Comparative analysis of diversity and environmental niches of soil bacterial, archaeal, fungal and protist communities reveal niche divergences along environmental gradients in the Alps

Lucie A Malard (lucie.malard@unil.ch)
University of Lausanne https://orcid.org/0000-0001-9988-2127
Heidi K Mod
University of Helsinki Faculty of Science: Helsingin yliopisto Matemaattis-luonnontieteellinen tiedekunta
Nicolas Guex
Universite de Lausanne
Olivier Broennimann
Université de Lausanne Faculté de biologie et médecine: Universite de Lausanne Faculte de biologie et medecine
Erika Yashiro
Université de Lausanne Faculté de biologie et médecine: Universite de Lausanne Faculte de biologie et medecine
Enrique Lara
Real Jardín Botánico: Real Jardin Botanico
Edward A. D. Mitchell
Université de Neuchâtel Institut de Biologie: Universite de Neuchatel Institut de Biologie
Helene Niculita-Hirzel
Unisanté Center for Primary Care and Public Health Department of Occupational and Environmental Medicine: Unisante Departement sante travail et environnement
Antoine Guisan
Université de Lausanne Faculté de biologie et médecine: Universite de Lausanne Faculte de biologie et medecine

Research Article

Keywords: microbial diversity, ecological niche, environmental filtering, ecological optimum

Posted Date: August 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-609984/v2

Version of Record: A version of this preprint was published at Soil Biology and Biochemistry on April 1st, 2022. See the published version at https://doi.org/10.1016/j.soilbio.2022.108674.

Abstract

Although widely used in ecology, comparative analyses of diversity and niche properties are still lacking for microorganisms, especially concerning niche variations. In this study, we identified important topoclimatic, edaphic, spatial and biotic drivers of the alpha and beta diversity of bacterial, archaeal, fungal and protist communities. Then, we calculated the niche breadth and position of each taxon along environmental gradients within all taxonomic groups, to determine how these vary within and between groups. Quantifying the niches of microbial taxa is necessary to then forecast how taxa and the communities they compose might respond to environmental changes. We found that edaphic properties were the most important drivers of both community diversity and composition for all microbial groups. Protists presented the largest niche breadths, followed by bacteria and archaea, with fungi displaying the smallest. Niche breadth generally decreased towards environmental extremes, especially along edaphic gradients, suggesting increased specialisation of microbial taxa in highly selective environments. Overall, we showed that microorganisms have well defined niches, as do macro-organisms, likely driving part of the observed spatial patterns of community variations. Assessing niche variation more widely in microbial ecology should open new perspectives, especially to tackle global change effects on microbes.

Introduction

A large number of studies have assessed patterns of community variations through analyses of spatial and environmental variations in alpha and beta diversity. Communities of macroorganisms have been shown to be driven by three main drivers: abiotic suitability (i.e. the species' environmental niches), dispersal and biotic interactions (Lortie *et al.*, 2004, D'Amen *et al.*, 2017). An increasing number of similar studies are now being conducted on microorganisms, yet many of them focused so far on diversity patterns and much less on niche quantifications and variations across taxa (Tedersoo *et al.*, 2014, Shi *et al.*, 2016, Delgado-Baquerizo *et al.*, 2018, Ning *et al.*, 2019, Oliverio *et al.*, 2020).

The environmental niche concept is fundamental in ecology and evolutionary biology (Hutchinson, 1957, Holt, 2009). It describes a taxon's distribution and performance in an environmental space and can be used to predict species distribution in a geographic space (Guisan *et al.*, 2017). It has been used to investigate biotic interactions (Wiens, 2011), community dynamics (Cabral & Kreft, 2012), geographic range limits (Kearney & Porter, 2009, Pagel & Schurr, 2012), conservation planning (Costa *et al.*, 2010, Schwartz, 2012), biological invasions (Guisan *et al.*, 2014) and species response to climate change (Thuiller *et al.*, 2005, Botkin *et al.*, 2007, Wiens *et al.*, 2009). The realized environmental niche considers the range of tolerance of taxa (measured by variations in occurrences, abundance, or fitness) along one or more environmental gradients while accounting for biotic interactions and dispersal limitations, defined as a multivariate environmental hypervolume (Hutchinson, 1957). The environmental niche is thus also a key filter of the assemblage of communities, by determining – at least in part - whether an assemblage of species (or operational taxonomic units) is adapted to thrive - now, in the past or in the future - at a given site (D'Amen *et al.*, 2017).

Niche properties such as niche position (average position along an environmental gradient) and breadth (amplitude of tolerance along an environmental gradient) are also strong indicators of a taxon's vulnerability to environmental change (Thuiller *et al.*, 2005, Cianfrani *et al.*, 2018). Niche position is often used to measure the marginality of a taxon's habitat distribution within a study area, with non-marginal taxa occupying typical conditions (i.e., dominant conditions in a landscape) and marginal taxa confined to atypical or extreme conditions (Dolédec *et al.*, 2000, Soininen & Heino, 2007, Muller, 2019). Niche breadth can be used to infer the species tolerance to changing conditions, so that specialist species with narrow niches are expected to experience a higher risk of extinction during times of stress (La Sorte & Jetz, 2010). On the other hand, generalist species with broad environmental tolerance are expected to maintain viable populations more easily even in unfavorable environmental conditions (Thuiller *et al.*, 2005, Binzer *et al.*, 2011, Slatyer *et al.*, 2013). The combination of narrow niche breath and marginal niche position especially increases vulnerability.

While niche theory is widely used in "macroorganisms" ecology (e.g. Gaston & Blackburn (2008)), it is still in its infancy in microbial ecology, and we know relatively little about the size and position of microbial niches along environmental gradients. Yet, microorganisms represent a large proportion of the global soil biodiversity and play essential roles in biogeochemical cycling, including carbon sequestration and soil fertility (Falkowski *et al.*, 2008, Yuan *et al.*, 2012, Bharti *et al.*, 2017, Crowther *et al.*, 2019) and therefore, are essential for ecosystem functioning. Hence, understanding and quantifying the niche of microbial taxa and communities is a pressing need to help understand and forecast the influence of environmental change on microorganisms (Mod *et al.*, 2021).

We deliberately focus on the microbial niche, as the overarching aim of this study is to characterise and compare the environmental niche of four microbial groups - bacteria, archaea, fungi and protists - in the same mountain region with wide environmental gradients. However, to achieve this, we first needed to identify key potential drivers of microorganism communities (and thus the main niche axes). We evaluated the correlation of topoclimatic, edaphic, spatial and biotic variables with the alpha and beta diversity of bacteria, archaea, fungi and protists. This first step also makes this work comparable to previous microbial ecology studies. Such comparisons are essential to understand how the soil biosphere differentially responds to ecosystem properties across landscapes, yet the number of studies comparing microbial groups within the same environment is indeed still limited (Dassen *et al.*, 2017, George *et al.*, 2019). As these four microbial groups have different ecological, physiological and phenotypic properties, we expected different variables to drive their niches and impact their diversity. Specifically, we hypothesised that bacterial and archaeal alpha and beta diversity would primarily be correlated to edaphic properties, especially soil pH and carbon content (Shi *et al.*, 2016, Delgado-Baquerizo *et al.*, 2018, Jiao *et al.*, 2019), while climatic variables, especially temperature and precipitations, would be strongly associated with fungal and protist communities (Tedersoo *et al.*, 2014, Oliverio *et al.*, 2020, Seppey *et al.*, 2020).

Next, based on the identified edaphic and topoclimatic drivers, we calculated the niche breadth and position of all individual microbial taxa in each of the four groups and assessed whether niche breadth

and position differed among the groups. While abiotic conditions have a key role in defining the niche breadth and position of species, other factors, such as biotic interactions and dispersal abilities influence the geographical range of taxa, and therefore, potentially, niche breadth (Gaston, 2000). For microorganisms, studies have shown that smaller body size (< 20 µm) might enhance dispersal capabilities via wind transport (Wilkinson *et al.*, 2012), potentially increasing the geographical range of taxa and the realised niche breadth observed (Rocha *et al.*, 2018). As bacteria and archaea have similar small size ranges (0.1 – 5 µm) (Tesson *et al.*, 2015, Stepanauskas *et al.*, 2017) and high ecological tolerance (Pikuta *et al.*, 2007, Belilla *et al.*, 2019), we hypothesised that they would present the largest niche breadths. On the other hand, we hypothesised that fungi and protists would present smaller niche breadths as they are generally larger (3-100 µm and 5-200 µm respectively) (Caron, 2009, Taylor & Sinsabaugh, 2015, Luan *et al.*, 2020), present higher spatial turnover (Fournier *et al.*, 2020, Kivlin & Hawkes, 2020) and generally have lower ecological tolerance (Pikuta *et al.*, 2007, Taylor & Sinsabaugh, 2015, Belilla *et al.*, 2019).

Finally, we calculated the niche breadth and position along the key environmental gradients identified as important for alpha and beta diversity to evaluate whether and how the two niche properties would be related. We hypothesised that niche breath would be related to niche position so that niche breadth would decrease towards the extremes of environmental gradients because of specialisations required for survival (Pandit *et al.*, 2009, Büchi & Vuilleumier, 2014, Okie *et al.*, 2015).

