore sought to assess if primer choice influenced fungal biodiversity across land use, soil type, and soil organic matter (SOM) class. Our next aim was to critically evaluate the influence of climatic and edaphic factors (e.g. soil pH, total carbon (C), nitrogen (N), phosphorus (P)) on fungal diversity arising from the use of the two different primer sets. Our final aim was to look for differences in coverage of taxonomic and functional diversity between the two primer sets across the broad range of land uses and soil types evaluated.

## MATERIALS AND METHODS

### *Study Design*

Data were collected as part of the Glastir Monitoring & Evaluation Programme (GMEP). The GMEP initiative was established by Welsh Government to monitor their most recent agri-environment scheme, Glastir, which involved 4,911 landowners over an area of 3,263 $km^2$ (Fig. 1). Through the GMEP framework, survey teams collected samples in 2013 and 2014 between April and October in each year (Emmett and the GMEP Team, 2017). Sampling protocols were based on those of the UK-wide ecosystem monitoring programme, Countryside Survey (Emmett et al., 2010). The survey design

randomly located 300, 1 km squares across 26 land classes in Wales which survey teams sampled with 5 plots in each square. A subset of samples were then randomly chosen from squares with a maximum of 3 selected in an individual square. A total of 437 samples were collected for biodiversity analyses.

At each sampling location, 2 cores were collected. One was a 15 cm deep by 4 cm diameter core from which measurements of soil physical and chemical properties were taken, including total C (%), N (%), P (mg/kg), organic matter (% loss-on-ignition), pH (measured in 0.01 M $CaCl_2$), mean soil water repellency (water drop penetration time in seconds), bulk density ($g/cm^3$), volume of rocks ($cm^3$), volumetric water content ($m^3/m^3$), as well as percentage sand and clay. For complete details on chemical analyses methodology, see Emmett et al. (2010). Soil texture data were measured by laser granulometry with a LS320 13 analyser (Beckman-Coulter) as described in George et al. (2019). The cut-off points for clay, silt, and sand were: 2.2 μm, 63 μm and 2000 μm respectively. Clay and sand percentages were selected for subsequent analyses and normalised using Aitchison's $log_{10}$-ratio transformation. Further geographic data including grid eastings, northings, and elevation were also collected. Mean temperature (°C) on date of sample collection and annual precipitation (mL) data were extracted from the Climate Hydrology and Ecology research Support System dataset (Robinson et al., 2017). Environmental variables were normalised (by $log_{10}$ or square root transformation) where appropriate (see Table 1).

Each sampling site was assigned to a land use category, soil type, and SOM class (based on percentage organic matter). The land use classification used in this study was originally developed for the UK Countryside Survey in 1990 (Bunce et al., 1999).

Briefly, vegetation was recorded by surveyors and used to classify each site into one of the 8 Aggregate Vegetation Classes (AVCs) as described in Bunce et al. (1999; for further details please see Supplementary Material). The AVCs have been shown to follow a gradient of soil nutrient content from which productivity and management intensity can also be inferred (see Supplementary Material and Bunce et al., 1999). There were 7 AVCs identified in the present study. The AVCs in descending order of productivity are: Crops/weeds (including arable land), Fertile grassland, Infertile grassland, Lowland woodland, Upland woodland, Moorland grass-mosaic, Heath/bog (Supplementary Table 1). Soil type based on the predominant major soil group classification was extracted from the National Soil Map (Supplementary Material; Avery, 1980). Additionally, we classified soils on a per sample basis by organic matter content. Each sample was grouped into one of four organic matter classes based on percent loss-on-ignition (LOI) following the protocols of the 2007 Countryside Survey (Emmett et al., 2010): mineral (0-8% LOI), humus-mineral (8-30% LOI), organo-mineral (30-60% LOI), and organic (60-100% LOI). Mean values for each environmental variable were recorded for each land use, soil organic matter class, and soil type.

*DNA Extraction*

Soils used in DNA extraction were collected from 15 cm deep by 8 cm diameter cores. Soil samples were transported in refrigerated boxes; samples were received at Environment Centre Wales, Bangor within an average of 48 h post-extraction and frozen at -80 °C upon arrival. Soils were then thawed and homogenised as they passed through a sterilised 2 mm stainless steel sieve after which they were returned to a -80 °C freezer until DNA extraction. Sieves were sterilised between samples by rinsing with tap water at

high pressure and an application of Vircon® laboratory disinfectant followed by UV-treating each side for 5 minutes. DNA was extracted by mechanical lysis from 0.25 g of soil per sample using a PowerLyzer PowerSoil DNA Isolation Kit (MO-BIO Inc.). Soils were pre-treated with 750 μL of a suspension of $CaCO_3$ (1 M) following Sagova-Mareckova et al. (2008) to improve PCR performances, especially for acidic soils. Extracted DNA was stored at -20 °C until amplicon library preparation began. The extractions and homogenisation steps were performed in triplicate. To check for contamination in sieves, 3 negative control DNA extractions were completed as well as 2 negative control kit extractions using the same technique but without the $CaCO_3$ pre-treatment. Aliquots of the resultant DNA were used to create amplicon libraries for sequencing with each primer set.

*Primer Selection and PCR Protocols for Library Preparation*

Amplicon libraries were created using primers for the ITS1 (ITS5/5.8S_fungi) area to specifically target fungi (Epp et al., 2012) and the V4 region of the 18S gene (TAReuk454FWD1/TAReukREV3) (Behnke et al., 2011) targeting a wide range of, but not all, eukaryotic organisms, including fungi. A two-step PCR following protocols devised in conjunction with the Liverpool Centre for Genome Research was used as described in George et al. (2019). Amplification of amplicon libraries was run in triplicate on DNA Engine Tetrad® 2 Peltier Thermal Cycler (BIO-RAD Laboratories Inc.) and thermocycling parameters for both PCR protocols started with 98 °C for 30 s and terminated with 72 °C for 10 min for final extension and held at 4 °C for a final 10 min. For the ITS1 locus, there were 15 cycles of 98 °C for 10 s; 58 °C for 30 s; 72 °C for 30 s.

For the 18S locus there were 15 cycles at 98 °C for 10 s; 50 °C for 30 s; 72 °C for 30 s. Twelve μL of each first-round PCR product were mixed with 0.1 μL of exonuclease I, 0.2 of μL thermosensitive alkaline phosphatase, and 0.7 μL of water and cleaned in the thermocycler with a programme of 37 °C for 15 min and 74 °C for 15 min and held at 4 °C. Addition of Illumina Nextera XT 384-way indexing primers to the cleaned first round PCR products were amplified following a single protocol which started with initial denaturation at 98 °C for 3 min; 15 cycles of 95 °C for 30 s; 55°C for 30 s; 72 °C for 30 s; final extension at 72 °C for 5 min and held at 4 °C. Twenty-five μL of second-round PCR products were purified with an equal amount of AMPure XP beads (Beckman Coulter). Library preparation for the 2013 samples was conducted at Bangor University. Illumina sequencing for both years and library preparation for 2014 samples were conducted at the Liverpool Centre for Genome Research.

*Bioinformatics*

Bioinformatics analyses were performed on the Supercomputing Wales cluster as previously described in George et al. (2019). A total of 104,276,828, and 98,999,009 raw reads were recovered from the ITS1 and 18S sequences, respectively. Illumina adapters were trimmed from sequences using Cutadapt (Martin, 2011) with 10% level mismatch for removal. Sequences were then de-multiplexed, filtered, quality-checked, and clustered using a combination of USEARCH v. 7.0 (Edgar, 2010) and VSEARCH v. 2.3.2 (Rognes et al., 2016). Open-reference clustering (97% sequence similarity) of operational taxonomic units (OTUs) was performed using VSEARCH; all other steps were conducted with USEARCH. Sequences with a maximum error greater than 1 and shorter than 200

231 bp were removed following the merging of forward and reverse reads for ITS1
232 sequences. A cut-off of 250 bp was used for 18S sequences, according to higher quality
233 scores. There were 7,242,508 (ITS1) and 9,163,754 (18S) cleaned reads following these
234 steps. Sequences were sorted and those that only appeared once in each dataset were
235 removed.
236 Remaining sequences were matched first against the UNITE 7.2 (Kõljalg et al.,
237 2013) and SILVA 128 (Quast et al., 2013) databases for the ITS1 and 18S sequences,
238 respectively. Ten per cent of sequences that failed to match were clustered *de novo* and
239 used as a new reference database for failed sequences. Sequences that failed to match
240 with the *de novo* database were subsequently also clustered *de novo*. All clusters were
241 collated and chimeras were removed using the uchime_ref command in VSEARCH.
242 Chimera-free clusters and taxonomy assignment summarised in an OTU table with
243 QIIME v. 1.9.1 (Caporaso et al., 2010) using RDP (Wang et al., 2007) methodology with
244 the UNITE database for ITS1 data. Taxonomy was assigned to the 18S OTU table using
245 BLAST (Altschul et al., 1990) against the SILVA database and OTUs appearing only
246 once or in only 1 sample were removed from each OTU table. Based on DNA quality and
247 read counts, 413 samples were used for analyses of the ITS1 data and 422 for 18S data
248 (from the total of 438).
249 A Newick tree was constructed for the 18S tables using 80% identity thresholds
250 and was paired with the 18S OTU table as part of analyses using the R package phyloseq
251 (McMurdie and Holmes, 2013). Non-fungi OTUs were removed from both OTU tables.
252 Read counts from each group were rarefied 100 times using phyloseq (as justified by
253 Weiss et al. (2017)) and the resulting mean richness was calculated for each sample. The

ITS1 table was rarefied at a depth of 4,000 reads whereas the 18S table was rarefied to 10,000 reads. A subset of the 18S data was rarefied to 400 reads across 398 samples to analyse Glomeromycetes OTUs separately. Samples with observed lower read counts were removed before rarefaction. To assess functional diversity, both OTU tables were processed using FUNGUILD (Nguyen et al., 2016) and the resulting matched OTU tables were used to investigate functional roles based on trophic mode. Sequences have been uploaded to The European Nucleotide Archive and can be accessed with the following primary accession codes after the end of the data embargo: PRJEB28028 (ITS1), and PRJEB28067 (18S).

*Statistical Analysis*

All statistical analyses were run using R v. 3.3.3 (R Core Team, 2017) following rarefaction. For each data set, NMDS ordinations using Bray-Curtis dissimilarity were created with the vegan package (Oksanen et al., 2016) to assess β-diversity. Environmental data was fitted linearly onto each ordination of AVCs using the envfit function. NMDS scores were plotted against these values for each variable to determine the direction of associations. Differences in β-diversity amongst AVCs were calculated with PERMANOVA and homogeneity of dispersion was also assessed.

Linear mixed models were constructed using package nlme (Pinheiro et al., 2016) to show the differences in α-diversity amongst AVCs, soil types, and LOI classification, for both ITS1 and 18S fungal data sets. Sample year as fixed factors; sample square identity was the random factor. This methodology was also used for the subsets of data that matched to the FUNGUILD database. For each model, significant differences were

assessed by ANOVA and pairwise differences were identified using Tukey's *post-hoc* tests from the multcomp package (Hothorn et al., 2008).

Partial least squares regressions from the pls package (Mevik et al., 2016) were used with the variable importance in projection (VIP) approach (Chong and Jun, 2005) to sort the original explanatory variables by order of importance to identify the most important environmental variables for richness. Such analysis is ideal for data where there are many more explanatory variables than sample numbers or where extreme multicollinearity is present (Lallias et al., 2015; George et al., 2019). Variables with VIP values > 1 were considered most important. Relationships between important variables and richness values for each group of organisms were investigated by linear regression. Richness was normalised before regression when necessary.

## RESULTS

### *Soil Properties*

Soil properties displayed a range of changes across land uses (Table 1). Notably, total C ($F_{6, 427} = 89.13$ $p < 0.001$), total N ($F_{6, 427} = 61.03$, $p < 0.001$), C:N ratio ($F_{6, 427} = 94.41$, $p < 0.001$), organic matter content ($F_{6, 428} = 107.02$, $p < 0.001$), elevation ($F_{6, 429} = 78.42$, $p < 0.001$), and mean annual precipitation ($F_{6, 429} = 72.6$, $p < 0.001$), and moisture ($F_{6, 427} = 33.74$, $p < 0.001$) increased with declining land use productivity. We also observed a reduction in pH ($F_{6, 428} = 69.56$, $p < 0.001$), bulk density ($F_{6, 428} = 79.87$, $p < 0.001$), and clay content ($F_{6, 344} = 19.54$, $p < 0.001$) across the land use productivity gradient. Trends in other variables such as soil water repellency ($F_{6, 428} = 22.08$, $p < 0.001$), total P ($F_{6, 424} = 7.1$, $p < 0.001$), sand content ($F_{6, 344} = 5.71$, $p < 0.001$), stone

content ($F_{6, 427} = 10.4$, $p < 0.001$), and temperature at time of sampling ($F_{6, 429} = 4.4$, $p < 0.001$), though significant, were less clear across land uses however. These findings were also apparent when samples were grouped from low-to-high organic matter content by organic matter class (Supplementary Table 2). Overall, no clear trends were evident across the different soil types (Supplementary Table 3).

*Sequencing Data*

A total of 7,582 and 4,408 fungal OTUs were recovered using the ITS1 and 18S primer sets, respectively. Of these, 5,666 were assigned an identifier at the class-level in the ITS1 dataset while 4,367 were assigned an identifier in the 18S dataset. There were 15 classes that were only found in the ITS1 dataset and 12 unique to the 18S data. Endogonomycetes was the most abundant class found only in the ITS dataset (19 OTUs), whereas Laboulbeniomycetes (17 OTUs) was the most abundant fungal class unique to the 18S data. A total of 24 classes were present in both ITS1 and 18S data (Fig. 2A).

As reported in George et al. (2019), Agaricomycetes were the most abundant class of fungi in the ITS1 dataset overall. There were also a large proportion of Sordariomycetes (Fig. 2B). Archaeorhizomycetes was the most abundant class in the 18S dataset (Fig. 2C). Proportionate abundances of Sordariomycetes and Agaricomycetes followed contrasting trends, with the dominance of the former replaced by the latter in lower productivity AVCs in the ITS1 data, as described previously (Fig. 3A). Although Agaricomycetes and Sordariomycetes comprised smaller fractions of the 18S dataset (Fig. 2C), this trend was still apparent (Fig. 3B). Additionally, the Archaeorhiozmycetes from 18S data generally followed the same trend as the Sordariomycetes (Fig. 3B). The

preceding trends observed across land uses are also evident across organic matter classes (Fig. S1) but are not as clear across soil types (Fig. S2).