Materials And Methods

Sample collection

The data were collected from 136 non-forested sites covering ca. 700 km^2 of a mountain area located in the western Swiss Alps spanning a 2695 m elevation range (425-3120 m) [Fig. 1]. Climatic and topographic conditions are heterogeneous, with annual mean temperatures ranging from -5°C to +8°C and mean precipitation ranging between 1200 mm and 2600 mm. The sampling sites were chosen following an equal random-stratified sampling design (Hirzel & Guisan, 2002), with elevation and slope as stratifying factors (Yashiro *et al.*, 2016). Detailed descriptions of the study area and the sampling protocol are provided in Yashiro *et al.* (2016) and Buri *et al.* (2020). In brief, at each site, the vegetation was inventoried in 4 m^2 quadrats by recording vascular plant species presence-absences and cover abundance. Soil samples (top 5 cm) were collected at the four corners and the centre of each quadrat using flame-sterilized shovels and Whirl-Pak bags (Nasco, Fort Atkinson, WI, USA) during the summer 2013. Samples were pooled by site and manually homogenised. A subset of the pooled soil was immediately flash frozen on site in liquid nitrogen and maintained at -80°C until ready for soil properties measurements, while the remainder was maintained on ice until processing in the lab. Within two days of collection in the field, part of the refrigerated soil was sieved and aliquoted, and then stored at -80°C until DNA extraction.

DNA extraction and amplicon sequencing

Soil DNA was extracted in triplicate for each sample using the PowerSoil DNA isolation kit (Qiagen, Carlsbad, CA, USA) following the manufacturer's instructions. Triplicate DNA extracts were pooled by sample for downstream analyses (Yashiro *et al.*, 2016). For bacterial and archaeal groups, the V5 region of the 16S rRNA gene was amplified for each DNA sample using the universal primers 784F-880R (Lazarevic *et al.*, 2009, Yashiro *et al.*, 2016, Mod *et al.*, 2021). The protocol is described in detail in Yashiro *et al.* (2016). For fungi, the internal transcribed spacer (ITS) region was amplified from each sample using the primers ITS1F and ITS2 (White *et al.*, 1990, Schmidt *et al.*, 2013). The protocol is detailed in Pinto-Figueroa *et al.* (2019). Finally, for protists, the V4 region of the 18S rRNA gene was amplified using universal eukaryotic primers TAReuk454FWD1 and TAReukREV3 (Stoeck *et al.*, 2010). The protocol is described in Seppey *et al.* (2020). The bacterial, archaeal and fungal libraries were sequenced on an Illumina HiSeq 2500 platform at the Genomic Technologies Facility of the University of Lausanne, while protist libraries were sequenced on an Illumina MiSeq platform at the University of Geneva (Molecular Systematics & Environmental Genomics Laboratory).

Bioinformatic processing

We opted for a standardized custom-made pipeline to avoid processing biases and allow for the best possible comparison among the groups, as described in Mod *et al.* (2021). In short, we demultiplexed the sequenced reads (of 16S and 18S), removed the barcodes, trimmed and merged the sequences. Amplicons were dereplicated to obtain zero-radius operational taxonomic units (zOTUs) (Edgar, 2018). Finally, we removed the zOTUs with less than 100 reads among all samples to remove chimeras, sequencing errors and potentially rare zOTUs which have been shown to have minimal effect on biodiversity indices (Brown *et al.*, 2015, Clare *et al.*, 2016, Malard, 2019). For protists, we removed all 18S sequences affiliated to Fungi, Metazoa and Embryophyta. For fungi, the same process was applied. However, since the length of the ITS insert is longer than the sequence of the reads, both ends were concatenated with a NNNN separation to permit subsequent alignments against a database of fungal ITS. These sequences were then dereplicated as described above and zOTUs with less than 100 reads across all samples were discarded. Taxonomic assignment was performed against the full SILVA v132 database (Pruesse *et al.*, 2007, Quast *et al.*, 2012, Yilmaz *et al.*, 2014) and with the RDP (ribosomal database project) naive Bayesian classifier v2.12 (Wang *et al.*, 2007) using the 16S rRNA (trainset 18) option for bacterial sequences, Warcup (fungal ITS trainset 2) and UNITE (fungal ITS trainset 07/04/2014) options for the fungi sequences.

Overall, we classified 59482 bacterial, 472 archaeal, 82375 fungal and 3419 protist zOTUs.

Edaphic, topoclimatic, spatial and biotic variables

All preparations related to the predictors, calculations of niche properties and statistical analyses were performed in the R environment (Team, 2013), primarily using the phyloseq (McMurdie & Holmes, 2013) and vegan packages (Dixon, 2003), and visualised using ggplot2 (Wickham & Wickham, 2007). The variables used in this study are detailed in Table S1.

A total of 44 local soil properties, including but not limited to electrical conductivity (EC), pH, bulk soil water content (SoilWaterC), total phosphorus (total P), total organic carbon content (TOC), nitrogen (N), C/N ratio, major elements (such as SiO_2, MnO, MgO and CaO), mineralogical composition (such as phyllosilicates, quartz and calcite) and five classes of soil texture were measured in the laboratory. The methods and variables measured are thoroughly detailed in Yashiro *et al.* (2016) and Buri *et al.* (2020).

Topographic variables included the slope angle and topographic position (convexity-concavity) and were calculated from the digital elevation models at 25 m resolution as in Dubuis *et al.* (2013). 19 climatic variables are based on the MeteoSuisse database and named following the WorldClim variables (Fick & Hijmans, 2017). We also derived variables such as growing degree days, moisture index and solar radiation based on Dubuis *et al.* (2013) with a resolution of 25 m. The snow cover duration index was calculated based on satellite images as described in Panchard et al. (in preparation).

To create spatial predictors, the geographic coordinates of the sampling sites were transformed to geodetic cartesian (x,y) coordinates using the SoDA package in R (Chambers, 2008) and the Euclidean distances among the sites was calculated using vegan. These were used for distance-decay curves and Mantel tests as they provide a tangible measure of spatial distance. Distance-based Moran's Eigenvector Maps (MEM) were constructed with the (x,y) coordinates using the adespatial R package (Dray *et al.*, 2017) to summarize spatial structures present in the study area in a few proxy variables. This method creates a series of variables that correspond to spatial structures at all spatial scales contained within a given sampling design (Borcard *et al.*, 2018). These vectors were used as predictors in random forests and variation partitioning analyses. In this instance, 27 MEM positive eigenvectors were constructed, representing a sequence of broad to fine scale variation over the extent of the study site (Borcard *et al.*, 2018).

To derive the biotic predictors of both the alpha and beta diversities, we used six biotic variables representing the richness (ACE) and diversity (Shannon index) of each of the three groups other than the one under investigation. We also included plant richness, using data sampled at the same sites and described in Dubuis *et al.* (2013).

Altogether, a total of 109 variables consisting of 44 soil, 29 climatic, 2 topographic, 27 spatial and 7 biotic variables were available for the models to determine which are the most important variables explaining alpha and beta diversity of each of the four microbial groups.

Alpha diversity analyses

Alpha diversity was measured using the ACE richness estimator and the Shannon index, both calculated using phyloseq. Differences in richness between groups were assessed with ANOVA and compared using Tukey's Honest Significant Difference (HSD) tests. The major edaphic, topoclimatic, spatial and biotic factors associated with the alpha diversity of each microbial group was evaluated with random forest regression models (Breiman, 2001). First, to identify the most important predictors, multivariate models including all the variables were computed with 2000 trees. To obtain the simplest and best model, only

the variables presenting the highest increase mean squared error (%IncMSE) and a minimum %IncMSE value of 5 were conserved for further analyses. Spearman correlation coefficients among these conserved variables were calculated with the corrplot package in R (Wei *et al.*, 2017). The variable retained in case of a correlation coefficient ≥ |0.70| (Dormann *et al.*, 2013) was the one with the highest %IncMSE; thus, only uncorrelated variables were included in the subsequent random forest models. Finally, the remaining variables were dropped using backward selection until the percentage of variation explained by the models stopped increasing and the models could not be improved further. The maximum number of variables within each final model was 12. The final models were computed with the rfPermute function with 5000 repetitions and 2000 trees, in the rfPermute package (Archer & Archer, 2020), to assess the importance and significance of each variable on the richness (ACE) and diversity (Shannon index) of each microbial community. Additionally, inter-group alpha diversity relationships were evaluated using regression models.

Beta diversity analyses

To assess beta diversity of microbial communities, each zOTU table was transformed to relative abundance (Weiss *et al.*, 2017) and Bray-Curtis distances were calculated using the phyloseq and vegan R packages. To determine whether the difference in community composition among locations was spatially correlated, we computed distance-decay curves (Morlon *et al.*, 2008) using the (x,y) coordinates with the betapart package (Baselga & Orme, 2012). Then, the relationship of edaphic, topoclimatic, spatial and biotic factors with microbial beta diversity of each community were identified by two methods. First, using vegan, we computed Mantel tests between the explanatory distance matrices (topoclimatic, edaphic and spatial) and the Bray-Curtis dissimilarity matrix of each microbial group.

Second, we used distance-based redundancy analysis (dbRDA), combined with ANOVA and variation partitioning. To obtain the best model, forward and backward selection of variables was conducted, starting from intercept-only and all-variables models, respectively, with the ordiR2step function. Once the best model was determined, Spearman correlation coefficients among the remaining variables were calculated. When pairs of variables had a correlation coefficient ≥ |0.70|, only the first variable selected by the model was conserved. The remaining uncorrelated variables were included in a new dbRDA model and the variance inflation factor (VIF) was computed to further confirm that multicollinearity was low or absent (VIF <10) and the varpart function was used to assess the proportions of variation explained by each variable. The statistical significance of each environmental variable was tested with ANOVA and only significant variables ($p<0.05$) were plotted on the dbRDA ordination plot.