When a class was present in both datasets, it was usually much more prevalent in one than the other (Supplementary Table 4). For example, there were 1858 Agaricomycetes and 915 Sordariomycetes OTUs in the ITS1, yet these numbers dropped to 646 and 417 OTUs in the 18S dataset. Similarly, Glomeromycetes accounted for 162 of the OTUs in the 18S data, but only 6 OTUs in the ITS1 dataset. Abundances of classes unique to the ITS1 and 18S datasets can be found in Supplementary Table 5 and Supplementary Table 6, respectively.

*Fungal Richness and β-Diversity from ITS1 and 18S Data*

We found that fungal richness followed the same trends across land use, irrespective of primer set. As previously demonstrated in George et al. (2019), fungal OTU richness from ITS1 metabarcoding significantly declined ($F_{6, 258} = 39.87$, $p < 0.001$; Fig. 4A) from high to low productivity/management intensity. Richness in Fertile grasslands was significantly greater than all other AVCs ($p < 0.001$) except Crops/weeds. In the 18S dataset, richness was also significantly higher ($F_{6,267} = 82.73$, $p < 0.001$) in more productive/managed land uses and declined along this gradient. However, richness in grasslands was highest in this dataset (Fig. 4B). For complete pairwise differences between land uses see Supplementary Material.

The trend of declining richness with productivity was also apparent when samples were categorised by organic matter content (Fig. 5). In both datasets, richness was significantly greater ($F_{3, 259} = 48.13$, $p < 0.001$; $F_{3, 269} = 46.71$, $p < 0.001$; for ITS1 and

18S, respectively) in mineral and humus-mineral than all other classifications (ITS1, Fig. 5A; 18S, Fig. 5B). There was no consistent pattern of richness when soils were categorised by soil type (Fig. S3). Again pairwise differences between organic matter classes and soil types are described in the Supplementary Material.

Community composition based on non-metric multidimensional scaling of Bray-Curtis distances also showed consistent trends between the datasets. Plots demonstrate tight clustering of Crops/weeds, and grassland AVCs in both ITS1 (Fig. 6A) and 18S (Fig. 6B) compared to the wide dispersal of other AVCs. Such results are supported by PERMANOVAs, which show significant differences ($F_{6, 406} = 10.74$, $p = 0.001$; $F_{6, 415} = 15.65$, $p = 0.001$); however, analyses of dispersion were also significant ($F_{6, 406} = 41.30$, $p = 0.001$; $F_{6, 415} = 10.69$, $p = 0.001$) as a result of the large disparity in replicates between land uses.

When these results are visualised by organic matter classification, the tight clusters are populated by mineral and humus-mineral samples, whereas organo-mineral and organic samples are more common in the widely dispersed areas of the plots (Fig. S4 and Fig. S5). Soil types are more widely dispersed but Brown and Surface-water gley soils are more common in the tightly grouped area (Fig. S6 and Fig. S7). Again, significant results were observed for both PERMANOVA and dispersion of variance across organic matter classes and soil types in both datasets.

*Relationships Between Soil Properties and Fungal Biodiversity*

Fungal richness showed similar relationships to soil properties in both datasets. Across samples, PLS and VIP analyses highlighted strong correlations between fungal

richness and soil properties. There were significant, positive relationships of richness with pH and bulk density; and significant, negative correlations between richness and C:N ratio, organic matter, elevation, and mean annual precipitation (Table 2). Although these results followed the same trend in ITS1 and 18S data, however, their relative rankings varied. For example, fungal richness from ITS1 data was most strongly correlated with bulk density and organic matter, while richness from 18S data was more strongly correlated to C:N ratio and elevation in addition to bulk density (Table 2). Furthermore, there were some relationships unique to each dataset. Significant negative relationships were observed between richness and soil water repellency. Similarly, richness derived from 18S data was negatively related to total C and sand content of soil but also positively related to clay content.

We found pH was the best predictor of β-diversity from linear fitting for fungi no matter what gene region is amplified (Table 3 and Table 4). All fitted variables were significantly correlated to β-diversity, though most of these only weakly. It is likely that they did not strongly influence the fungal communities. Variables followed similar rankings in both the ITS1 and 18S data. Elevation, annual precipitation, soil moisture, C:N ratio, organic matter, and bulk density all had $R^2$ values greater than 0.35, but their relative order differed between datasets (Table 3 and Table 4).

*Effect of Land Use on Functional Diversity*

There was a distinct difference in trophic modes of OTUs that were successfully matched to the FUNGUILD database between ITS1 and 18S datasets. In total, 3,402 and 1,783 OTUs from the ITS1 and 18S datasets respectively were matched to the

392 FUNGUILD database. Overall, saprotrophs were the most abundant trophic mode in
393 both datasets (Fig. 6); however, pathotrophs ranked second in ITS1 (Fig. 6A) data while
394 the pathotroph-saprotroph-symbiotroph multi-trophic group was second-most abundant in
395 18S data (Fig. 6B). Across land uses, proportions of pathotrophs and pathotroph-
396 saprotroph-symbiotrophs fell with declining productivity (Fig. 7). In matches from the
397 ITS1 data, pathotroph-saprotrophs increased across the productivity gradient (Fig. 7A), as
398 did saprotrophs in the 18S data (Fig. 7B). The aforementioned trend in proportional
399 abundance of pathotrophs and pathotroph-saprotroph-symbiotrophs was also present
400 across organic matter classes (Fig. S8). Symbiotrophs appeared to follow an opposite
401 trend, increasing as productivity fell. Interestingly, this was the case for saprotrophs in
402 the 18S (Fig. S8B) but not the ITS1 (Fig. S8A) dataset. Proportional abundances of
403 fungal OTUs grouped by trophic modes did not follow a discernable pattern across
404 changing soil types (Fig. S9). For simplicity, we focused further analyses only on the
405 broadly defined saprotroph, pathotroph, and symbiotroph groups, ignoring all
406 combination groups; pairwise differences for all of the following comparisons are
407 described in the Supplementary Material.
408 Across land uses, significant differences were observed in the richness of
409 saprotrophic fungi in both the ITS1 ($F_{6,258} = 25.14$, $p < 0.001$) and 18S ($F_{6, 267} = 31.10$, $p$
410 $< 0.001$) data; however, there were differences between datasets (Fig. 8). In the ITS1
411 dataset, richness followed the same trend as overall fungal richness, with the highest and
412 lowest values in the Crops/weeds and Heath/bog AVCs respectively (Fig. 8A). Although
413 this pattern was preserved in the 18S data (Fig. 8B), richness of saprotrophs was much
414 more even across AVCs in this case. Indeed, rather than the linear decline of richness

415 along the productivity gradient, there appeared to be 3 distinct levels in the data affiliated
416 with (i) grassland/agricultural sites, (ii) woodlands, and (iii) bogs.
417 The same pattern was also apparent across organic matter classifications in both
418 datasets (ITS1: $F_{3,\ 260} = 32.86$, $p < 0.001$; 18S: $F_{3,\ 269} = 41.13$, $p < 0.001$; Fig. 9). In the
419 ITS1 dataset, each class was significantly different from the others (Fig. 9A). In the 18S
420 data, saprotroph richness was significantly higher in mineral and humus-mineral soils
421 than organo-mineral and organic soils (all $p < 0.001$ except mineral – organo-mineral $p =$
422 0.02) (Fig. 9B). Again, the overarching trend of fungal richness was not apparent when
423 samples were grouped by soil type. Although there were significant differences across
424 soil types in both the ITS1 ($F_{5,\ 259} = 9.7$, $p < 0.001$) and 18S ($F_{5,\ 268} = 10.73$, $p < 0.001$)
425 datasets, these differences did demonstrate consistent patterns across soil types (Fig.
426 S10).
427 In the case of pathotrophic fungi, richness also followed a similar trend to the
428 saprotrophs across both datasets. In the ITS1 data, significantly ($F_{6,\ 258} = 26.11$, $p <$
429 0.001) greater richness values were observed in Crops/weeds and grassland samples (Fig.
430 8A). Richness of pathotrophs was significantly highest in Crops/weeds sites. Again, this
431 trend was present, though not as clear, in the 18S dataset (Fig. 8B). Significant
432 differences ($F_{6,\ 267} = 52.26$, $p < 0.001$) were observed between AVCs, with the highest
433 richness of pathotrophs occurring in the Fertile grassland and Crop/weeds land uses.
434 Across organic matter classes, significant differences were also observed in
435 pathotroph richness in the ITS1 ($F_{3,\ 250} = 24.91$, $p < 0.001$) and 18S ($_{F3,\ 269} = 30.49$, $p <$
436 0.001) datasets. However, in this case the trends were more apparent in the 18S data than
437 the ITS1 data (Fig. 9). Pathotroph richness was highest in mineral soils and lowest in

organic soils when compared to all other classes in the ITS1 data (Fig. 9A). However, all organic matter classifications were statistically different from each other in the 18S data (Fig. 9B), in descending order from mineral to peat soils. Again, trends were less clear across soil types (Fig. S10). Significant differences were observed in the ITS1 data ($F_{5, 259} = 6.93$, $p < 0.001$) with the lowest pathotroph richness found in peat soils (Fig. S10A). In the 18S data, differences between pathotrophic fungi across soil types were more similar to those observed in other groups (Fig. S10B). Pathotroph richness was significantly ($F_{5, 268} = 13.6$, $p < 0.001$) different across soil types with the highest values found in brown soils and the lowest in peats.

The previously described trend of declining richness across the land use productivity gradient (i.e. Fig. 4) was not apparent when considering symbiotrophs. Furthermore, although significant differences were apparent in both the ITS1 ($F_{6, 258} = 14.88$, $p < 0.001$) and 18S ($F_{6, 267} = 55.13$, $p < 0.001$) datasets they were by no means identical (Fig. 8). Symbiotroph richness was highest in Lowland wood sites followed by Upland wood. This trend was not apparent in the 18S dataset, however (Fig. 8B). Here richness of symbiotrophs was greatest in grassland AVCs and lowest in Heath/bog sites much like the overarching trend of total fungal OTU richness.

When samples were grouped by organic matter class, further discrepancies became apparent between the datasets. Whereas the previously described trend of decreasing richness with increasing organic matter content held true in the 18S data ($F_{3, 269} = 36.28$, $p < 0.001$; Fig. 9B), no significant differences were observed in the ITS1 dataset ($F_{3, 260} = 1.88$, $p = 0.13$; Fig 9A). In the 18S data, richness of symbiotrophs was greater in mineral and humus-mineral soils when compared to organo-mineral ($p = 0.002$,

461 $p = 0.04$, respectively) and organic ($p < 0.001$) soils (Fig. 9B). There were also no
462 significant differences ($F_{5, 259} = 1.43$, $p = 0.21$) in symbiotroph richness across soil types
463 in ITS1 data (Fig. S10A), though there were in 18S data ($F_{5, 259} = 12.52$, $p < 0.001$; Fig.
464 S10B). As described previously, richness was lowest in peat soils and highest in brown
465 soils.
466 We suspected that the differences in functional diversity observed between
467 datasets might be a result of differential coverage of important groups. We were able to
468 confirm this when we analysed the richness of OTUs identified as Glomeromycetes
469 present in the 18S dataset (Fig. 10). All of the 162 Glomeromycetes OTUs were assigned
470 as highly-probable symbiotrophs through FUNGUILD. Across land uses, richness of
471 Glomeromycetes followed similar trends to those of symbiotrophs and saprotrophs from
472 18S data. There were significant ($F_{6, 244} = 33.47$, $p < 0.001$) differences across land uses,
473 though they appeared, like the saprotroph richness to be tiered between grasslands,
474 woods, and bogs (Fig. 10A). Richness of Glomeromycetes was higher in grasslands than
475 all other AVCs except Crops/weeds and lowest in Heath/bog sites. Again, when grouped
476 by organic matter class (Fig. 10B) and soil type (Fig. 10C), Glomeromycetes richness
477 followed the same trend as saprotrophs and symbiotrophs from the 18S dataset. Richness
478 was significantly ($F_{3, 246} = 37.65$, $p < 0.001$) greater in mineral and humus-mineral soils
479 than all others. Across soil types, richness of Glomeromycetes was significantly ($F_{5, 245} =$
480 $8.65$, $p < 0.001$) lower in peat soils when compared to most other soil types.
481
482 *Relationships Between Soil Properties and Fungal Functional Diversity*

Across all samples, PLS and VIP analyses highlighted strong correlations between fungal richness and soil properties by trophic groups. Richness of pathotrophs showed similar relationships to soil properties in both datasets. There were significant, positive relationships of richness with pH and bulk density; and significant negative correlations between richness and total C, C:N ratio, organic matter, elevation, and mean annual precipitation (Table 5). As with the total fungal data, the relative rankings of the strength of relationships between pathotroph and each property varied between datasets. Organic matter was most strongly correlated with pathotroph richness from ITS1 data whereas pH was most strongly correlated with pathotroph richness in the 18S data (Table 5). Also soil moisture content was also negatively correlated with pathotroph richness in the ITS1 dataset only.

Organic matter, elevation (both negative), pH, and bulk density (both positive) all showed significant relationships with saprotroph richness in both datasets (Table 5). The correlations between richness of saprotrophs and both bulk density and pH were the strongest observed in the ITS1 data. There were also negative correlations between saprotroph richness and total C, mean annual precipitation, soil moisture, soil water repellency, and mite abundance in the ITS1 data. However, it again should be noted that the correlation with mites was extremely weak. C:N ratio was strongly and positively correlated with saprotroph richness in the 18S data. Similarly, richness from 18S data was negatively related to total C and sand content of soil but also positively related to clay content. In addition, there was a significant, positive, but weak correlation between sand content and saprotroph richness.

In both datasets, symbiotroph richness was significantly correlated with pH and C:N ratio (Table 5). Interestingly, the relationships were positive in the case of C:N ratio and negative for pH in ITS1 data but the opposite was apparent in the 18S data. There were also many more relationships unique to each dataset. Weak but significant positive relationships were observed between symbiotroph richness and rock volume, Collembola abundance, and temperature as well as a negative correlation to soil moisture. In the 18S data, stronger relationships were observed between symbiotroph richness and bulk density (positive) and elevation (negative). Furthermore a weakly negative correlation was observed with sand content in addition to weak positive correlations with clay content and total P.

## DISCUSSION

### *Primer Choice and the Total Fungal Community*

We observed congruent patterns in total fungal OTU richness across land uses, organic matter classes and soil type when measured with either ITS1 or 18S primer sets. Richness was greater in arable and grassland land uses, which are highly productive, intensively managed and declined in the less productive, largely unmanaged bogs. Although these findings had been previously known from the ITS1 dataset (George et al., 2019), it is important to note that the trend was also present in the fungal OTUs identified from 18S sequencing. A similar trend was observed across organic matter classes. Here, fungal richness fell as organic matter increased. Fungal α-diversity is known to be greater in arable soils than in grasslands or forests (Szoboszlay et al., 2016). Potential mechanisms for this include: (i) increased nutrient availability due to fertiliser input

528 (Szoboszlay et al., 2016), and (ii) beneficial disturbance from tillage and other standard
529 agricultural practices. The latter is consistent with the intermediate disturbance
530 hypothesis whereby high levels of diversity are maintained by consistent interruption of
531 successional processes (Connell, 1978).

532 Soils rich in organic matter, especially peats, found in upland moors, bogs, and
533 other wetlands across harbour distinct fungal communities from neighbouring habitats
534 (Anderson et al., 2003). Fungi dominate microbial communities in bogs (Thormann and
535 Rice, 2007) although their proportional abundance drops sharply below the first 5 cm of
536 bog habitats (Potter et al., 2017). Yet, richness in bogs is consistently low, perhaps due to
537 environmental pressures such as high acidity, highly recalcitrant SOM, low nutrients and
538 oxygen levels (Rousk et al., 2010; Tedersoo et al., 2010) or reduced competition within
539 the fungal community.

540 In comparison to AVC and SOM levels, differences in fungal communities were
541 not as clear across soil types as defined by the National Soil Map (Avery, 1980), which is
542 inline with previous work on microbial activity across the UK (Jones et al., 2014).
543 Richness was highest in brown soils and was lowest in peats. Brown soils commonly
544 support grassland communities across Wales (Avery, 1980; Rudeforth et al., 1984).
545 Nearly half of the Fertile and Infertile grasslands surveyed in GMEP were categorised as
546 brown soils. The absence of other major trends besides these may be due to the use of the
547 dominant soil type and lack of resolution for the soil classification. The soils map used in
548 this study simply does not provide enough resolution (1:63, 360; Avery, 1980) for soil
549 type to be an effective category. Furthermore, this system heavily uses subsoil properties
550 to determine soil type (Avery, 1980), while our work only involved the upper 15 cm.

551 However, it is our opinion that the use of organic matter classification is more effective
552 and simple metric that can be easily implemented in large-scale studies in lieu of fine-
553 scale maps.

554 Results of PLS analyses demonstrates that soil properties and associated
555 environmental factors influencing fungal richness are consistent across ITS1 and 18S
556 datasets. Major drivers included pH, bulk density, C:N ratio, organic matter, elevation,
557 and mean annual temperature (Table 2). Such results from 18S data are consistent with
558 previous findings from the ITS1 data (George et al., 2019). However, there were certain
559 properties that were significant in only one of the datasets and the relative importance of
560 these properties does vary between the two datasets. There are several possible
561 explanations for this. Firstly, 9 more samples were used in the 18S dataset (n = 422) than
562 the ITS1 data (n = 413), which may have introduced the discrepancy in relative
563 importance of the data. However, it is much more likely that a differential coverage of
564 fungal groups between the two datasets caused these discrepancies.

565 Community composition showed consistent clustering across land uses, organic
566 matter classes, and soil types in both data sets. As in George et al. (2019), communities
567 were most similar in the grassland and arable sites and more spread out across woodlands
568 and upland habitats. This was likely driven by environmental factors across Wales. In
569 both datasets, pH was the most important environmental variable influencing community
570 composition and although the remaining properties followed similar patterns, their
571 relative importance again differed in the dataset. The importance of pH, elevation, C:N
572 ratio, and precipitation in determining fungal community composition fits well in the
573 wider context of soil fungi biogeography. Tederoo et al. (2014) previously highlighted

574 the importance of these variables in the distribution of fungi at the global scale.
575 Furthermore, the strong positive correlation with C:N ratio is indicative of the expected
576 fungal dominance (de Vries et al., 2006) of nutrient-poor, acidic soils (Bloem et al.,
577 1997).
578
579 *Primer Choice and Fungal Functional Diversity*
580 Differences between richness of trophic modes of fungi, used here as a proxy for
581 functional diversity, showed some discrepancies across land uses and soil classification
582 between data sets. Saprotrophs made up the largest proportion of the 3 functional groups
583 studied and generally exhibited the same trends as total richness across soils and land
584 uses. This was also the case for pathotrophs. Indeed, correlations between environmental
585 variables with pathotroph and saprotroph richness were largely consistent across datasets.
586 However, we observed divergent trends in symbiotroph richness across land uses and
587 soils. Symbiotroph richness was highest in woodlands in the ITS1 dataset whereas it was
588 highest in grasslands according to the 18S data (Fig. 7A and 7B). A similar increase in
589 richness within grasslands in the 18S data is repeated when Glomeromycetes were
590 considered on their own (Fig. 9); AMF are the predominant mycorrhizal fungi in
591 grassland systems (Smith and Read, 2008). The symbiotroph peak in the ITS1 data may
592 be explained by an increase in coverage of ectomycorrhizas which are the most common
593 group to associate with trees and shrubs (Smith and Read, 2008). Despite these
594 differences, both datasets suggest that symbiotroph richness was low in arable land,
595 which is in line with previous findings demonstrating high susceptibility of mycorrhizal
596 fungi to disturbance, for example tillage (Schnoor et al., 2011; Säle et al., 2015), and the

addition of fertilizers, which decreases the receptiveness of many agricultural plants to mycorrhizal infection (Smith and Read, 2008).

The divergent trend in symbiotroph richness and discrepancies in relationships between functional groups and environmental variables likely stem from primer biases. Primer biases have been well recognised as a confounding factor in categorising communities from environmental DNA (Cai et al., 2013; Elbrecht and Leese, 2015; Tedersoo et al., 2015). Tedersoo et al. (2015) assessed the effectiveness of fungal barcodes from the ITS, 18S, and 28S rDNA regions and found that primer choice did not affect richness or β-diversity results of soil fungi communities from Papua New Guinea, although fewer OTUs were recovered by 18S primers than ITS primers. *In silico* analyses suggests such findings are the result of lumping of sequences in the 18S that may predominantly affect rare sequences, thereby strengthening community matrices. Similarly, results were similar enough for all primers to be suitable for analyses at the class-level (Tedersoo et al., 2015). Although the 18S primers used here were designed to cover the breadth of eukaryotes and may lack specificity to fungi (Behnke et al., 2011), our results show strong congruence to the ITS1 data across total richness and indeed most functional groups.

Unlike Tedersoo et al. (2015) we observed considerable differences in the proportions of fungal classes between the ITS1 and 18S data sets. We suspect that such differences stem from the need to use appropriate databases to assign taxonomy to OTUs to each dataset (Xue et al., 2019). Perhaps only 30%-35% of Glomeromycetes are present in 18S and ITS databases, respectively (Hart et al., 2015), and although sequences are continuously being uploaded to such repositories, it is likely the majority of AMF are not

identifiable from environmental samples (but see Öpik et al., 2014). Similarly we suspect that, although not studied in detail, primer choice may lead to biases in other groups. Archaeorhizomycetes accounted for nearly 25% of the 18S sequences but less than 1% from the ITS1 data (Fig. 1B). Primer bias has been recognised for Archaeorhizomycetes even before the class' formal description; approximately 19% of 18S sequences collected from Anderson et al. (2003), have been matched to Archaeorhizomycetes, whereas none were recovered from the same samples using ITS primers. Despite its recent description, Archaeorhizomycetes are ubiquitous components of soil communities. Strong associations have been observed with trees, yet precise functional roles of these fungi have yet to be determined (Rosling et al., 2011). Subsequently, such biases likely account for divergent relationships between functional group richness and environmental properties.

*Conclusions*

Our comparison of the use of ITS1 and 18S primers and their respective databases in a nationwide metabarcoding survey of fungi yielded 3 major findings. First, the congruent findings of total richness and β-diversity across land use and their relationships to environmental variables confirmed our previous research (George et al., 2019). Second, soil organic matter was found to be a more sensitive metric than soil type in our survey design. Third, biases from the combination of primer and database choice became apparent for certain classes of fungi, including Glomeromycetes and Archaeorhizomycetes, which strongly influenced functional group richness across land uses as well as their relationships with environmental variables. It is therefore important

to recognise the sensitivity of metabarcoding to primer choice, even when using universal primers. Without simultaneous analyses of environmental DNA using both primers and databases, the presence of AM fungi as well as the newly characterised Archaeorhizomycetes would have been overlooked and unquantified in this survey. Furthermore, since the majority of soil biodiversity is undescribed (Ramirez et al., 2015), utilising multiple primers will elucidate a more complete picture of belowground biodiversity by revealing shortcomings in existing probes and revealing the presence of as yet undescribed organisms. We therefore advocate that future nation-wide surveys included both a sample-based metric of soil type (i.e. organic matter classification) and multiple primers for fungal biodiversity. Such measures should not be arduous to implement, especially if researchers can identify specific fungal groups of particular interest to accommodate.

## AUTHOR CONTRIBUTIONS

P.B.L.G., D.L.J., D.A.R. and S.C. conceived this project. Bioinformatics and statistical analyses were led by P.B.L.G. with assistance from S.C. and R.I.G. P.B.L.G. wrote the first draft of the manuscript and S.C., D.A.R., and D.L.J. contributed to subsequent revisions. All authors read and approved the final draft of the manuscript.

## FUNDING

This project was supported as part of the Welsh Government funded Glastir Monitoring & Evaluation Programme (Contract reference: C147/2010/11) and by the UK Natural Environment Research Council (NERC) through the RC/Centre for Ecology &

Hydrology (CEH Project: NEC04782). PBLG was supported by a Soils Training and Research Studentship (STARS) grant from the Biotechnology and Biological Sciences Research Council and NERC [Grant number NE/M009106/1]. STARS is a consortium consisting of Bangor University, British Geological Survey, Centre for Ecology and Hydrology, Cranfield University, James Hutton Institute, Lancaster University, Rothamsted Research and the University of Nottingham.

**ACKNOWLEDGEMENTS**

This research was supported under the Glastir Monitoring & Evaluation Programme (Contract reference: C147/2010/11) and by the UK Natural Environment Research Council (NERC) through the RC/Centre for Ecology & Hydrology (CEH Project: NEC04782). We thank the GMEP Team for their contribution in collecting the data and the laboratory staff of Environment Centre Wales, especially Delphine Lallias, for sample processing and DNA extraction. Thanks also to Fiona M. Seaton for assistance with statistical analyses. We thank Supercomputing Wales for support and training on their system for bioinformatics analyses. We thank John G. Kenny and Richard M. Eccles for generating sequence data at the Centre for Genomic Research, University of Liverpool.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be access online at: []

**DATA AVAILABILITY**

Data associated with this paper will be publically published in the National Environment Research Council (NERC) Environmental Information Data Centre (EIDC). Data are also available from the authors upon reasonable request with permission from the Welsh Government. Sequences with limited sample metadata have been uploaded to the

European Nucleotide Archive and can be accessed with the following primary accession codes after the end of data embargo: PRJEB28028 (ITS1), and PRJEB28067 (18S).

## REFERENCES


Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410. doi: 10.1016/S0022-2836(05)80360-2

Anderson, I. C., Campbell, C. D., and Prosser, J. I. (2003). Diversity of fungi in organic soils under a moorland – Scots pine (*Pinus sylvestris* L). *Environ. Microbiol.* 5, 1121-1132. doi: 10.1046/j.1462-2920.2003.005522.x

Avery, .W. (1980). Soil classifications for England and Wales (Higher Categories). Soil Survey Technical Monograph No 14. Harpenden, UK.

Behnke, A., Engel, M., Christen, R., Nebel, M., Klein, R. R., and Stoeck, T. (2011). Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ. Microbiol.* 13, 340-349. doi: 10.1111/j.1462-2920.2010.02332.x

Berruti, A., Desirò, A., Visentin, S., Zecca, O., Bonfante, P. (2017). ITS fungal barcoding primers versus 18S AMF-specific primers reveal similar AMF-based diversity patterns in roots and soils of three mountain vineyards. *Environ. Microbiol. Rep.* 9, 658-667.

Bloem, J., de Ruiter, P. C., and Bouwman, L. A. (1997). "Soil food webs and nutrient cycling in agroecosystems," in *Modern soil microbiology*, eds. van Elsas, J. D., Jansson, J. K., and Trevors, J. T. (Boca Raton, FL: CRC Press), 245–278.

Bunce, R. G. H., Carr, C. J., Gillespie, M. K., Howard, D. C., Scott, W. A., Smart, S. M., Van de Poll, H. M., and Watkins, J. W. (1999). Vegetation of the British countryside – the countryside vegetation system. (Department of the Environment, Transport and the Regions, London).

Cai, L., Ye, L., Tong, A. H. Y., Lok, S., and Zhang, T. (2013). Biased diversity metrics revealed by bacterial 16S phylotags derived from different primer sets. *PLoS ONE* 8, e53649. doi: 10.1371/journal.pone.0053649

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Wildmann, J., Yaysunenko, T., Zaneveld, J., and Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335-336. doi: 10.1038/nmeth.f.303

Cavender-Bares, J., Kozak, K, H,, Fine, P. V. A., Kembel, S. W. (2009) The merging of community ecology and phylogenetic biology. *Ecol. Lett.* 12, 693-715. doi: 10.1111/j.1461-02548.2009.01314.x

Connell, J. H. (1978). Diversity in tropical rain forests and coral reefs. *Science* **199**, 1302-1310. doi: 10.1126/science.199.4335.1302

Chong, I.-G., and Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometr. Intell. Lab. Syst.* 78, 103-112. doi: 10.1016/j.chemolab.2004.12.011

Deshpande, V., Wang, Q., Greenfield, P., Charleston, M., Porras-Alfaro, A., Kuske, C. R., Cole, J. R., Midgley, D. J., and Tran-Dinh, N. (2016). Fungal identification using a Bayesian classifier and the Warcup training set of internal transcribed spacer sequences. *Mycologia* 108, 1-5. doi: 10.3852/14-293.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **10**, 2460-2461. doi: 10.1093/bioinformatics/btq461

Elbrecht, V., and Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE* 10, e0130324. doi: 10.1371/journal.pone.0130324

Emmett, B. A. & the GMEP Team. (2017). Glastir Monitoring & Evaluation Programme. Final Report to Welsh Government (Contract reference: C147/2010/11. NERC/Centre for Ecology & Hydrology (CEH Project: NEC04780/NEC05371/NEC05782).

Emmett, B. A., Reynolds, B., Chamberlain, P. M., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Lawlor, A. J., Poskitt, J., Potter, E., Robinson, D. A., Scott, A., Wood, C., and Woods, C. (2010). Countryside Survey: Soils Report from 2007. NERC/Centre for Ecology and Hydrology (CS Technical Report No. 9/07, CEH Project Number: C03259).