Finally, we evaluated the relative importance of stochastic processes using the modified stochasticity ratio (MST) (Ning *et al.*, 2019, Guo *et al.*, 2021). This metric estimates ecological stochasticity according to a null-model-based framework. The MST index value indicates the dominance of stochastic processes (> 50%) or deterministic processes (< 50%) on community assembly. The Bray–Curtis dissimilarity-based MST index was calculated for each microbial group using the NST R package (Ning *et al.*, 2019).

Environmental niche

The niche breadth was calculated as the standard deviation of each abiotic variable, weighted by the relative abundance of the zOTUs at each sampling site. The niche position was calculated as the mean of the variable, weighted by the relative abundance of the zOTUs at each sampling site. If more than one variable was investigated (e.g. PC1 and PC2), the values of niche breadth and position were averaged to obtain mean niche properties. The niche function ecospat.nichePOSNB used in this study is available on Github (https://github.com/ecospat/ecospat) and will be available in the ecospat v3.3 R package (Di Cola *et al.*, 2017).

To characterise the environmental niches of zOTUs, we first defined a PCA space based on the soil and topoclimatic variables measured across the 136 sampling sites. The first two axes of the PCA jointly explained 34.8 % of the total environmental variation in the study sites. We calculated the niche breadth and position of each zOTU within this environmental space. We also repeated this analysis along each environmental gradient of interest identified as influencing alpha and beta diversity. Quadratic regressions were performed and plotted as a rasterised density plot using the geom_bin_2d function.

Elevational gradient

As this study was conducted in the Alps, elevation was an important gradient to consider. However, it was not used as a variable in any of the previously mentioned analysis since it is not a direct causal driver for taxa, and it correlates already with the gradients of many other climatic and edaphic variables. Nevertheless, we independently evaluated the relationship between elevation and alpha or beta diversity by repeating the above analyses and calculated changes in niche characteristics along the elevation gradient.

Results

Predictors of alpha diversity

Bacterial richness (ACE) and diversity (Shannon index) were significantly higher than fungal, protist and archaeal richness and diversity (Table S2). Interestingly, while mean fungal richness was significantly higher than mean protist richness, the diversity of both groups was within the same range.

The random forest models revealed strong relationships between edaphic, biotic and topoclimatic variables and the richness and the diversity of each of the four microbial groups investigated [Fig.2]. In contrast, only six spatial vectors (MEM 1, 3, 7, 10, 14, 21) significantly explained alpha diversity, primarily for fungi and protists [Fig. 2]. Overall, the models best explained bacterial alpha diversity, with over 67% of the variance in both richness and diversity explained. The variance was also well explained for archaeal and fungal richness (>66%) but less for diversity (<45%). Finally, variation in protist richness (19%) and diversity (27%) were the least explained by the models.

The diversity and richness of each microbial group were related to a unique set of variables representing mainly edaphic and biotic properties. Specifically, soil pH (and highly correlated calcium oxide content

(CaO) (Spearman = 0.89)) were the most important edaphic variables driving richness and diversity of all microbial groups, except protist richness [Fig. 2]. Overall, bell-shaped curves were observed, with decrease in alpha-diversity towards both ends of the pH gradient [Fig. S1]. Soil water content (SoilWaterC) was also a key variable, explaining archaeal, fungal and protist richness and/or diversity [Fig. 2]. Other variables such as silica (SiO_2) were also important edaphic variables differentially correlated to the alpha diversity of various groups [Fig. 2]. The increase in silica was correlated with the increase in fungal richness as well as with the increases in archaeal, fungal and protists diversities [Fig. S2].

Out of topoclimatic variables, predictors related to winter conditions (snow cover duration (SCD) and number of freezing degree days (FDD)) were repeatedly identified as important in shaping alpha-diversity of all groups [Fig. 2]. In all the cases, alpha diversity was higher at the warm than at the cold extreme of the gradient [Fig. S3-4], a trend also reflected along the elevation gradient [Fig. S5].

Among the biotic variables, plant richness was identified as a key factor in random forest models [Fig. 2]. This was especially clear with fungal alpha diversity, where an increase in plant richness was related to an increase in fungal richness and diversity ($R^2 > 0.38$, $p < 0.001$) [Fig. S6]. Other biotic variables explained some of the variance, differentially for each group [Fig. 2]. These alpha diversity relationships among the groups using regression models are detailed in figures S7 and S8 and highlight the strongest positive correlation between bacterial and archaeal communities for both richness (ACE) and diversity (Shannon).

Predictors of beta diversity

The Mantel tests determined that the beta diversity of all microbial communities was more correlated to edaphic variables than to topoclimatic variables [Table 1]. However, topoclimatic properties better explained the beta diversity of fungal and protist communities than that of bacterial and archaeal communities. Geography (through the Euclidean distances of sampling sites) was also a significant factor correlated with fungal and protists communities [Table 1] and this relationship was further supported by distance-decay curves and power models [Fig. S9]. While the increase in community dissimilarity with increasing spatial distance was weak at the scale of the sampling area, these curves also highlighted the spatial turnover of these communities. Fungi presented the highest turnover with very dissimilar communities overall, followed by archaea, protists and bacteria [Fig. S9], as supported by the distance-decay curves (Ranjard *et al.*, 2013).

For beta diversity, the distance-based RDA identified pH, soil water content and electrical conductivity (EC) as key variables across all groups [Fig. 3]. Total phosphorus (TotalP) was also important for all groups except protists. Topoclimatic variables presented a lower degree of correlation, with snow cover duration (SCD) related to bacterial and archaeal community composition and temperature (gdd, bio5) related to fungal and protist communities. Spatial factors were confirmed to mostly have a limited relationship with microbial communities, as previously calculated by the Mantel tests. They were significant only for fungal and protist communities, with one and five significant vectors respectively, mainly across medium

to broad spatial scales [Fig. 3]. Biotic factors explained most of the variance after edaphic variables, with plant richness correlated with all microbial groups [Fig. 3]. We observed strong correlations in inter-group beta-diversity especially between bacterial-fungal, bacterial-archaeal and fungal-protists communities as supported also by the mantel tests results [Table 1]. Finally, all communities had MST values < 50% [Fig. 3], confirming the dominance of deterministic processes, even for fungal and protist communities that had a high variance remaining unexplained in the variation partitioning [Fig. 3].

The environmental niche

The environmental niche of microbial zOTUs in the PCA environmental space were primarily defined by topoclimatic variables (represented mainly by PC1), and secondarily by edaphic properties (represented by both PC1 and PC2) [Fig. S10]. PC1 explained 26.2 % of the variance while PC2 accounted for 8.6 % of the environmental variance among the sites. Differences in mean niche breadth were significant (ANOVA, $p < 0.001$) with protists presenting the largest average niche breadth, followed by bacteria, archaea and fungi [Fig. 4A]. Fungal and bacterial groups had longer tails of zOTUs with very large niche breadths, suggesting the presence of highly generalist taxa, able to survive in a wide range of environmental conditions. Differences in average niche position among the four groups were also significant (ANOVA, $p < 0.001$), although less visually evident [Fig. 4B]. The pairwise comparisons defined that fungi and protists had similar average niche positions (Tukey's HSD, $p = 0.77$) while all other comparisons were significantly different (Tukey's HSD, $p < 0.05$).

Assessing the relationships between niche breadth and position of each zOTU along selected environmental gradients (previously identified as important factors for alpha and beta diversity of each microbial group [Fig. 2, Fig. 3]), revealed that smaller niche breadths were, in general, recorded toward the environmental extremes of the gradients. This was especially clear along edaphic gradients such as pH, soil water content (SoilWaterC) and total phosphorus content (TotalP) [Fig. 5], for which the sampling area covered most of the possible variation in the gradients. The strongest relationships were observed for soil water content, where zOTUs representing organisms living in drier soil exhibited small niche breadth, which rapidly increased in range with increasing water content. Similarly, zOTUs living in environments with low phosphorus content exhibited smaller niche breadths. Interestingly, while niche breadth decreased for bacterial, archaeal and fungal zOTUs towards the pH extremes values, this was not observed for protists, which exhibited increasing niche breadths towards higher pH values. Soil electrical conductivity was also identified as a key variable for the microbial communities [Fig. S11], with increasing niche breadth associated to increasing conductivity. Similar patterns of decreasing niche breadths toward environmental extremes were less clear for climatic gradients [Fig. S11]. A decrease in niche breadth with increasing snow cover duration was observed for bacterial, archaeal and fungal zOTUs. In contrast, the niche breadth of protists increased with snow cover duration. No clear trends were recorded for the other climatic variables identified as important for the alpha and beta diversity.

Discussion

Environmental predictors of microbial diversity

Using tools traditionally applied to macro-organisms, we showed their applicability to detect spatial and environmental biogeographic trends also when studying microorganisms both at the level of individual taxonomic units (niches) and communities (alpha and beta diversity). Assessing alpha and beta diversity of four microbial groups within the same study area allowed us to identify and compare the key factors associated with each community. In all cases, deterministic processes drove community assembly (Ning *et al.*, 2019) and as hypothesised, the diversity of each microbial group was explained/governed by a different set of variables, but contrary to our hypothesis, alpha and beta diversity of all groups were mainly associated with edaphic properties. Indeed, while soil properties were expected to weight more on the bacterial and archaeal assemblies (Shi *et al.*, 2016, Delgado-Baquerizo *et al.*, 2018, Malard *et al.*, 2021, Starke *et al.*, 2021), we had hypothesised that topoclimatic variables would be more important for fungal and protist communities (Mod *et al.*, Bates *et al.*, 2013, Tedersoo *et al.*, 2014, Mod *et al.*, 2020, Oliverio *et al.*, 2020). However, fungi and protists had a high proportion of unexplained variance but the lowest MST values, suggesting that especially for these groups, some key variables structuring these communities were missing from the models, such as soil microclimate (Lembrechts *et al.*, 2019). Biotic variables were also key to the diversity of all groups, suggesting strong inter-kingdom interactions. Additional important biotic interactions with plants (Yashiro *et al.*, 2016), or with other soil organisms such as arthropods, rotifers or nematodes (Coleman & Wall, 2014, Singer *et al.*, 2020, van den Hoogen *et al.*, 2020) could also drive the diversity of these microbial groups. Finally, the MST index indicated that stochastic events such as ecological drift or diversification account for some of the unexplained variance (Zhou & Ning, 2017). Overall, these results highlight the importance of cross kingdom comparisons.