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., Erséus, C., Gusarov, V. I., Edwards, M. E., Johnsen, A., Stenøien, H. K., Hassel, K., Kauserud, H., Yoccoz, N. G., Bråthen, K. A., Willerslev, E., Taberlet, P., Coissac, E., and Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol. Ecol.* 8, 1821-1833. doi: 10.1111/j.1365-294X.2012.05537.x

Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. (2012). Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484, 186-194. doi: 10.1038/nature10947

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., and Jones, D. L. (2019). Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. *Nat. Commun.* 10, 1107. doi:10.1038/s41467-019-09031-1

Hart, M. M., Aleklett, K., Chagnon, P.-L., Egan, C., Ghignone, S., Helgasson, T., Lekberg, Y., Öpik, M., Pickles, B. J., and Waller, L. (2015). Navigating the labyrinth: a guide to sequence-based, community ecology of arbuscular mycorrhizal fungi. *New Phytol.* 207, 235-247. doi: 10.1111/nph.13340

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general Parametric Models. *Biometrical J.* 50, 346-363. doi: 10.1002/bimj.200810425

Jones, D. L., Simfukwe, P., Hill, P. W., Mills, R. T. E., Emmett, B. A. Evaluation of dissolved organic carbon as a soil quality indicator in national monitoring schemes. *PLoS One* 9(3):e90882. doi: 10.1371/journal.pone.0090882

Kõljalg, U., Nilsson, R. H., Abarenkov, K., Tedersoo, L., Taylor, A. F. S., Bahram, M., Bates, S. T., Burns, T. D., Bengtsson-Palme, J., Callaghan, T. M., Douglas, B., Drenkhan, T., Eberhardt, U., Dueñas, M., Grebenc, T., Griffith, G. W., Hartmann, M., Kirk, P. M., Kohout, P., Larsson, E., Lindahl, B. D., Lücking, R., Martín, M. P., Matheny, P. B., Nguyen, N. H., Niskanen, T., Oja, J., Peay, K. G., Peintner, U.,

778 Peterson, M., Põldmaa, K., Saag, L., Saar, I., Senés, C., Smith, M. E., Suija, A.,
779 Taylor, D. L., Telleria, M. T., Weiss, M., and Larsson, K.-H. (2013). Towards a
780 unified paradigm for sequence-based identification of fungi. *Mol. Ecol.* 22, 5271-
781 5277. doi: 10.1111/mec.12481
782 Lallias, D., Hiddink, J., G., Fonseca, V. G., Gaspar, J. M., Sung, W., Neill, S. P., Barnes,
783 N., Ferrero, T., Hall, N., Lambshead, P. J. D., Packer, M., Thomas, W. K., and Creer,
784 S. (2015). Environmental metabarcoding reveals heterogeneous drivers of microbial
785 eukaryote diversity in contrasting estuarine ecosystems. *ISME J.* **9**, 1208-1221. doi:
786 10.1038/ismej.2014.213
787 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput
788 sequencing reads. EMBnet.journal. 2011. Date of access 04/07/2017. *EMBnet* **17**, 10-
789 12. doi: 10.14806/ej.17.1.200
790 McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible
791 interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217.
792 doi: 10.1371/journal.pone.0061217
793 Mevik, B-H., Wehrens, R., Liland, K. H., and Hiemstra, P. (2016). pls: Partial Least
794 Squares and Principal Component Regression. R package version 2.6-0.
795 https://CRAN.R-project.org/package=pls
796 Nguyen, N. H., Song, Z., Bates, S. T., Branco, S., Tedersoo, L., Menke, J., Schilling, J.
797 S., and Kennedy, P. G. (2016). FUNGuild: an open annotation tool for parsing fungal
798 community datasets by ecological guild. *Fungal Ecol.* 20, 241-248. doi:
799 10.1016/j.funeco.2015.06.006
800 Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D.,
801 Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H.,
802 Szoecs, E., and, Wagner, H. (2016). Vegan: community ecology package. R package
803 version 2.4-0. https://CRAN.R-project.org/package1/4vegan.
804 Öpik, M., Davison, J., Moora, M., and Zobel, M. (2014). DNA-based detection and
805 identification of Glomeromycota: the virtual taxonomy of environmental sequences[1].
806 *Botany* 92, 135-147. doi: 10.1139/cjb-2013-0110
807 Peay, K. G., Kennedy, G., and Bruns, T. D. (2008). Fungal community ecology: a hybrid
808 beast with a molecular master. *BioScience* 58, 799-810. doi: 10.1641/B580907
809 Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., EISPAK authors, Heisterkamp, S., Van
810 Willigen, B., and R Core Team (2016). nlme: Lineral and Nonlinear Mixed Effects
811 Models. R package version 3.1e138. https://CRAN.R-project.org/package1/4nlme
812 Potter, C., Freeman, C., Golyshin, P. N., Ackermann, G., Fenner, N., McDonald, J. E.,
813 Ehbair, A., Jones, T. G., Murphy, L. M., and Creer, S. (2017). Subtle shifts in
814 microbial communities occur alongside the release of carbon induced by drought and
815 rewetting in contrasting peatland ecosystems. *Sci. Rep.-UK* 7, 11314. doi:
816 10.1038/s41598-017-11546
817 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and
818 Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved
819 data processing and web-based tools. *Nucl. Acids Res.* 41 (D1), D590-D596. doi:
820 10.1093/nar/gks1219
821 R Core Team (2017). R: A language and environment for statistical computing. R
822 Foundation for Statistical Computing, Vienna, Austria (2017). https://www.R-
823 project.org/.

824 Ramirez, K. S., Leff, J. W., Barberán, A., Bates, S. T., Betley, J., Crowther, T. W., Kelly,
825 E. F., Oldfield, E. E., Shaw, E. A., Steenbock, C., Bradford, M. A., Wall, D. H.,
826 Fierer, N. (2015). Biogeographic patterns in below-ground diversity in New York
827 City's Central Park are similar to those observed globally. *Proc. R. Soc. B.* 281,
828 20141988. doi: 10.1098/rspb.2014.1988
829 Robinson, E. L., Blyth, E., Clark, D. B., Comyn-Platt, E., Finch, J., and Rudd, A. C.
830 (2017). Climate hydrology and ecology research support system meteorology dataset
831 for Great Britain (1961-2015) [CHESS-met] v1.2. NERC Environmental Information
832 Data Centre.
833 Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a
834 versatile open source tool for metagenomics. *PeerJ* **4**, e2584. doi: 10.7717/peerj.2584
835 Rosling, A., Cox, F., Cruz-Martinez, K., Ihrmark, K., Grelet, G.-A., Lindahl, B. D.,
836 Menkis, A., and James, T. Y. (2011). Archaeorhizomycetes: unearthing an ancient
837 class of ubiquitous soil fungi. *Science* 333, 876-879. doi: 10.1126/science.1206958
838 Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G.,
839 Knight, R., and Fierer, N. (2010). Soil bacterial and fungal communities across a pH
840 gradient in an arable soil. *ISME J.* 4, 1340-1351. doi: 10.1038/ismej.2010.58
841 Rudeforth, C. C., Hartnup, R., Lea, J. W., Thompson, T. R. E., and Wright, P. S. (1984).
842 Soils and their use in Wales. Harpenden: Lawes Agricultural Trust (Soil Survey of
843 England and Wales).
844 Sagova-Mareckova, M., Cermak, L., Novotna, J., Plhackova, K, Forstova, J., and
845 Kopecky, J. (2008). Innovative methods for soil DNA purification tested in soils with
846 widely differing characteristics. *Appl. Environ. Microb.* 74, 2902-2907. doi:
847 10.1128/AEM.02161-07
848 Säle, V., Aguilera, P., Laczko, E., Mäder, P., Berner, A., Zihlmann, U., van der Heijden,
849 M. G. A., and Oehl, F. (2015). Impact of conservation tillage and organic farming on
850 the diversity of arbuscular mycorrhizal fungi. *Soil Biol. Biochem.* 84, 38-52.
851 doi:10.1016/j.soilbio.2015.02.005
852 Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A.,
853 Chen, W., and the Fungal Barcoding Consortium, (2012). Nuclear ribosomal internal
854 transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *P. Natl.*
855 *Acad. Sci. USA.* 109, 6241-6246. doi: 10.1073/pnas.1117018109
856 Schnoor, T. K., Lekberg, Y., Rosendahl, S., and Olsson, P. A. (2011) Mechanical soil
857 disturbance as a determinant of arbuscular mycorrhizal fungal communities in semi-
858 natural grassland. *Mycorrhia* 21, 211-220. doi: 10.1007/s00572-010-0325-3
859 Schüβler, A., Schwarzott, D., and Walker, C. (2001). A new fungal phylum, the
860 *Glomeromycota*: phylogeny and evolution. *Mycol. Res.* 105, 1413-1421. doi:
861 10.1017/S0953756201005196
862 Smith, S. E., and Read, D. J. (2008). Mycorrhizal Symbiosis 3rd ed. London: Academic
863 Press.
864 Stockinger, H., Krüger, M., and Schüβler, A. (2010). DNA barcoding of arbuscular
865 mycorrhizal fungi. *New Phytol.* 187, 461-474. doi: 10.111/j.1469-8137.2010.03262.x
866 Szoboszlay, M., Dohrmann, A. B., Poeplau, C., Don, A., and Tebbe, C. C. (2017). Impact
867 of land-use change and soil organic carbon quality on microbial diversity in soils
868 across Europe. *FEMS Microbiol. Ecol.* 93, fix146. doi: 10.1093/femsec/fix146

Tedersoo, L., Anslan, S., Bahram, M., Pólme, S., Riit, T., Liiv, I., Kõljalg, U., Kisand, V., Nilsson, R. H., Hildebrand, F., Bork, P., and Abarenkov, K. (2015). Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoing analyses of fungi. *MycoKeys* 10, 1-43. doi: 10.3897/mycokeys.104852

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., Ruiz, L. V., Vasco-Palacios, A. M., Thu, P. Q., Suija, A., Smith, M. E., Sharp, C., Saluveer, E., Saitta, Al., Rosas, M., Riit, T., Ratkowsky, D., Pritsch, K., Põldmaa, K., Piepenbring, M., Phosri, C., Peterson, M., Parts, K., Pärtel, K., Otsing, E., Nouhra, E., Njouonkou, A. L., Nilsson, R. H., Morgado, L. N., Mayor, J., May, T. W., Majuakim, L., Lodge, D. J., Lee, S. S., Larsson, K.-H., Kohout, P., Hosaka, K., Hiiesahu, I., Henkel, T. W., Harend, H., Guo, L., Greslebin, A., Grelet, G., Geml, J., Gates, G., Dunstan, W., Dunk, C., Drenkhan, R., Dearnaley, J., de Kesel, A., Dang, T., Chen, X., Buegger, F., Brearley, F. Q., Bonito, G., Anslan, S., Abell, S., and Abarenkov, K. (2014). Global diversity and geography of soil fungi. *Science* 346(6213): 1256688. doi: 10.1126/science.1256688

Thiéry, O., Vasar, M., Jairus, T., Davison, J., Roux, C., Kivistik, P.-A., Metspalu, A., Milani, L., Saks, Ü., Moora, M., Zobel, M., and Öpik, M. Sequence variation in nuclear ribosomal small subunit, internal transcribed spacer and large subunit regions of *Rhizophagus irregularis* and *Gigaspora margarita* is high and isolate-dependent. *Mol. Ecol.* 25, 2816-2832. doi: 10.1111/mec.13655

Thormann, M. N., and Rice, A. V. (2007). Fungi from peatlands. *Fungal Divers.* 24, 241-299.

de Vries, F. T., Hoffland, E., van Eekeren, N., Brussard, L., and Bloem, J. (2006). Fungal/bacterial ratios in grasslands with contrasting nitrogen management. *Soil Biol. Biochem.* 38, 2092-2103. doi: 10.1016/j.soilbio.2006.01.008

Wang, B., and Qui, Y.-L. (2006). Phylogeneitc distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* 16, 299-363. doi: 10.1007/s00572-005-0033-6

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microb.* 73, 5261-5267. doi: 10.1128/AEM.00062-07

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J. R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E. R., and Knight, R. (2017). Normailzation and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5, 1-18. doi: 10.1186/s40168-017-0237-y

Xue, C., Hao, Y., Pu, X., Penton, C. R., Wang, Q., Zhao, M., Zhang, B., Ran, W., Huang, Q., Shen, Q., and Tiedje, J. M. (2019). Effect of LSU and ITS genetic markers and reference databases on analyses of fungal communities. *Biol. Fert. Soils* 55, 79-88. doi: 10.1007/s00374-018-1331-4

Yuste, J. C., Peñuelas, J., Estiarte, M., Garcia-Mas, J., Mattana, S., Ogaya, R., Pujol, M., and Sardans, J. (2011). Drought-resistant fungi control soil organic matter decomposition and its response to temperature. *Glob. Change Biol.* 17, 1475-1486. doi 10.1111/j.1365-2486.2010.02300.x

*Captions*

**Fig. 1.** Map of sites selected for GMEP monitoring. To protect landowner anonymity, each triangle gives an approximate location of every 1 $km^2$ plot from which samples were taken

**Fig. 2.** Composition of fungal classes from ITS1 and 18S datasets. **A)** Venn diagram denoting total number of shared and unique classes in each data set, following exclusion of unknown sequences. Sankey diagrams of proportional abundances of fungal OTUs from all samples from **B)** ITS1 data and **C)** 18S data. Arms denote proportions of OTUs of the most populous classes.

**Fig. 3.** Proportionate abundances of fungal OTUs for **A)** ITS1 and **B)** 18S data across Aggregate Vegetation Class. Aggregate Vegetation Classes are ordered from most (Crops/weeds) to least (Heath/bog) productive.

**Fig. 4.** Boxplots of fungal OTU richness for **A)** ITS1 and **B)** 18S datasets plotted against Aggregate Vegetation Class. Aggregate Vegetation Classes are ordered from most (Crops/weeds) to least (Heath/bog) productive. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping notches are a proxy for non-significant differences between medians. Black dots are outliers.

**Fig. 5.** Boxplots of fungal OTU richness for **A)** ITS1 and **B)** 18S datasets plotted against organic matter class. Organic matter classes are listed in order of increasing percent organic matter. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping notches are a proxy for non-significant differences between medians. Black dots are outliers.

**Fig. 6.** Non-metric dimensional scaling ordinations of fungal community composition across GMEP sites. Samples are coloured by Aggregate Vegetation Class. Data from ITS1 (stress = 0.13) is shown in **A)**; Data from 18S (stress = 0.11) is shown in **B)**.

**Fig. 7.** Proportionate abundances of fungal OTUs matched to FUNGuild trophic groups for **A)** ITS1 and **B)** 18S data across Aggregate Vegetation Classes. Aggregate Vegetation Classes are ordered from most (Crops/weeds) to least (Heath/bog) productive. Abbreviations for multi-trophic mode groups are as follows: Path.-Sap. (Pathotroph-Saprotroph); Path.-Sap.-Sym. (Pathotroph-Saprotroph-Symbiotroph); Path.-Sym. (Pathotroph-Symbiotroph); Sap.-Path.-Sym (Saprotroph-Pathotroph-Symbiotroph); Sap.-Sym. (Saprotroph-Symbiotroph).

**Fig. 8.** Boxplots of richness of fungal OTUs matched to the pathotrophic, saprotroph, and symbiotroph trophic modes in FUNGuild for **A)** ITS1 and **B)** 18S datasets plotted against Aggregate Vegetation Class. Aggregate Vegetation Classes are ordered from most

(Crops/weeds) to least (Heath/bog) productive. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping notches are a proxy for non-significant differences between medians. Black dots are outliers.

**Fig. 9.** Boxplots of richness of fungal OTUs matched to the pathotrophic, saprotroph, and symbiotroph trophic modes in FUNGuild for **A)** ITS1 and **B)** 18S datasets plotted against organic matter class. Organic matter classes are listed in order of increasing percent organic matter. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping notches are a proxy for non-significant differences between medians. Black dots are outliers.

**Fig. 10.** Boxplots of richness of Glomeromycetes OTUs plotted against **A)** Aggregate Vegetation Class; **B)** organic matter class; **C)** soil type. Aggregate Vegetation Classes are ordered from most (Crops/weeds) to least (Heath/bog) productive. Organic matter classes are listed in order of increasing percent organic matter. Soils are listed in increasing order of moisture retention. Boxes cover the first and third quartiles and horizontal lines denote the median. Black dots represent outliers beyond the whiskers, which cover 1.5X the interquartile range. Notches indicate confidence interval around the median. Overlapping

981 notches are a proxy for non-significant differences between medians. Black dots are
982 outliers.
983
984

985 **Table 1.** Mean values (± SE) of soil physical and chemical variables for each Aggregate Vegetation Class. Following normalisation
986 on selected variables (see below), ANOVAs and Tukey's *post-hoc* tests were performed.

| Environmental variable | Crops/weeds | Fertile grassland | Infertile grassland | Lowland wood | Upland wood | Moorland grass-mosaic | Heath/bog |
|---|---|---|---|---|---|---|---|
| **Total C (%)$^{L}$** | 3.87 (± 0.83)d | 4.75 (± 0.2)d | 5.85 (± 0.33)d | 5.78 (± 1.07)d | 9.7 (± 2.25)c | 12.19 (± 2.07)b | 23.57 (± 1.88)a |
| **Total N (%)$^{L}$** | 0.32 (± 0.05)d | 0.45 (± 0.02)d | 0.49 (± 0.02)d | 0.4 (± 0.06)d | 0.58 (± 0.1)c | 0.83 (± 0.11)b | 1.05 (± 0.09)a |
| **C:N ratio$^{S}$** | 11.44 (± 0.81)cd | 10.49 (± 0.13)d | 11.62 (± 0.27)cd | 13.92 (± 0.75)bc | 15.86 (± 0.7)b | 14.41 (± 0.42)b | 20.65 (± 0.94)a |
| **Total P (mg/kg)$^{S}$** | 1103.44 (± 145.47)ab | 1194.9 (± 45.53)a | 1045.5 (± 43.3)ab | 601.68 (± 77.68)c | 762.45 (± 61.95)bc | 930.49 (± 57.5)ab | 769.63 (± 50.04)ab |
| **Organic matter (% LOI)$^{L}$** | 7.53 (± 1.62)d | 9.39 (± 0.34)d | 11.25 (± 0.55)d | 10.71 (± 1.7)d | 18.79 (± 4.16)c | 22.99 (± 3.72)b | 39.26 (± 3.6)a |
| **pH ($CaCl_2$)** | 4.73 (± 0.26)b | 5.2 (± 0.08)a | 4.73 (± 0.05)b | 4.31 (± 0.26)b | 3.57 (± 0.1)cd | 3.85 (± 0.09)c | 3.84 (± 0.1)d |
| **Soil water repellency$^{*}$** | 4077.56 (± 3990.72)abc | 264.01 (± 73.28)c | 781.68 (± 137.58)b | 2975.47 (± 2108.12)abc | 1965.87 (± 698.61)a | 4186.13 (± 798.48)a | 3186.4 (± 812.15)a |
| **Volumetric water content ($m^3/m^3$)** | 0.23 (± 0.03)bc | 0.35 (± 0.01)b | 0.34 (± 0.01)b | 0.22 (± 0.02)c | 0.36 (± 0.03)b | 0.46 (± 0.02)a | 0.52 (± 0.02)a |
| **Rock volume (mL)** | 3.95 (± 1.11)abc | 5.25 (± 0.45)b | 5.44 (± 0.42)b | 9.13 (± 2.49)a | 4.41 (± 0.57)ab | 3.25 (± 0.39)c | 1.87 (± 0.21)c |
| **Bulk density ($g/cm^3$)** | 1.03 (± 0.09)a | 0.9 (± 0.02)a | 0.8 (± 0.02)b | 0.71 (± 0.08)b | 0.56 (± 0.04)c | 0.5 (± 0.04)c | 0.47 (± 0.03)d |
| **Clay content (%)$^{A}$** | 22.25 (± 1.85)ab | 25.46 (± 0.65)a | 23.18 (± 0.64)ab | 17.47 (± 1.34)ab | 17.82 (± 1.82)ab | 18.12 (± 1.27)c | 11.76 (± 2.24)d |
| **Sand content (%)$^{A}$** | 30.97 (± 4.66)ad | 24.88 (± 1.25)d | 29.21 (± 1.44)bd | 42.99 (± 4.01)ac | 40.23 (± 4.15)abc | 29.5 (± 3.0)b | 45.15 (± 7.61)a |
| **Elevation (m)** | 88.71 (± 47.69)cd | 109.38 (± 8.62)d | 167.28 (± 8.65)c | 119.06 (± 16.38)cd | 297.83 (± 20.62)b | 406.63 (± 19.22)a | 380.55 (± 19.7)a |
| **Mean annual precipitation (mL)** | 968.44 (± 69.01)c | 1078.19 (± 24.71)c | 1177.05 (± 18.91)c | 1100.12 (± 52.28)c | 1405.33 (± 65.35)b | 2027.23 (± 74.39)a | 1771.2 (± 58.19)a |
| **Temperature (°C)** | 12.64 (± 1.18)ab | 12.09 (± 0.41)b | 13.44 (± 0.29)a | 15.80 (± 0.87)a | 14.53 (± 0.53)a | 14.51 (± 0.36)a | 13.87 (± 0.29)a |

987 Note: $^{A}$ denotes Aitchison's $log_{10}$-ratio transformation; $^{L}$ denotes $log_{10}$-transformation; square-root-transformation; *Soil water
988 repellency was derived from median water drop penetration times (s) and $log_{10}$ transformed
989

**Table 2.** Results of partial least squares regressions for fungal richness against environmental variables. Positive relationships are underlined; negative relationships are written in italics. *** indicates P < 0.001, ** 0.001 > P < 0.01, * 0.01 > P < 0.05, blank indicates P > 0.05.

| Soil and environmental variables | Fungi (ITS) | Fungi (18S) |
|---|---|---|
| **Total $C^{L}$** | 0.44 | *1.03 ($R^2$ = 0.38$^{***}$)* |
| **Total $N^{L}$** | 0.93 | 0.56 |
| **C:N ratio$^{S}$** | *1.64 ($R^2$ = 0.28$^{***}$)* | *1.71 ($R^2$ = 0.41$^{***}$)* |
| **Total $P^{S}$** | 0.70 | 0.87 |
| **Organic matter (% LOI)$^{L}$** | *1.13 ($R^2$ = 0.29$^{***}$)* | *1.17 ($R^2$ = 0.38$^{***}$)* |
| **pH ($CaCl_2$)** | 1.52 ($R^2$ = 0.23$^{***}$) | 1.55 ($R^2$ = 0.37$^{***}$) |
| **Soil water repellency$^{L}$** | *1.23 ($R^2$ = 0.13$^{***}$)* | 0.82 |
| **Volumetric water content ($m^3/m^3$)** | 0.60 | 0.70 |
| **Rock volume (mL)** | 0.64 | 0.43 |
| **Bulk density ($g/cm^3$)** | 1.41 ($R^2$ = 0.29$^{***}$) | 1.33 ($R^2$ = 0.41$^{***}$) |
| **Clay content (%)$^{A}$** | 0.84 | 1.19 ($R^2$ = 0.11$^{***}$) |
| **Sand content (%)$^{A}$** | 0.6 | *1.11 ($R^2$ = 0.1$^{***}$)* |
| **Elevation (m)** | *1.68 ($R^2$ = 0.22$^{***}$)* | *1.83 ($R^2$ = 0.41$^{***}$)* |
| **Mean annual precipitation (mL)** | *1.44 ($R^2$ = 0.18$^{***}$)* | *1.52 ($R^2$ = 0.27$^{***}$)* |
| **Temperature (°C)** | 0.56 | 0.52 |

Note: $^{A}$ denotes Aitchison's $\log_{10}$-ratio transformation; $^{L}$ denotes $\log_{10}$-transformation; $^{S}$ denotes square-root-transformation.

**Table 3.** Summary of relationships amongst environmental factors and fungal communities based on ITS data. +/- signify the direction of association between each variable and respective NMDS axes.

| **Variable** | | **Correlation** | |
|---|---|---|---|
| | $R^2$ | Axis1 | Axis2 |
| **pH ($CaCl_2$)** | $0.6^{***}$ | - | + |
| **C:N ratio**$^{S}$ | $0.47^{***}$ | + | - |
| **Elevation (m)** | $0.41^{***}$ | + | - |
| **Volumetric water content ($m^3/m^3$)** | $0.41^{***}$ | + | - |
| **Mean annual precipitation (mL)** | $0.39^{***}$ | + | - |
| **Bulk density ($g/cm^3$)** | $0.38^{***}$ | - | + |
| **Organic matter (% LOI)**$^{L}$ | $0.37^{***}$ | + | - |
| **Total C**$^{L}$ | $0.31^{***}$ | + | - |
| **Clay content (%)**$^{A}$ | $0.28^{***}$ | - | + |
| **Soil water repellency**$^{L}$ | $0.24^{***}$ | + | - |
| **Total N (%)**$^{L}$ | $0.21^{***}$ | + | - |
| **Sand content (%)**$^{A}$ | $0.19^{***}$ | + | + |
| **Total P (mg/kg)**$^{S}$ | $0.11^{***}$ | - | - |
| **Rock volume (mL)** | $0.07^{***}$ | - | + |
| **Temperature (°C)** | $0.04^{***}$ | - | + |

Note: $^{A}$ denotes Aitchison's $\log_{10}$-ratio transformation; $^{L}$ denotes $\log_{10}$-transformation; $^{S}$ denotes square-root-transformation

**Table 4.** Summary of relationships amongst environmental factors and fungal communities based on 18S data. +/- signify the direction of association between each variable and respective NMDS axes.

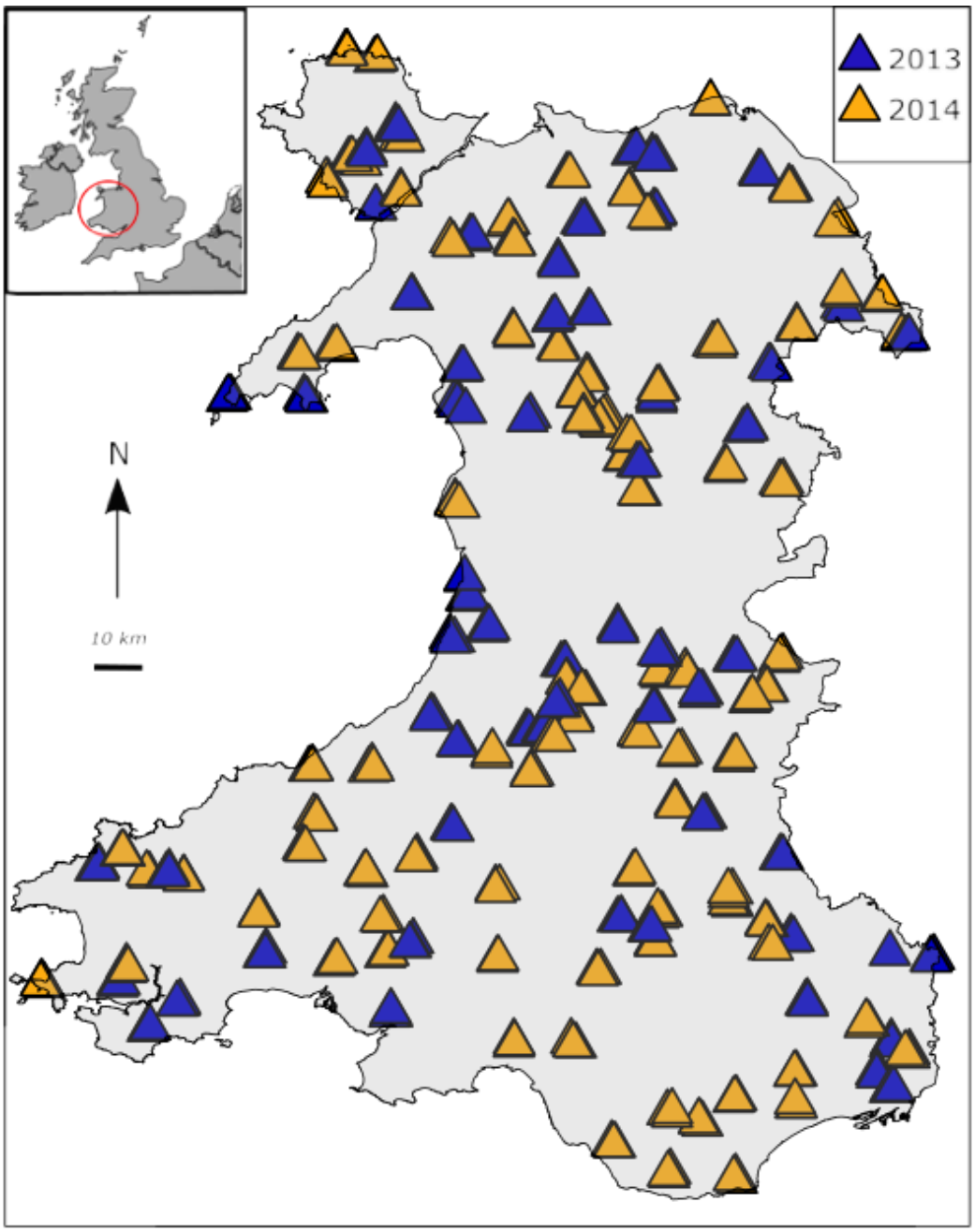
| **Variable** | | **Correlation** | |
|---|---|---|---|
| | $R^2$ | Axis1 | Axis2 |
| **pH ($CaCl_2$)** | $0.61^{***}$ | - | + |
| **Elevation (m)** | $0.50^{***}$ | + | - |
| **Mean annual precipitation (mL)** | $0.46^{***}$ | + | - |
| **Volumetric water content ($m^3/m^3$)** | $0.45^{***}$ | + | - |
| **C:N ratio$^S$** | $0.43^{***}$ | + | + |
| **Organic matter (% LOI)$^L$** | $0.43^{***}$ | + | + |
| **Bulk density ($g/cm^3$)** | $0.39^{***}$ | - | - |
| **Total C$^L$** | $0.34^{***}$ | + | + |
| **Clay content (%)$^A$** | $0.30^{***}$ | - | + |
| **Total N (%)$^L$** | $0.28^{***}$ | + | - |
| **Soil water repellency$^L$** | $0.21^{***}$ | + | - |
| **Sand content (%)$^A$** | $0.14^{***}$ | + | + |
| **Total P (mg/kg)$^S$** | $0.10^{***}$ | - | - |
| **Rock volume (mL)** | $0.06^{***}$ | - | + |
| **Temperature (°C)** | $0.05^{***}$ | - | + |