Microbial niche breadth

We hypothesised that fungi and protists would have smaller niche breadths than bacteria and archaea due to higher spatial turnover and lower ecological tolerance (Pikuta *et al.*, 2007, Taylor & Sinsabaugh, 2015, Belilla *et al.*, 2019, Fournier *et al.*, 2020, Kivlin & Hawkes, 2020). While fungi indeed presented the lowest niche breadths recorded and the highest spatial turnover, the environmental niche breadths of protists was, on average, larger than that of any other microbial group investigated. In soils, many bacteria, archaea and fungi form obligate associations with other taxa (consortia), a fact that is illustrated by the limited number of taxa that can be cultured axenically, as compared with the total diversity (Stewart, 2012, Pande & Kost, 2017). This means that the presence of each taxon depends on the occurrence of associated organisms nearby, as illustrated by co-occurrence networks (Barberán *et al.*, 2012, de Vries *et al.*, 2018, Shi *et al.*, 2019), thus likely reducing niche breadth to that of the whole consortium (Bar-Massada, 2015) or at least part of it. In contrast, most soil protists are heterotrophic (Singer *et al.*, 2020, Mazel *et al.*, 2021) and depend on the presence of prey or host , even though food specialization occurs to various degrees (Geisen *et al.*, 2015, Seppey *et al.*, 2017). Furthermore, both bacteria and archaea have similar cell size ranges, on average 1-3 orders of magnitude lower than the size of protists (Tesson *et al.*, 2015, Stepanauskas *et al.*, 2017), allowing more efficient

passive wind dispersal (Wilkinson *et al.*, 2012). As a result, spatial factors did not account for any of the variation in alpha and beta diversity of these two groups, suggesting that dispersal limitation only weakly structured these communities within the studied region. Therefore, their niche breadth might rather depend on their survival based on environmental conditions and biotic interactions, such as co-occurrence, rather than dispersal capabilities.

Niche specialisation

The few fungal and bacterial zOTUs with very large niche breadths indicate the presence of some highly generalist taxa, able to survive in many environmental conditions and likely with strong dispersal capabilities. These may include spore producing fungal and bacterial taxa increasing dispersal range and the available environmental space (Aguilar-Trigueros *et al.*, 2019). Unlike the three other groups, the distribution [Fig. 4] of fungal niche breadths showed a high abundance of taxa with very small niche breadths, likely reflecting highly specialised taxa. Theory suggests that generalist taxa are primarily influenced by dispersal processes leading to larger realized niche breadths, while specialists are primarily influenced by environmental filtering and biotic interactions, and accordingly present smaller realized niche breadths (Pandit *et al.*, 2009, Büchi & Vuilleumier, 2014). While we did not actively differentiate generalists from specialist taxa, this initial theory (Pandit *et al.*, 2009) would suggest, based on the average larger realized niche breadth, that protist, bacterial and archaeal communities harbor more generalist taxa than fungal communities. Therefore, the variance in fungal communities should be best explained by environmental variables. Instead, however, fungal and protist communities had the largest unexplained variance and edaphic properties played the most important role on the alpha and beta diversity of all groups. This is in accordance with other studies on microbial communities which have actually shown that biogeographic patterns among generalist microbial assemblages might be better explained by environmental conditions than for specialists, suggesting that environmental processes significantly impact all taxa and not only specialists (Székely & Langenheder, 2014, Lindh *et al.*, 2016, Hu *et al.*, 2019, Luo *et al.*, 2019, Malard *et al.*, 2019). Future studies could also explore how the environmental niche varies, within a same taxonomic group, between distinct functional groups (e.g. for protists, between consumers, parasites and phototrophs) as their distributions was shown to differ along environmental gradients (Mazel *et al.*, 2021).

Niche breadth and position along the gradients

The role of edaphic variables on all microbial groups was also reflected when assessing the relationships between niche breadth and position along the environmental gradients. As hypothesised, the niche breadth decreased towards environmental extremes, suggesting more specialised taxa towards environmental edges (Büchi & Vuilleumier, 2014). This trend was clearly observed when most of the range of the gradient was represented in the study area and was conserved across microbial groups. For instance, taxa with niche position close to either end of the pH gradient generally had a narrow niche, suggesting the increase in specialist taxa in basic and acidic soils. Only protists did not follow this trend, but the lack of study sites with pH over 9 might have obscured a change in pattern potentially further up.

Electrical conductivity, which reflects the salt concentration in the soil, also showed decreasing niche breadth with decreasing conductivity. However, the sampling sites only covered a small section of the salinity gradient given that no sites were recorded over 500 $\mu S.cm^{-1}$; as a general reference, electrical conductivity over 500 $\mu S.cm^{-1}$ represent relatively saline systems, while values can exceed 4000 $\mu S.cm^{-1}$ in certain soils (Hardie & Doyle, 2012). As salt concentration is an important variable for microbial communities (Logares *et al.*, 2009), we would expect the niche breadth to also decrease in hypersaline soils. Similarly, organisms living under growth-limiting low available water or nutrient (phosphorus) contents (Schmidt & Lipson, 2004, Darcy *et al.*, 2018) need adaptations to survive in these conditions (Torsvik & Øvreås, 2008, Malard & Pearce, 2018, Singer *et al.*, 2018), hence the lower niche breadths observed for taxa with niche position close to the lower ends of soil water content and nutrient gradients.

Interestingly, these decreasing trends in niche breadths towards the ends of the gradient were not as clear along climatic variables, likely because the study area only covered a portion of these gradients. Only snow cover duration was relatively well covered, with sites having short snow cover duration (1 day in average) and sites with up to 235 days of snow cover (out of 365 annually). The general trend observed, although quite weak, was also a decrease in niche breadth towards longer snow cover, suggesting more specialised taxa, as observed for ectomycorrhizal fungi and plants (Björk & Molau, 2007, Yao *et al.*, 2013). Furthermore, the recurrent detection of snow cover duration as an important climatic variable for alpha and beta diversity highlights the importance of considering the ecosystems being investigated (Edwards *et al.*, 2007). Snow cover duration is rarely considered in most studies on alpine microorganisms (exceptions include Zinger *et al.* (2009) and Xia *et al.* (2014)), despite its ecologically important role in these ecosystems (Edwards *et al.*, 2007). For all other climatic variables, as well as elevation, only partial gradients were covered and niche breadth trends may thus not fully reflect the reality in all parts of the gradients due to niche truncation, resulting from the limited range of environmental conditions existing in a study region (Regos *et al.*, 2019), highlighting the need for global scale analyses.

Perspectives

Quantifying and comparing the niche properties of microorganisms has a wide range of potential uses, especially to model spatial distribution and evaluate the potential influence of climate and environmental change on taxa and communities (Thuiller *et al.*, 2005, Muller, 2019, Finn *et al.*, 2020, Mod *et al.*, 2021). Specialised taxa with limited niche breadth may be at greater risk of extinction as they benefit from homogeneous environments (in space and/or time). Devictor *et al.* (2008) already highlighted some overwhelming evidence of specialist declines among macroscopic organisms in fragmented landscapes. However, whether similar trends await microbial taxa remains undetermined, especially as speciation and diversification occur at much faster rates than in macro-organisms, potentially allowing faster acclimation and adaptation to changing environmental conditions (Cavicchioli *et al.*, 2019). The theory of functional redundancy in microbial communities (Rousk *et al.*, 2009, Louca *et al.*, 2018) might also mitigate the impact of widespread extinctions following environmental change. However, emerging

evidence suggests that functional redundancy may not actually be as widespread as previously thought (Allison & Martiny, 2008, Delgado-Baquerizo *et al.*, 2016, Galand *et al.*, 2018). Therefore, the extinction of specialist taxa harbouring rare and/or key functions for an ecosystem (Mariadassou *et al.*, 2015) could disrupt ecosystem functioning (Jousset *et al.*, 2017), especially in fast changing ecosystems such as alpine regions.

Declarations

Ethics approval

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All zOTU tables and associated metadata have been deposited on Figshare (https://figshare.com/account/home#/projects/97421).

The niche function ecospat.nichePOSNB used in this study is available on Github (https://github.com/ecospat/ecospat) and will be available in the ecospat v3.3 R package (Di Cola *et al.*, 2017).

Competing interests

The authors declare that they have no competing interests

Funding

This study was supported by the Swiss national science foundation (SNSF, grant PDFMP3_135129 to AG and HNH and grant 315230_184908 to AG). EL was supported by an “Atracción de talentos” grant from the Community of Madrid, project 2017-T1/AMB-5210 and by the project MYXOTROPIC VI (PGC2018-094660-B-I00) awarded by the Spanish Government.

Authors’ contributions

AG and HNH designed the initial project leading to the sampling of the data used here. AG, EY and HNH planned the initial sampling, and EY led the laboratory work for conditioning the field samples and sequencing for bacteria, archaea and fungi. EL and EADM designed and conducted the laboratory work and sequencing for protists. NG and EY conducted the bioinformatic processing while HM compiled and prepared the datasets, in coordination with AG. AG conceived the present project based on the data, with help from HM, NG and EY. LAM and AG designed the present comparative study. LAM conducted the

statistical analyses, with help from OB, and wrote the manuscript. All authors revised and approved the final version.

Acknowledgements

We thank Aline Buri, Eric Pinto, Christophe Seppey, Florent Mazel, Jan van der Meer and Ian Sanders for their help with previous steps that contributed to generate the data used here, and for interesting discussions and feedbacks.

References

Aguilar-Trigueros CA, Hempel S, Powell JR, Cornwell WK & Rillig MC (2019) Bridging reproductive and microbial ecology: a case study in arbuscular mycorrhizal fungi. *The ISME journal* **13**: 873-884.

Allison SD & Martiny JB (2008) Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences* **105**: 11512-11519.

Archer E & Archer ME (2020) Package 'rfPermute'. p.^pp.

Bar-Massada A (2015) Complex relationships between species niches and environmental heterogeneity affect species co-occurrence patterns in modelled and real communities. *Proceedings of the Royal Society B: Biological Sciences* **282**: 20150927.

Barberán A, Bates ST, Casamayor EO & Fierer N (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME journal* **6**: 343.

Baselga A & Orme CDL (2012) betapart: an R package for the study of beta diversity. *Methods in ecology and evolution* **3**: 808-812.

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

Belilla J, Moreira D, Jardillier L, Reboul G, Benzerara K, López-García JM, Bertolino P, López-Archilla AI & López-García P (2019) Hyperdiverse archaea near life limits at the polyextreme geothermal Dallol area. *Nature ecology & evolution* **3**: 1552-1561.

Bharti V, Dotaniya M, Shukla S & Yadav V (2017) Managing soil fertility through microbes: prospects, challenges and future strategies. *Agro-environmental sustainability,*p.^pp. 81-111. Springer.

Binzer A, Brose U, Curtsdotter A, Eklöf A, Rall BC, Riede JO & de Castro F (2011) The susceptibility of species to extinctions in model communities. *Basic and Applied Ecology* **12**: 590-599.

Björk RG & Molau U (2007) Ecology of alpine snowbeds and the impact of global change. *Arctic, Antarctic, and Alpine Research* **39**: 34-43.

Borcard D, Gillet F & Legendre P (2018) *Numerical ecology with R*. Springer.

Botkin DB, Saxe H, Araujo MB, Betts R, Bradshaw RH, Cedhagen T, Chesson P, Dawson TP, Etterson JR & Faith DPJB (2007) Forecasting the effects of global warming on biodiversity. **57**: 227-236.

Breiman L (2001) Random forests. *Machine learning* **45**: 5-32.

Brown SP, Veach AM, Rigdon-Huss AR, Grond K, Lickteig SK, Lothamer K, Oliver AK & Jumpponen A (2015) Scraping the bottom of the barrel: are rare high throughput sequences artifacts? *Fungal Ecology* **13**: 221-225.

Büchi L & Vuilleumier S (2014) Coexistence of specialist and generalist species is shaped by dispersal and environmental factors. *The American Naturalist* **183**: 612-624.

Buri A, Grand S, Yashiro E, Adatte T, Spangenberg JE, Pinto-Figueroa E, Verrecchia E & Guisan A (2020) What are the most crucial soil variables for predicting the distribution of mountain plant species? A comprehensive study in the Swiss Alps. *Journal of Biogeography* **47**: 1143-1153.

Cabral JS & Kreft H (2012) Linking ecological niche, community ecology and biogeography: insights from a mechanistic niche model. *Journal of Biogeography* **39**: 2212-2224.

Caron DA (2009) Past President's Address: Protistan Biogeography: Why All The Fuss? 1. *Journal of Eukaryotic Microbiology* **56**: 105-112.

Cavicchioli R, Ripple WJ, Timmis KN, Azam F, Bakken LR, Baylis M, Behrenfeld MJ, Boetius A, Boyd PW & Classen AT (2019) Scientists' warning to humanity: microorganisms and climate change. *Nature Reviews Microbiology* **17**: 569-586.

Chambers J (2008) *Software for data analysis: programming with R*. Springer Science & Business Media.

Cianfrani C, Broennimann O, Loy A & Guisan A (2018) More than range exposure: Global otter vulnerability to climate change. *Biological Conservation* **221**: 103-113.

Clare EL, Chain FJ, Littlefair JE & Cristescu ME (2016) The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data. *Genome* **59**: 981-990.

Coleman DC & Wall DH (2014) Soil fauna: Occurence, biodiversity. and roles in ecosystem function. *Soil microbiology, ecology and biochemistry,*(Paul EA, ed.) p.^pp. Academic Press.

Costa GC, Nogueira C, Machado RB & Colli GR (2010) Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. *Biodiversity Conservation* **19**: 883-899.

Crowther TW, Van den Hoogen J, Wan J, Mayes MA, Keiser A, Mo L, Averill C & Maynard DS (2019) The global soil community and its influence on biogeochemistry. *Science* **365**: eaav0550.

D'Amen M, Rahbek C, Zimmermann NE & Guisan A (2017) Spatial predictions at the community level: from current approaches to future frameworks. *Biological Reviews* **92**: 169-187.

Darcy JL, Schmidt SK, Knelman JE, Cleveland CC, Castle SC & Nemergut DR (2018) Phosphorus, not nitrogen, limits plants and microbial primary producers following glacial retreat. *Science advances* **4**: eaaq0942.

Dassen S, Cortois R, Martens H, de Hollander M, Kowalchuk GA, van der Putten WH & De Deyn GB (2017) Differential responses of soil bacteria, fungi, archaea and protists to plant species richness and plant functional group identity. *Molecular Ecology* **26**: 4085-4098.

de Vries FT, Griffiths RI, Bailey M, Craig H, Girlanda M, Gweon HS, Hallin S, Kaisermann A, Keith AM & Kretzschmar M (2018) Soil bacterial networks are less stable under drought than fungal networks. *Nature Communications* **9**: 1-12.

Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett RD, Maestre FT, Singh BK & Fierer N (2018) A global atlas of the dominant bacteria found in soil. *Science* **359**: 320-325.

Delgado-Baquerizo M, Giaramida L, Reich PB, Khachane AN, Hamonts K, Edwards C, Lawton LA & Singh BK (2016) Lack of functional redundancy in the relationship between microbial diversity and ecosystem functioning. *Journal of Ecology* **104**: 936-946.

Devictor V, Julliard R & Jiguet F (2008) Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. *Oikos* **117**: 507-514.

Di Cola V, Broennimann O, Petitpierre B, Breiner FT, d'Amen M, Randin C, Engler R, Pottier J, Pio D & Dubuis A (2017) ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography* **40**: 774-787.

Dixon P (2003) VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14**: 927-930.

Dolédec S, Chessel D & Gimaret-Carpentier C (2000) Niche separation in community analysis: a new method. *Ecology* **81**: 2914-2927.

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B & Leitão PJ (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**: 27-46.

Dray S, Blanchet G, Borcard D, Clappe S, Guenard G, Jombart T, Larocque G, Legendre P, Madi N & Wagner H (2017) Adespatial: multivariate multiscale spatial analysis. R package version 0.0-9. p.^pp.

Dubuis A, Giovanettina S, Pellissier L, Pottier J, Vittoz P & Guisan A (2013) Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables. *Journal of Vegetation Science* **24**: 593-606.

Edgar RC (2018) Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* **34**: 2371-2375.

Edwards AC, Scalenghe R & Freppaz M (2007) Changes in the seasonal snow cover of alpine regions and its effect on soil processes: a review. *Quaternary international* **162**: 172-181.

Falkowski PG, Fenchel T & Delong EF (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**: 1034-1039.

Fick SE & Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology* **37**: 4302-4315.

Finn D, Yu J, Ilhan Z, Fernandes V, Penton C, Krajmalnik-Brown R, Garcia-Pichel F & Vogel T (2020) MicroNiche: an R package for assessing microbial niche breadth and overlap from amplicon sequencing data. *FEMS microbiology ecology* **96**: fiaa131.

Fournier B, Samaritani E, Frey B, Seppey CV, Lara E, Heger TJ & Mitchell EA (2020) Higher spatial than seasonal variation in floodplain soil eukaryotic microbial communities. *Soil Biology and Biochemistry* **147**: 107842.

Galand PE, Pereira O, Hochart C, Auguet JC & Debroas D (2018) A strong link between marine microbial community composition and function challenges the idea of functional redundancy. *The ISME journal* **12**: 2470-2478.

Gaston K & Blackburn T (2008) *Pattern and process in macroecology*. John Wiley & Sons.

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

Geisen S, Laros I, Vizcaíno A, Bonkowski M & De Groot G (2015) Not all are free-living: High-throughput DNA metabarcoding reveals a diverse community of protists parasitizing soil metazoa. *Molecular Ecology* **24**: 4556-4569.

George PB, Lallias D, Creer S, Seaton FM, Kenny JG, Eccles RM, Griffiths RI, Lebron I, Emmett BA & Robinson DA (2019) Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. *Nature Communications* **10**: 1-11.

Guisan A, Thuiller W & Zimmermann NE (2017) *Habitat suitability and distribution models: with applications in R*. Cambridge University Press.

Guisan A, Petitpierre B, Broennimann O, Daehler C & Kueffer C (2014) Unifying niche shift studies: insights from biological invasions. *Trends in ecology & evolution* **29**: 260-269.