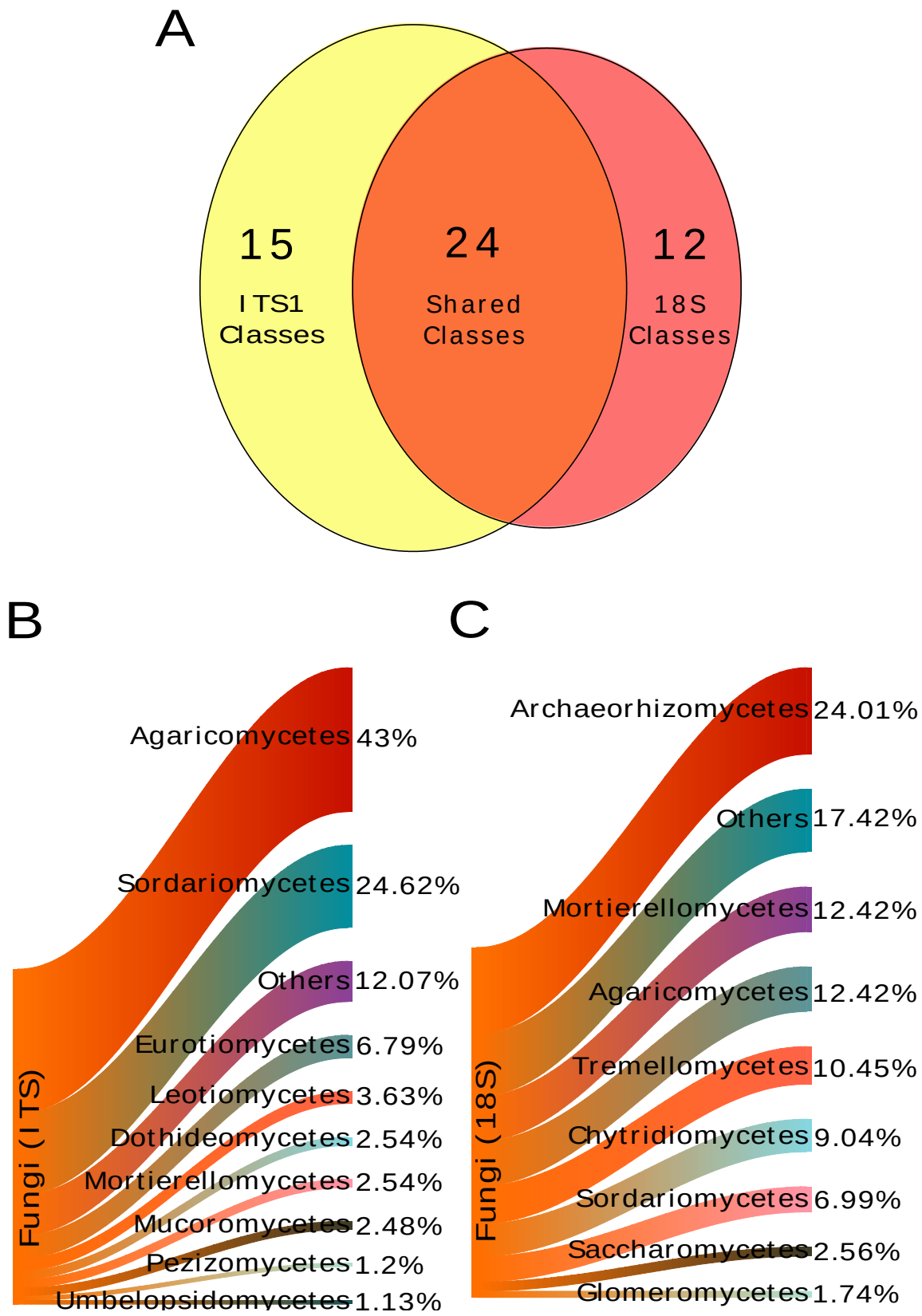
Note: $^A$ denotes Aitchison's $\log_{10}$-ratio transformation; $^L$ denotes $\log_{10}$-transformation; $^S$ denotes square-root-transformation

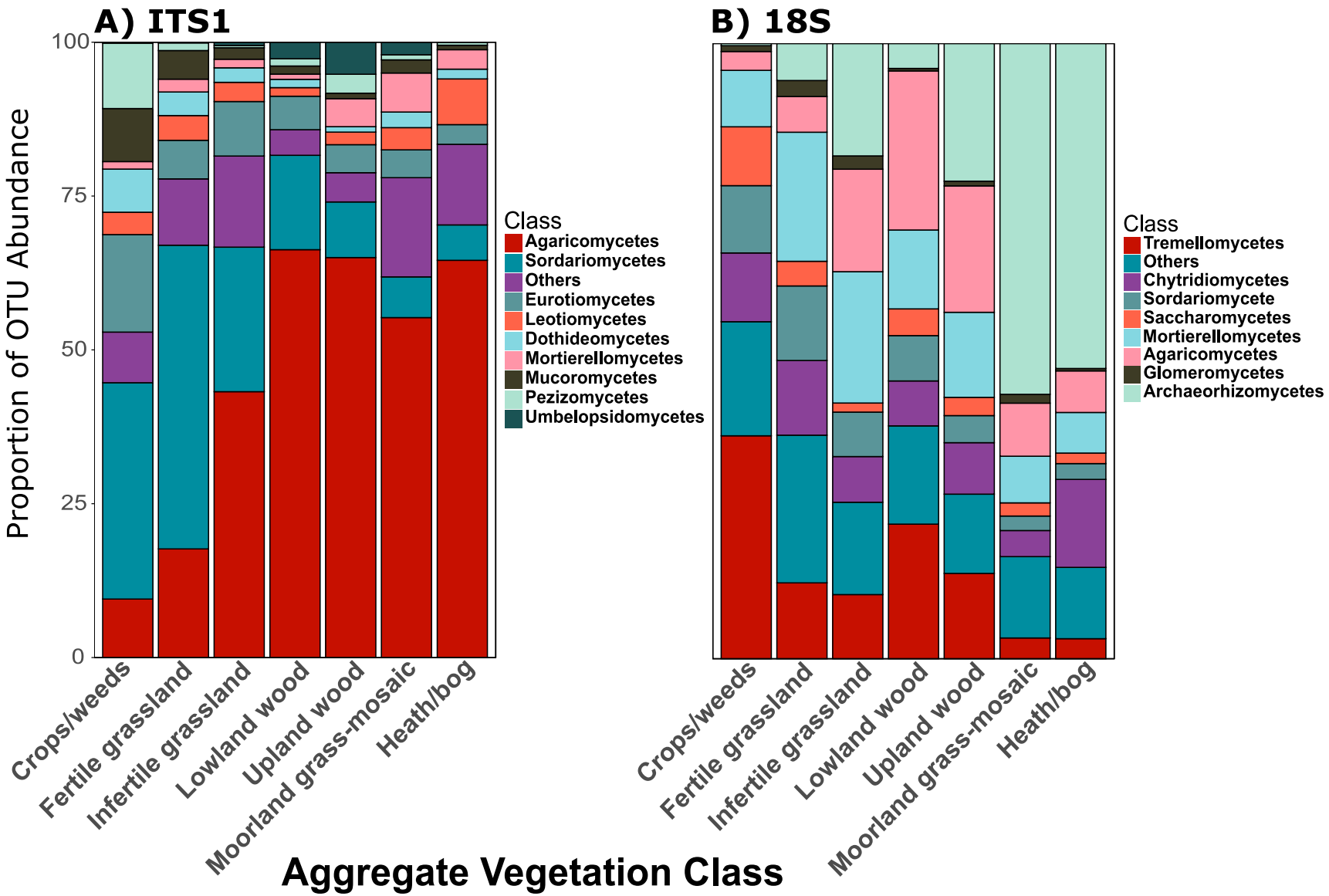
**Table 5.** Results of partial least squares regressions for richness of OTUs classified by trophic mode from FUNGUILD analyses against environmental variables. Positive relationships are underlined; negative relationships are written in italics. *** indicates P < 0.001, ** 0.001 > P < 0.01, * 0.01 > P < 0.05, blank indicates P > 0.05.

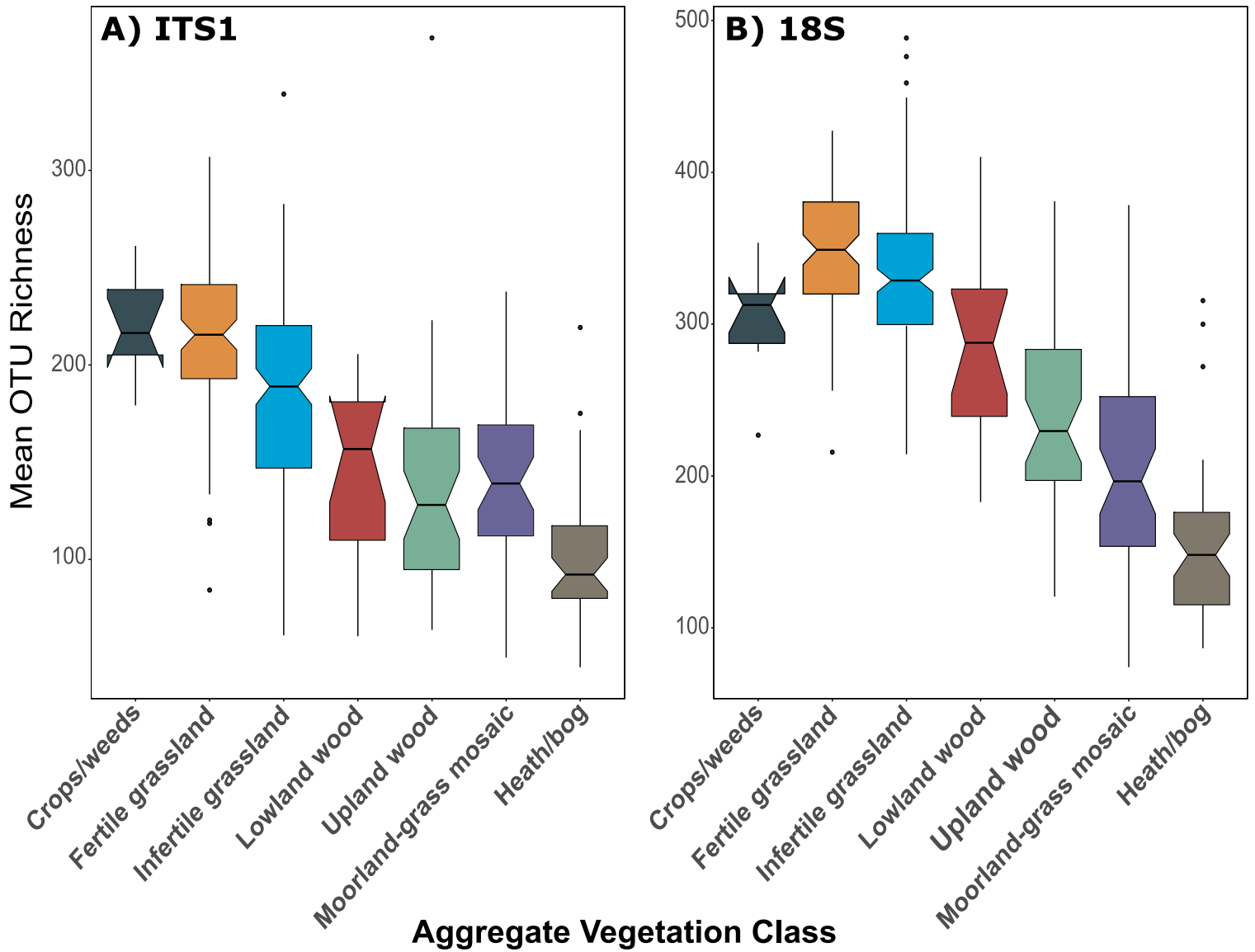
| Soil and environmental variables | Saprotrophs (ITS) | Saprotrophs (18S) | Pathotrophs (ITS) | Pathotrophs (18S) | Symbiotrophs (ITS) | Symbiotrophs (18S) |
|---|---|---|---|---|---|---|
| **Total C (%)$^{L}$** | *1.1 ($R^2$ = 0.24$^{***}$)* | 0.89 | *1.07 ($R^2$ = 0.17$^{***}$)* | *1.0 ($R^2$ = 0.25$^{***}$)* | 0.24 | 0.99 |
| **Total N (%)$^{L}$** | 0.99 | 0.10 | 0.82 | 0.64 | *1.17 ($R^2$ = 0.02$^{**}$)* | 0.10 |
| **C :N ratio$^{S}$** | 0.95 | *2.31 ($R^2$ = 0.28$^{***}$)* | *1.22 ($R^2$ = 0.16$^{***}$)* | *1.41 ($R^2$ = 0.25$^{***}$)* | 1.69 ($R^2$ = 0.01$^{*}$) | *2.47 ($R^2$ = 0.34$^{***}$)* |
| **Total P (mg/kg)$^{S}$** | 0.07 | 0.86 | 0.75 | 0.75 | 1.38 | 1.31 ($R^2$ = 0.02$^{*}$) |
| **Organic matter (% LOI)$^{L}$** | *1.36 ($R^2$ = 0.28$^{***}$)* | *1.02 ($R^2$ = 0.24$^{***}$)* | *1.38 ($R^2$ = 0.21$^{***}$)* | *1.16 ($R^2$ = 0.28$^{***}$)* | 0.37 | 0.92 |
| **pH ($CaCl_2$)** | 1.34 ($R^2$ = 0.21$^{***}$) | 1.27 ($R^2$ = 0.14$^{***}$) | 1.4 ($R^2$ = 0.16$^{***}$) | 1.98 ($R^2$ = 0.4$^{***}$) | *2.35 ($R^2$ = 0.05$^{***}$)* | 1.45 ($R^2$ = 0.2$^{***}$) |
| **Soil water repellency$^{L}$** | *1.28 ($R^2$ = 0.15$^{***}$)* | 0.36 | 0.84 | 0.98 | 0.3 | 0.62 |
| **Volumetric water content ($m^3/m^3$)** | *1.46 ($R^2$ = 0.22$^{***}$)* | 0.56 | *1.38 ($R^2$ = 0.17$^{***}$)* | 0.99 | *1.42 ($R^2$ = 0.05$^{***}$)* | 0.40 |
| **Rock volume (mL)** | 0.68 | 0.06 | 0.8 | 0.59 | 1.09 ($R^2$ = 0.02$^{**}$) | 0.10 |
| **Bulk density ($g/cm^3$)** | 1.42 ($R^2$ = 0.28$^{***}$) | 1.23 ($R^2$ = 0.2$^{***}$) | 1.71 ($R^2$ = 0.12$^{***}$) | 1.29 ($R^2$ = 0.27$^{***}$) | 0.51 | 1.48 ($R^2$ = 0.26$^{***}$) |
| **Clay content (%)$^{A}$** | 0.71 | 0.74 | 0.90 | 1.17 ($R^2$ = 0.1$^{***}$) | 0.49 | 1.05 ($R^2$ = 0.03$^{**}$) |
| **Sand content (%)$^{A}$** | 0.18 | *1.71 ($R^2$ = 0.05$^{***}$)* | 0.05 | 0.32 | 0.21 | 1.63 ($R^2$ = 0.08$^{***}$) |
| **Elevation (m)** | *1.58 ($R^2$ = 0.25$^{***}$)* | *1.13 ($R^2$ = 0.13$^{***}$)* | *1.6 ($R^2$ = 0.19$^{***}$)* | *1.98 ($R^2$ = 0.39$^{***}$)* | 0.37 | 1.07 ($R^2$ = 0.17$^{***}$) |
| **Mean annual precipitation (mL)** | *1.45 ($R^2$ = 0.23$^{***}$)* | 0.81 | *1.38 ($R^2$ = 0.16$^{***}$)* | *1.49 ($R^2$ = 0.24$^{***}$)* | 0.00 | 0.69 |
| **Temperature (°C)** | 0.09 | 0.49 | 0.21 | 0.43 | 1.17 ($R^2$ = 0.01$^{*}$) | 0.53 |