Guo J, Ling N, Li Y, Li K, Ning H, Shen Q, Guo S & Vandenkoornhuyse P (2021) Seed-borne, endospheric and rhizospheric core microbiota as predictors of plant functional traits across rice cultivars are dominated by deterministic processes. *New Phytologist* **230**: 2047-2060.

Hardie M & Doyle R (2012) Measuring soil salinity. *Plant salt tolerance*,p.^pp. 415-425. Springer.

Hirzel A & Guisan A (2002) Which is the optimal sampling strategy for habitat suitability modelling. *Ecological modelling* **157**: 331-341.

Holt RD (2009) Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences* **106**: 19659-19665.

Hu A, Wang H, Cao M, Rashid A, Li M & Yu C-P (2019) Environmental Filtering Drives the Assembly of Habitat Generalists and Specialists in the Coastal Sand Microbial Communities of Southern China. *Microorganisms* **7**: 598.

Hutchinson GE (1957) The multivariate niche. Vol. 22 p.^pp. 415-421.

Jiao S, Xu Y, Zhang J & Lu Y (2019) Environmental filtering drives distinct continental atlases of soil archaea between dryland and wetland agricultural ecosystems. *Microbiome* **7**: 1-13.

Jousset A, Bienhold C, Chatzinotas A, Gallien L, Gobet A, Kurm V, Küsel K, Rillig MC, Rivett DW & Salles JF (2017) Where less may be more: how the rare biosphere pulls ecosystems strings. *The ISME journal* **11**: 853.

Kearney M & Porter W (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology letters* **12**: 334-350.

Kivlin SN & Hawkes CV (2020) Spatial and temporal turnover of soil microbial communities is not linked to function in a primary tropical forest. *Ecology* **101**: e02985.

La Sorte FA & Jetz W (2010) Projected range contractions of montane biodiversity under global warming. *Proceedings of the Royal Society B: Biological Sciences* **277**: 3401-3410.

Lazarevic V, Whiteson K, Huse S, Hernandez D, Farinelli L, Østerås M, Schrenzel J & François P (2009) Metagenomic study of the oral microbiota by Illumina high-throughput sequencing. *Journal of microbiological methods* **79**: 266-271.

Lembrechts JJ, Nijs I & Lenoir J (2019) Incorporating microclimate into species distribution models. *Ecography* **42**: 1267-1279.

Lindh MV, Sjöstedt J, Casini M, Andersson A, Legrand C & Pinhassi J (2016) Local environmental conditions shape generalist but not specialist components of microbial metacommunities in the Baltic Sea. *Frontiers in Microbiology* **7**: 2078.

Logares R, Bråte J, Bertilsson S, Clasen JL, Shalchian-Tabrizi K & Rengefors K (2009) Infrequent marine–freshwater transitions in the microbial world. *Trends in microbiology* **17**: 414-422.

Lortie CJ, Brooker RW, Choler P, Kikvidze Z, Michalet R, Pugnaire FI & Callaway RM (2004) Rethinking plant community theory. *Oikos* **107**: 433-438.

Louca S, Polz MF, Mazel F, Albright MB, Huber JA, O'Connor MI, Ackermann M, Hahn AS, Srivastava DS & Crowe SA (2018) Function and functional redundancy in microbial systems. *Nature ecology & evolution* **2**: 936-943.

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**: 1-11.

Luo Z, Liu J, Zhao P, Jia T, Li C & Chai B (2019) Biogeographic patterns and assembly mechanisms of bacterial communities differ between habitat generalists and specialists across elevational gradients. *Frontiers in microbiology* **10**: 169.

Malard L (2019) *Microbial biogeography and colonisation in Arctic terrestrial ecosystems*. University of Northumbria at Newcastle (United Kingdom).

Malard LA & Pearce DA (2018) Microbial diversity and biogeography in Arctic soils. *Environmental microbiology reports* **10**: 611-625.

Malard LA, Anwar MZ, Jacobsen CS & Pearce DA (2019) Biogeographical patterns in soil bacterial communities across the Arctic region. *FEMS microbiology ecology* **95**.

Malard LA, Anwar MZ, Jacobsen CS & Pearce DA (2021) Influence of Spatial Scale on Structure of Soil Bacterial Communities across an Arctic Landscape. *Applied and Environmental Microbiology* **87**.

Mariadassou M, Pichon S & Ebert D (2015) Microbial ecosystems are dominated by specialist taxa. *Ecology Letters* **18**: 974-982.

Mazel F, Malard L, Niculita-Hirzel H*, et al.* (2021) Soil protist function varies with elevation in the Swiss Alps. *Environmental Microbiology*.

McMurdie PJ & Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one* **8**.

Mod HK, Scherrer D, Di Cola V, Broennimann O, Blandenier Q, Breiner FT, Buri A, Goudet J, Guex N & Lara EJGCB Greater topoclimatic control of above-Versus belowground communities.

Mod HK, Buri A, Yashiro E, Guex N, Malard L, Pinto-Figueroa E, Pagni M, Niculita-Hirzel H, van der Meer JR & Guisan A (2021) Predicting spatial patterns of soil bacteria under current and future environmental conditions. *The ISME Journal* 1-14.

Mod HK, Scherrer D, Di Cola V*, et al.* (2020) Greater topoclimatic control of above-versus below-ground communities. *Global change biology* **26**: 6715.

Morlon H, Chuyong G, Condit R, Hubbell S, Kenfack D, Thomas D, Valencia R & Green JL (2008) A general framework for the distance–decay of similarity in ecological communities. *Ecology letters* **11**: 904-917.

Muller EEJM (2019) Determining microbial niche breadth in the environment for better ecosystem fate predictions. **4**.

Ning D, Deng Y, Tiedje JM & Zhou J (2019) A general framework for quantitatively assessing ecological stochasticity. *Proceedings of the National Academy of Sciences* **116**: 16892-16898.

Okie JG, Van Horn DJ, Storch D, Barrett JE, Gooseff MN, Kopsova L & Takacs-Vesbach CD (2015) Niche and metabolic principles explain patterns of diversity and distribution: theory and a case study with soil bacterial communities. *Proceedings of the Royal Society B: Biological Sciences* **282**: 20142630.

Oliverio AM, Geisen S, Delgado-Baquerizo M, Maestre FT, Turner BL & Fierer N (2020) The global-scale distributions of soil protists and their contributions to belowground systems. *Science advances* **6**: eaax8787.

Pagel J & Schurr FM (2012) Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography* **21**: 293-304.

Pande S & Kost C (2017) Bacterial unculturability and the formation of intercellular metabolic networks. *Trends in microbiology* **25**: 349-361.

Pandit SN, Kolasa J & Cottenie K (2009) Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology* **90**: 2253-2262.

Pikuta EV, Hoover RB & Tang J (2007) Microbial extremophiles at the limits of life. *Critical reviews in microbiology* **33**: 183-209.

Pinto-Figueroa EA, Seddon E, Yashiro E, Buri A, Niculita-Hirzel H, Van Der Meer JR & Guisan A (2019) Archaeorhizomycetes spatial distribution in soils along wide elevational and environmental gradients reveal co-abundance patterns with other fungal saprobes and potential weathering capacities. *Frontiers in microbiology* **10**: 656.

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

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J & Glöckner FO (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* **41**: D590-D596.

Ranjard L, Dequiedt S, Prévost-Bouré NC, Thioulouse J, Saby N, Lelievre M, Maron P, Morin F, Bispo A & Jolivet C (2013) Turnover of soil bacterial diversity driven by wide-scale environmental heterogeneity. *Nature communications* **4**: 1-10.

Regos A, Gagne L, Alcaraz-Segura D, Honrado JP & Domínguez J (2019) Effects of species traits and environmental predictors on performance and transferability of ecological niche models. *Scientific reports* **9**: 1-14.

Rocha MP, Bini LM, Siqueira T, Hjort J, Grönroos M, Lindholm M, Karjalainen S-M & Heino J (2018) Predicting occupancy and abundance by niche position, niche breadth and body size in stream organisms. *Oecologia* **186**: 205-216.

Rousk J, Brookes PC & Bååth E (2009) Contrasting soil pH effects on fungal and bacterial growth suggest functional redundancy in carbon mineralization. *Applied and Environmental Microbiology* **75**: 1589-1596.

Schmidt P-A, Bálint M, Greshake B, Bandow C, Römbke J & Schmitt I (2013) Illumina metabarcoding of a soil fungal community. *Soil Biology and Biochemistry* **65**: 128-132.

Schmidt S & Lipson D (2004) Microbial growth under the snow: implications for nutrient and allelochemical availability in temperate soils. *Plant and Soil* **259**: 1-7.

Schwartz MW (2012) Using niche models with climate projections to inform conservation management decisions. *Biological Conservation* **155**: 149-156.

Seppey CV, Singer D, Dumack K, Fournier B, Belbahri L, Mitchell EA & Lara E (2017) Distribution patterns of soil microbial eukaryotes suggests widespread algivory by phagotrophic protists as an alternative pathway for nutrient cycling. *Soil Biology and Biochemistry* **112**: 68-76.

Seppey CV, Broennimann O, Buri A, Yashiro E, Pinto-Figueroa E, Singer D, Blandenier Q, Mitchell EA, Niculita-Hirzel H & Guisan A (2020) Soil protist diversity in the Swiss western Alps is better predicted by topo-climatic than by edaphic variables. *Journal of Biogeography* **47**: 866-878.

Shi Y, Fan K, Li Y, Yang T, He JS & Chu H (2019) Archaea Enhance the Robustness of Microbial Co-occurrence Networks in Tibetan Plateau Soils. *Soil Science Society of America Journal* **83**: 1093-1099.

Shi Y, Adams JM, Ni Y, Yang T, Jing X, Chen L, He J-S & Chu H (2016) The biogeography of soil archaeal communities on the eastern Tibetan Plateau. *Scientific reports* **6**: 38893.

Singer D, Duckert C, Heděnec P, Lara E, Hiltbrunner E & Mitchell EA (2020) High-throughput sequencing of litter and moss eDNA reveals a positive correlation between the diversity of Apicomplexa and their invertebrate hosts across alpine habitats. *Soil Biology and Biochemistry* **147**: 107837.

Singer D, Kosakyan A, Seppey CV, Pillonel A, Fernández LD, Fontaneto D, Mitchell EA & Lara E (2018) Environmental filtering and phylogenetic clustering correlate with the distribution patterns of cryptic protist species. *Ecology* **99**: 904-914.

Singer D, Seppey CV, Lentendu G, Dunthorn M, Bass D, Belbahri L, Blandenier Q, Debroas D, de Groot GA & De Vargas C (2020) Protist taxonomic and functional diversity in soil, freshwater and marine ecosystems. *Environment International* **146**: 106262.

Slatyer RA, Hirst M & Sexton JP (2013) Niche breadth predicts geographical range size: a general ecological pattern. *Ecology letters* **16**: 1104-1114.

Soininen J & Heino J (2007) Variation in niche parameters along the diversity gradient of unicellular eukaryote assemblages. *Protist* **158**: 181-191.

Starke R, Siles JA, Fernandes MLP, Schallert K, Benndorf D, Plaza C, Jehmlich N, Delgado-Baquerizo M & Bastida F (2021) The structure and function of soil archaea across biomes. *Journal of Proteomics* **237**: 104147.

Stepanauskas R, Fergusson EA, Brown J, Poulton NJ, Tupper B, Labonté JM, Becraft ED, Brown JM, Pachiadaki MG & Povilaitis T (2017) Improved genome recovery and integrated cell-size analyses of individual uncultured microbial cells and viral particles. *Nature communications* **8**: 1-10.

Stewart EJ (2012) Growing unculturable bacteria. *Journal of bacteriology* **194**: 4151-4160.

Stoeck T, Bass D, Nebel M, Christen R, Jones MD, BREINER HW & Richards TA (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular ecology* **19**: 21-31.

Székely AJ & Langenheder S (2014) The importance of species sorting differs between habitat generalists and specialists in bacterial communities. *FEMS microbiology ecology* **87**: 102-112.

Taylor DL & Sinsabaugh RL (2015) The soil fungi: occurrence, phylogeny, and ecology. *Soil microbiology, ecology, and biochemistry* 77-109.

Team RC (2013) R: A language and environment for statistical computing. 2013, Vienna, Austria: R Foundation for Statistical Computing. *URL http://www R-project org*.

Tedersoo L, Bahram M, Põlme S, Kõljalg U, Yorou NS, Wijesundera R, Ruiz LV, Vasco-Palacios AM, Thu PQ & Suija A (2014) Global diversity and geography of soil fungi. *Science* **346**: 1256688.

Tesson SV, Okamura B, Dudaniec RY, Vyverman W, Löndahl J, Rushing C, Valentini A & Green AJ (2015) Integrating microorganism and macroorganism dispersal: modes, techniques and challenges with particular focus on co-dispersal. *Ecoscience* **22**: 109-124.

Thuiller W, Lavorel S & Araújo MB (2005) Niche properties and geographical extent as predictors of species sensitivity to climate change. *Global Ecology Biogeography* **14**: 347-357.

Torsvik V & Øvreås L (2008) Microbial diversity, life strategies, and adaptation to life in extreme soils. *Microbiology of extreme soils*,p.^pp. 15-43. Springer.

van den Hoogen J, Geisen S, Wall DH, Wardle DA, Traunspurger W, de Goede RG, Adams BJ, Ahmad W, Ferris H & Bardgett RD (2020) A global database of soil nematode abundance and functional group composition. *Scientific data* **7**: 1-8.

Wang Q, Garrity GM, Tiedje JM & Cole JR (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology* **73**: 5261-5267.

Wei T, Simko V, Levy M, Xie Y, Jin Y & Zemla J (2017) Package 'corrplot'. *Statistician* **56**: 316-324.

Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y & Birmingham A (2017) Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5**: 27.

White TJ, Bruns T, Lee S & Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications* **18**: 315-322.

Wickham H & Wickham MH (2007) The ggplot package. p.^pp.

Wiens JA, Stralberg D, Jongsomjit D, Howell CA & Snyder MA (2009) Niches, models, and climate change: assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences* **106**: 19729-19736.

Wiens JJ (2011) The niche, biogeography and species interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**: 2336-2350.

Wilkinson DM, Koumoutsaris S, Mitchell EA & Bey I (2012) Modelling the effect of size on the aerial dispersal of microorganisms. *Journal of biogeography* **39**: 89-97.

Xia H, Peng Y, Yan W & Ning W (2014) Effect of snow depth and snow duration on soil N dynamics and microbial activity in the alpine areas of the eastern Tibetan Plateau. *Russian journal of ecology* **45**: 263-268.

Yao F, Vik U, Brysting AK, Carlsen T, Halvorsen R & Kauserud H (2013) Substantial compositional turnover of fungal communities in an alpine ridge-to-snowbed gradient. *Molecular ecology* **22**: 5040-5052.

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

Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, Schweer T, Peplies J, Ludwig W & Glöckner FO (2014) The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic acids research* **42**: D643-D648.

Yuan H, Ge T, Chen C, O'Donnell AG & Wu J (2012) Significant role for microbial autotrophy in the sequestration of soil carbon. *Applied Environmental Microbiology* **78**: 2328-2336.

Zhou J & Ning D (2017) Stochastic community assembly: does it matter in microbial ecology? *Microbiology Molecular Biology Reviews* **81**: e00002-00017.

Zinger L, Shahnavaz B, Baptist F, Geremia RA & Choler P (2009) Microbial diversity in alpine tundra soils correlates with snow cover dynamics. *The ISME journal* **3**: 850-859.

Table

Table 1: Summary of significant Mantel test results investigating the correlation of each predictor matrix and microbial community structures (Bray-Curtis matrices).

	Bacteria		Archaea		Fungi		Protists	
Mantel statistics	R	p value	R	p value	R	p value	R	p value
Edaphic	0.39	1.10^{-4}	0.20	1.10^{-4}	0.53	1.10^{-4}	0.43	1.10^{-4}
Topoclimatic	0.17	1.10^{-4}	0.09	1.10^{-4}	0.34	1.10^{-4}	0.30	1.10^{-4}
Spatial (x,y)	0.08	0.006	0.06	0.03	0.12	0.003	0.12	0.003
Bacteria			0.73	1.10^{-4}	0.78	1.10^{-4}	0.57	1.10^{-4}
Archaea	0.73	1.10^{-4}			0.54	1.10^{-4}	0.36	1.10^{-4}
Fungi	0.78	1.10^{-4}	0.54	1.10^{-4}			0.64	1.10^{-4}
Protists	0.57	1.10^{-4}	0.36	1.10^{-4}	0.64	1.10^{-4}		