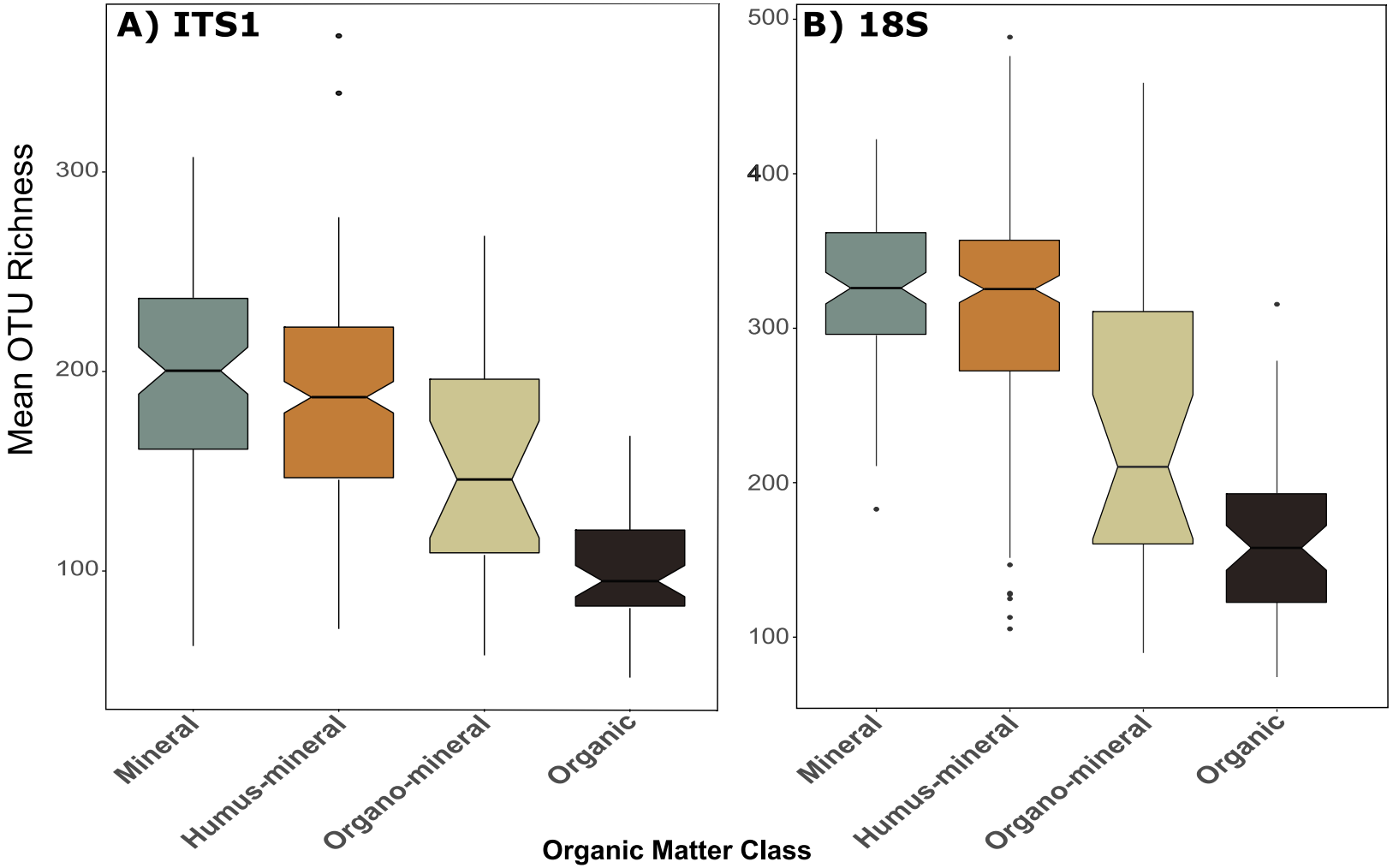
Note: $^{A}$ denotes Aitchison's $log_{10}$-ratio transformation; $^{L}$ denotes $log_{10}$-transformation; $^{S}$ denotes square-root-transformation

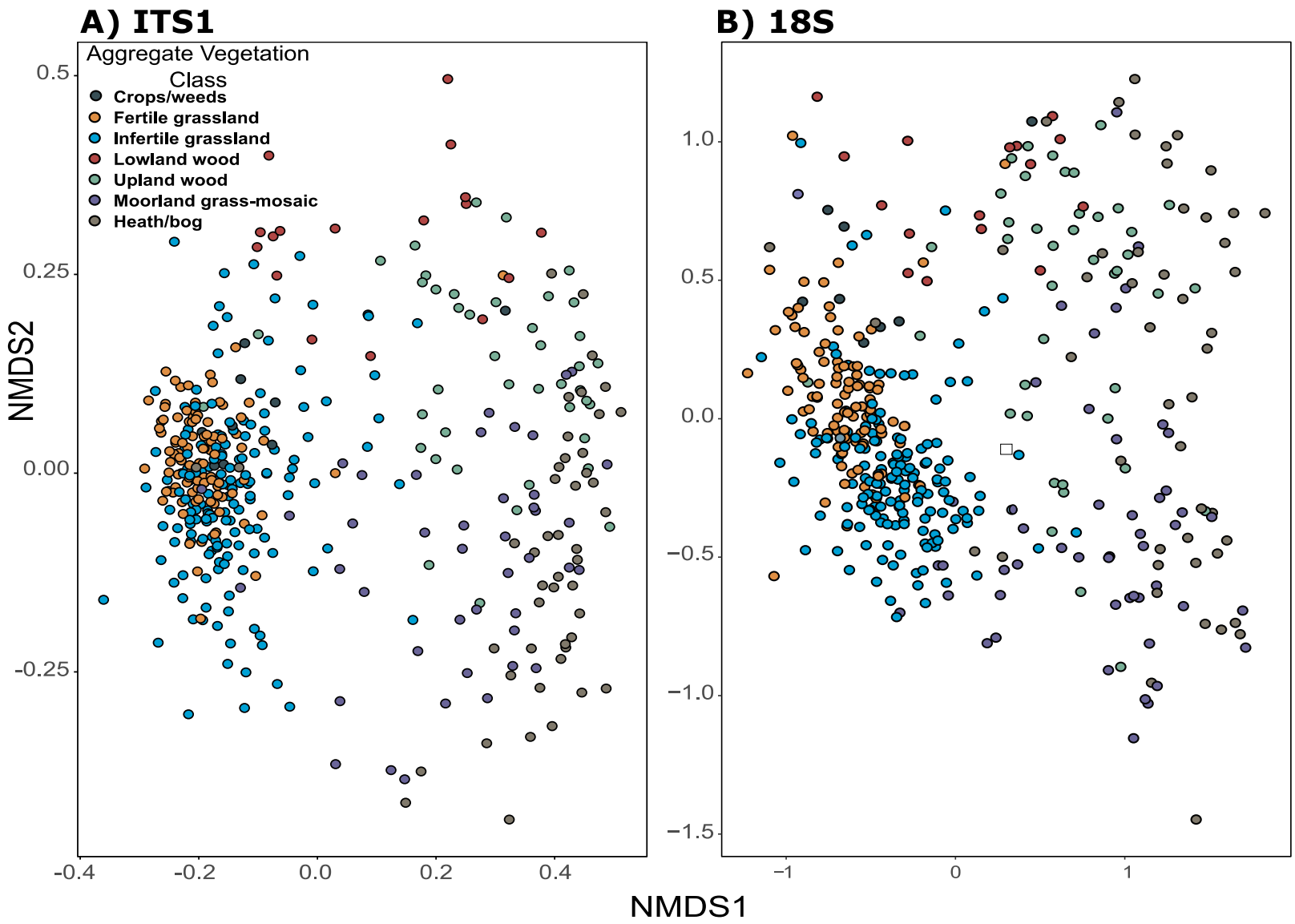

2013
2014
N
10 km

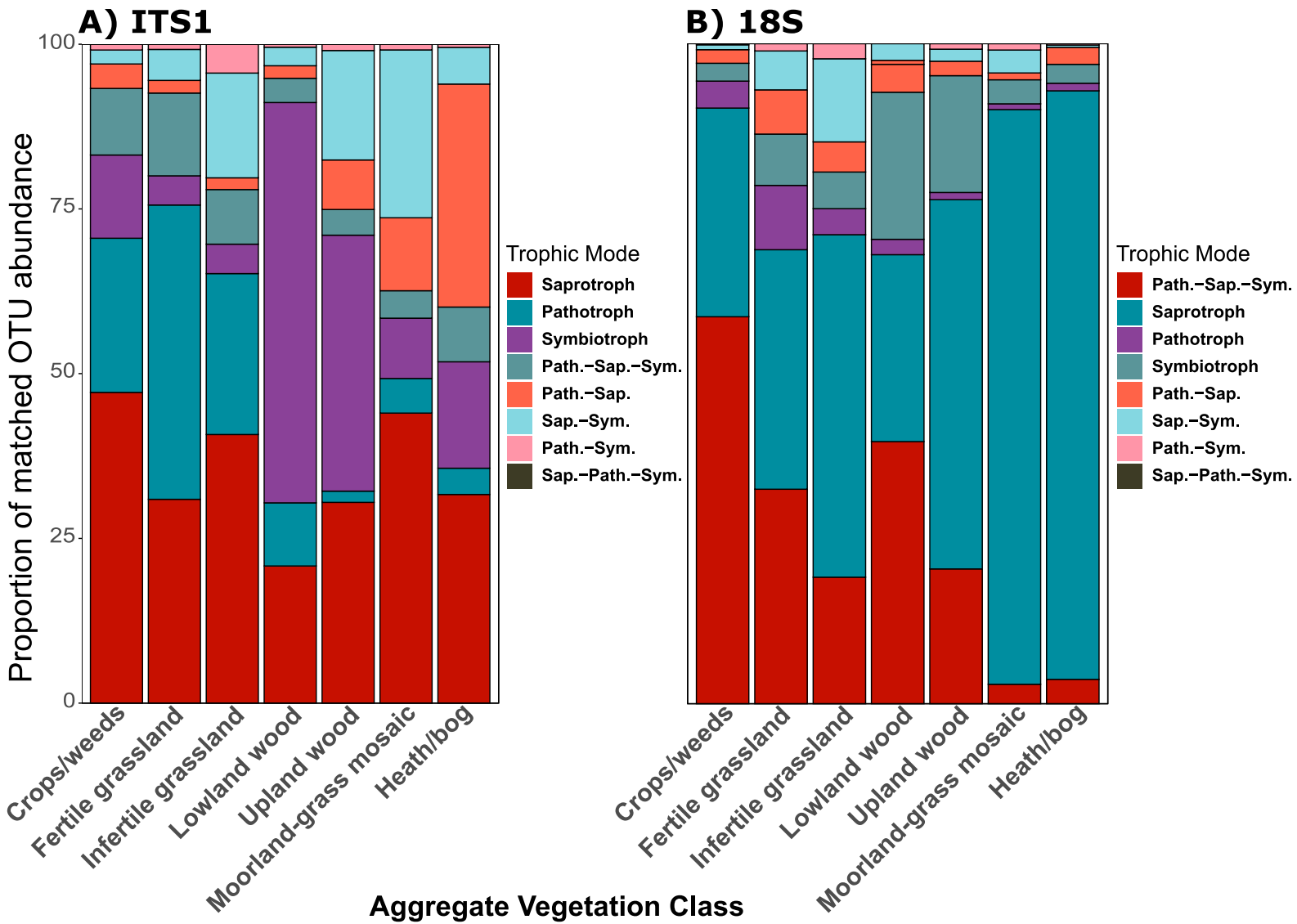

A
15
ITS1
Classes
24
Shared
Classes
12
18S
Classes
B
Fungi (ITS)
Agaricomycetes 43%
Sordariomycetes 24.62%
Others 12.07%
Eurotiomycetes 6.79%
Leotiomycetes 3.63%
Dothideomycetes 2.54%
Mortierellomycetes 2.54%
Mucoromycetes 2.48%
Pezizomycetes 1.2%
Umbelopsidomycetes 1.13%
C
Fungi (18S)
Archaeorhizomycetes 24.01%
Others 17.42%
Mortierellomycetes 12.42%
Agaricomycetes 12.42%
Tremellomycetes 10.45%
Chytridiomycetes 9.04%
Sordariomycetes 6.99%
Saccharomycetes 2.56%
Glomeromycetes 1.74%

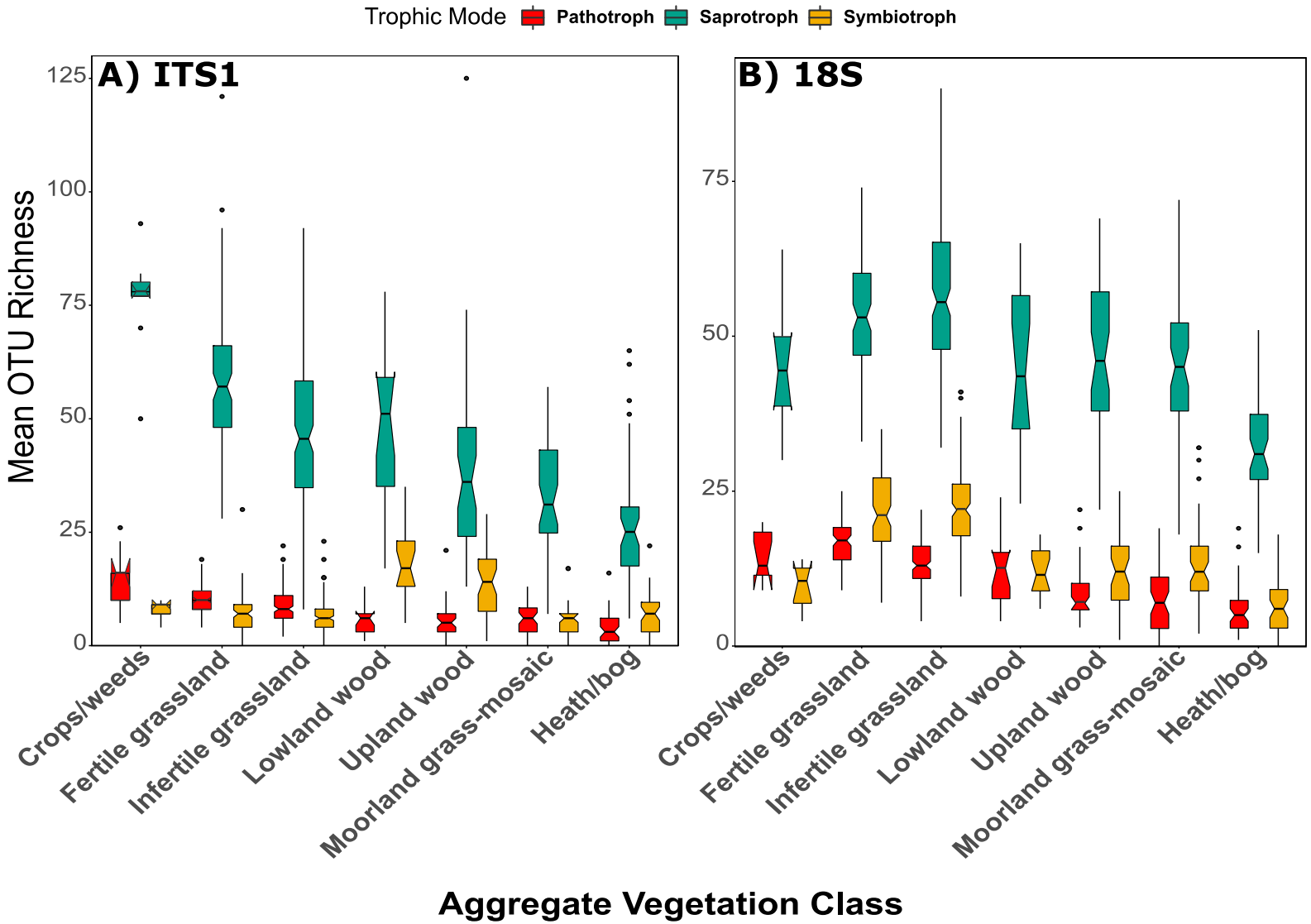

A) ITS1
B) 18S
Proportion of OTU Abundance
100
75
50
25
0
Class
Agaricomycetes
Sordariomycetes
Others
Eurotiomycetes
Leotiomycetes
Dothideomycetes
Mortierellomycetes
Mucoromycetes
Pezizomycetes
Umbelopsidomycetes
Class
Tremellomycetes
Others
Chytridiomycetes
Sordariomycete
Saccharomycetes
Mortierellomycetes
Agaricomycetes
Glomeromycetes
Archaeorhizomycetes
Crops/weeds
Fertile grassland
Infertile grassland
Lowland wood
Upland wood
Moorland grass-mosaic
Heath/bog
Aggregate Vegetation Class

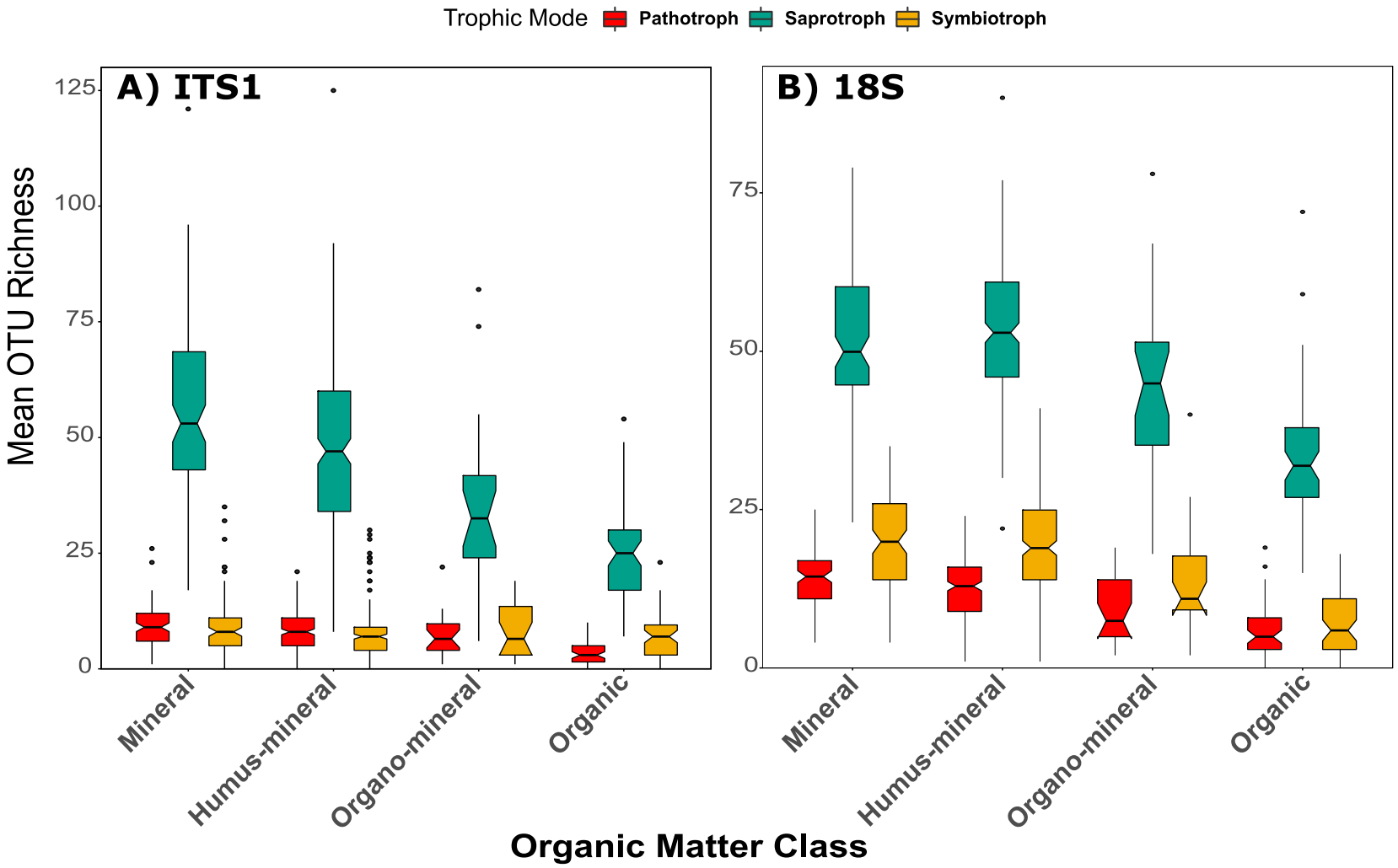

A) ITS1
B) 18S
Mean OTU Richness
300
200
100
500
400
300
200
100
Crops/weeds
Fertile grassland
Infertile grassland
Lowland wood
Upland wood
Moorland-grass mosaic
Heath/bog
Aggregate Vegetation Class

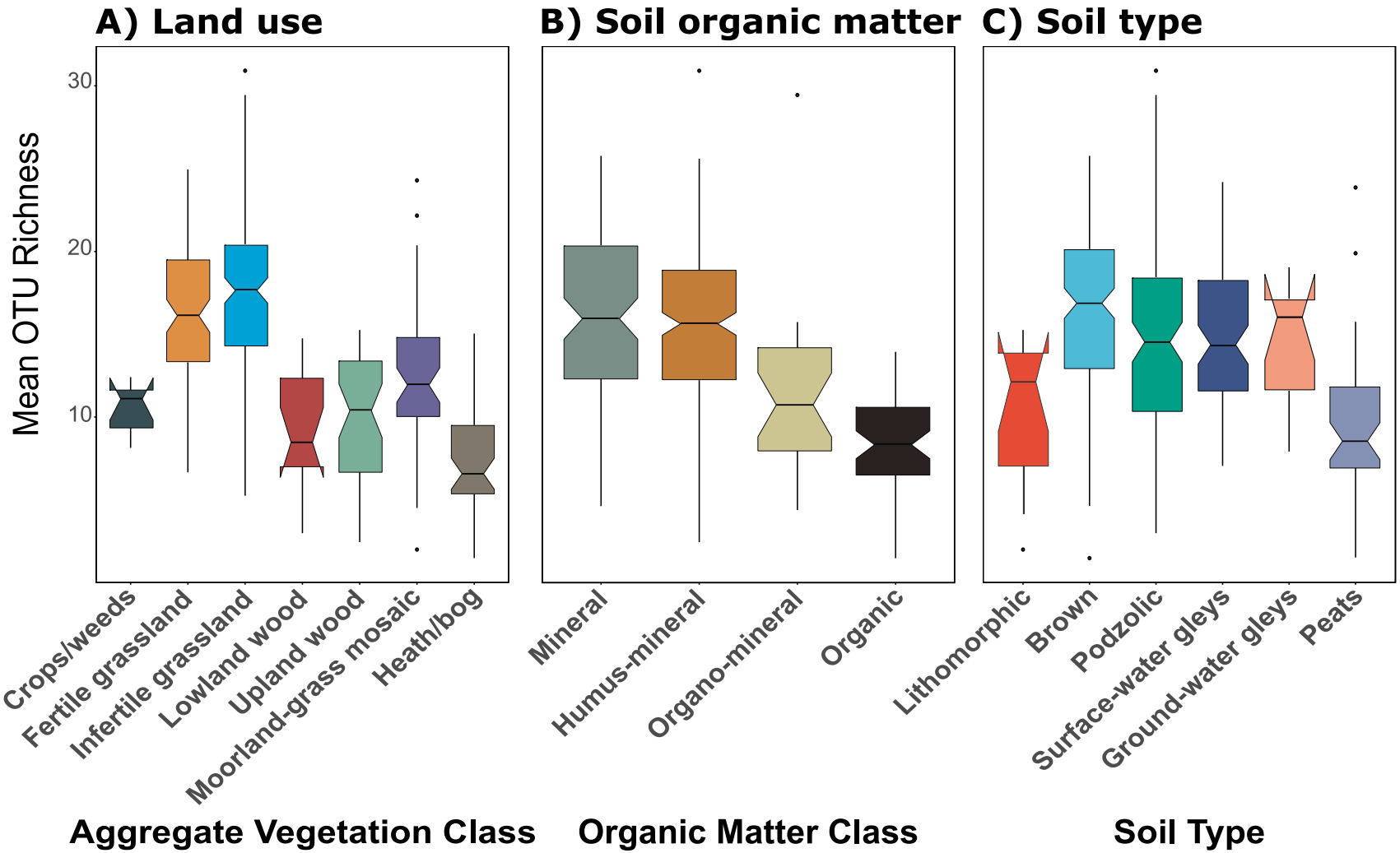

A) ITS1
B) 18S
Mean OTU Richness
300
200
100
500
400
300
200
100
Mineral
Humus-mineral
Organo-mineral
Organic
Organic Matter Class


A) ITS1
B) 18S
Aggregate Vegetation Class
Crops/weeds
Fertile grassland
Infertile grassland
Lowland wood
Upland wood
Moorland grass-mosaic
Heath/bog
NMDS2
NMDS1
0.5
0.25
0.00
-0.25
-0.4
-0.2
0.0
0.2
0.4
1.0
0.5
0.0
−0.5
−1.0
−1.5
−1
0
1


A) ITS1
B) 18S
Proportion of matched OTU abundance
100
75
50
25
0
Crops/weeds
Fertile grassland
Infertile grassland
Lowland wood
Upland wood
Moorland-grass mosaic
Heath/bog
Aggregate Vegetation Class
Trophic Mode
Saprotroph
Pathotroph
Symbiotroph
Path.−Sap.−Sym.
Path.−Sap.
Sap.−Sym.
Path.−Sym.
Sap.−Path.−Sym.
Trophic Mode
Path.−Sap.−Sym.
Saprotroph
Pathotroph
Symbiotroph
Path.−Sap.
Sap.−Sym.
Path.−Sym.
Sap.−Path.−Sym.


Trophic Mode
Pathotroph
Saprotroph
Symbiotroph
A) ITS1
B) 18S
Mean OTU Richness
125
100
75
50
25
0
75
50
25
0
Crops/weeds
Fertile grassland
Infertile grassland
Lowland wood
Upland wood
Moorland grass-mosaic
Heath/bog
Aggregate Vegetation Class


Trophic Mode
Pathotroph
Saprotroph
Symbiotroph
A) ITS1
B) 18S
Mean OTU Richness
125
100
75
50
25
0
75
50
25
0
Mineral
Humus-mineral
Organo-mineral
Organic
Organic Matter Class


A) Land use
B) Soil organic matter
C) Soil type
Mean OTU Richness
30
20
10
Crops/weeds
Fertile grassland
Infertile grassland
Lowland wood
Upland wood
Moorland-grass mosaic
Heath/bog
Mineral
Humus-mineral
Organo-mineral
Organic
Lithomorphic
Brown
Podzolic
Surface-water gleys
Ground-water gleys
Peats
Aggregate Vegetation Class
Organic Matter Class
Soil Type