Figures

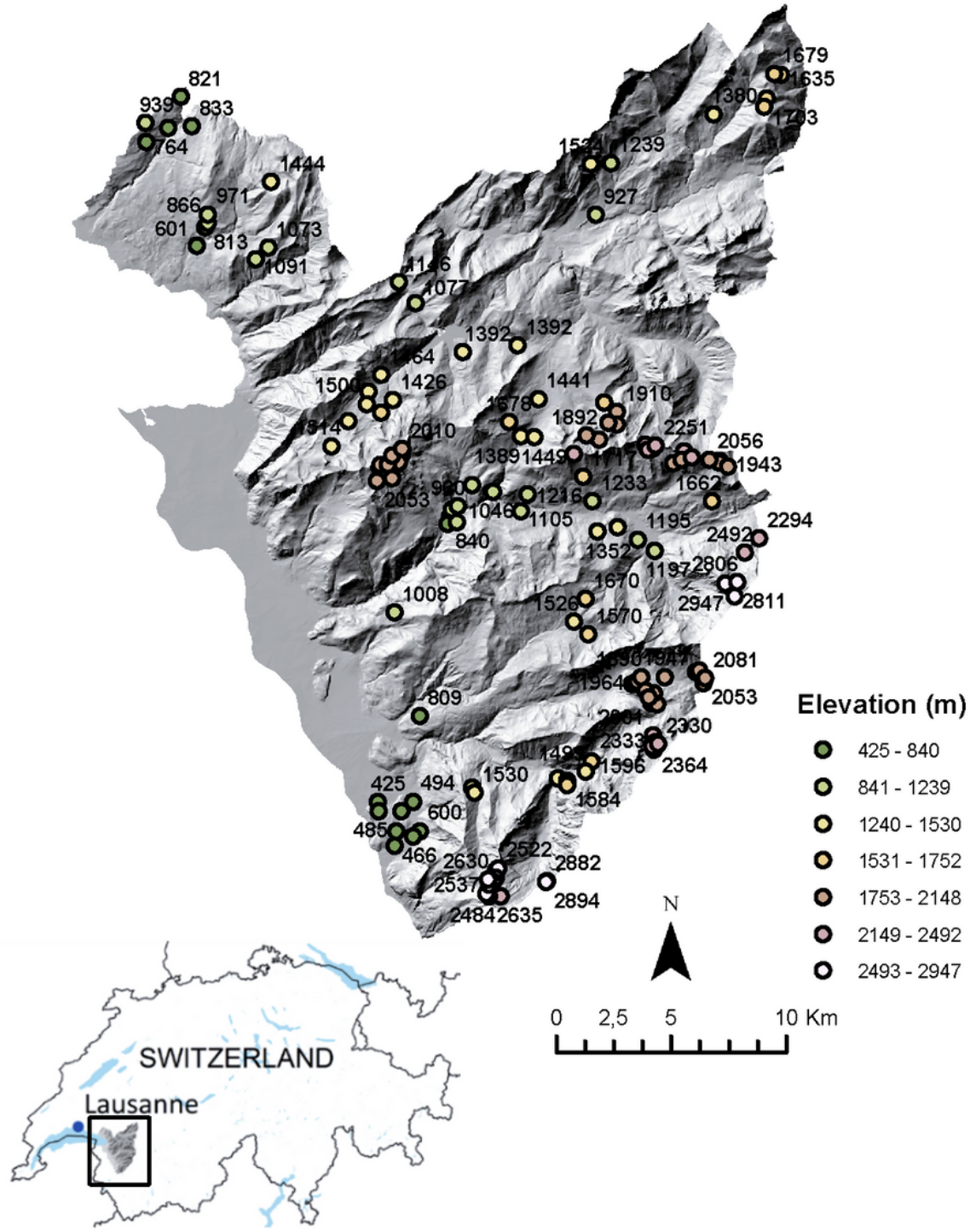


Figure 1

Locations and elevations of the 136 study sites in the western Swiss Alps across 700 km2 from which the bacterial, archaeal, fungal and protists soil communities were determined.

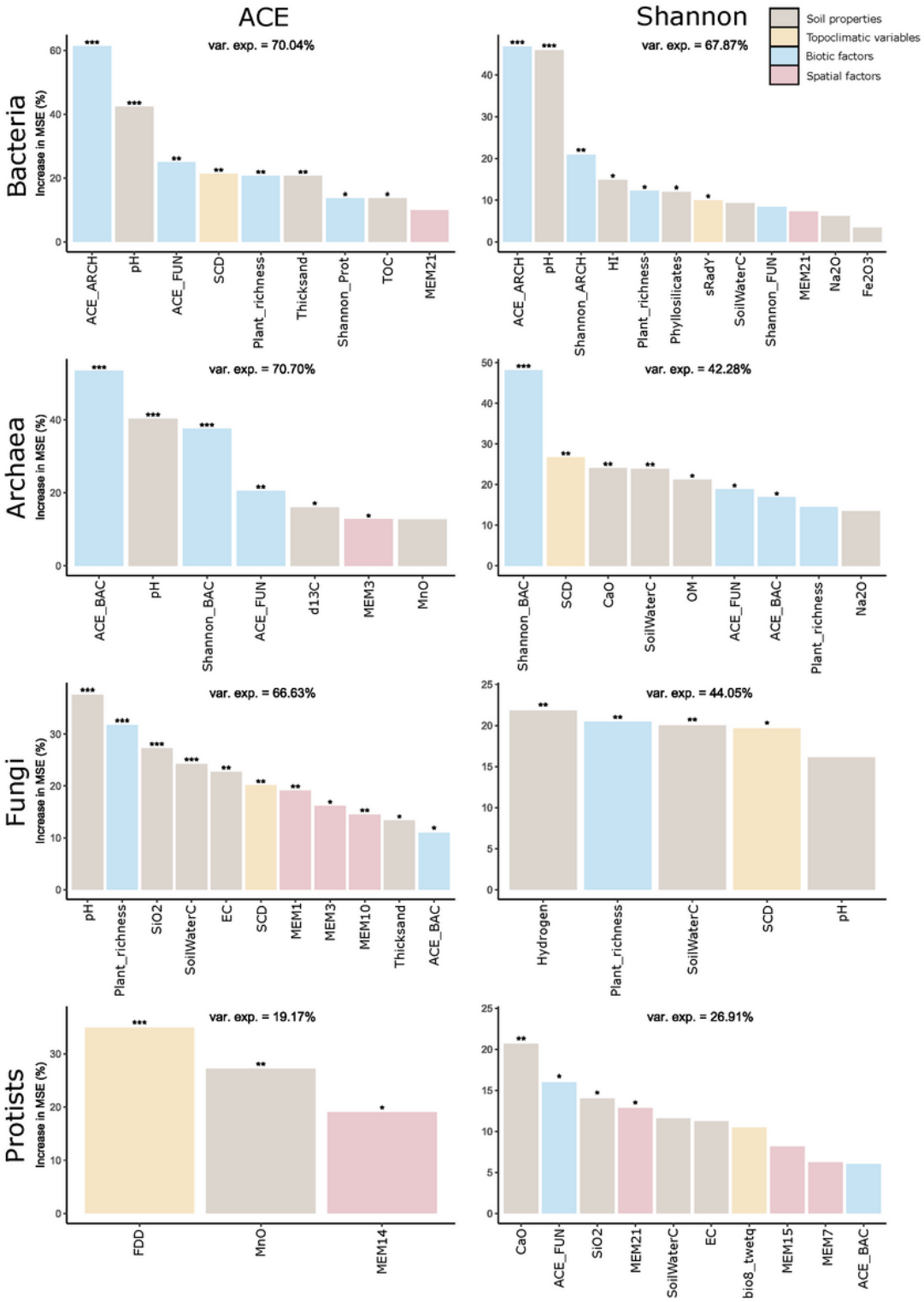
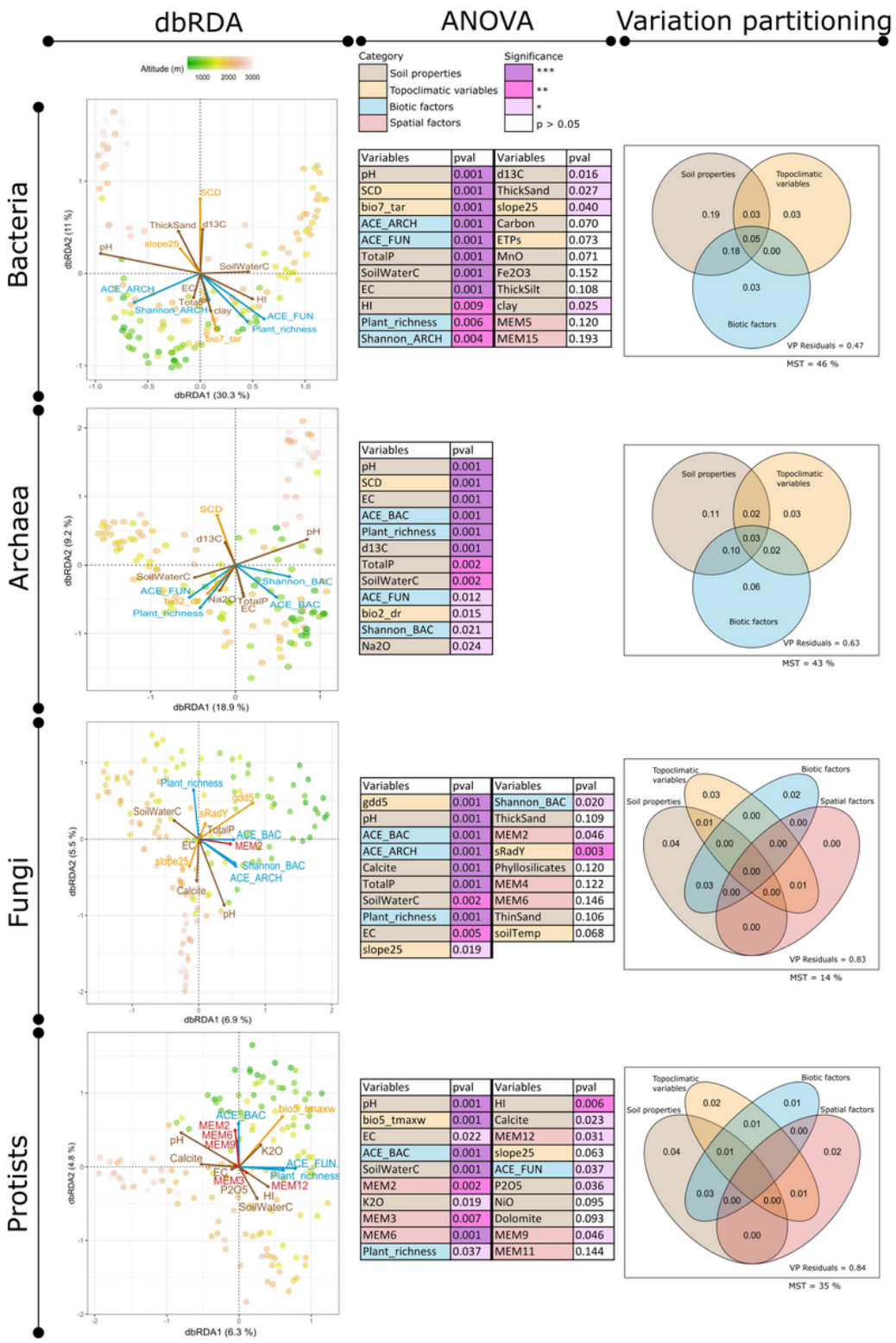


Figure 2

Mean predictor importance (% of increase in mean square error) of variables on microbial alpha-diversity (Shannon) and richness (ACE) predicted by the best random forest models. The accuracy was computed for each tree and averaged over the forest (2000 trees). The variance explained (var. exp.) is indicated for each model. Significance levels of each variable included in the model are as follows: *P < 0.05 and ** P < 0.01 and *** P < 0.001.



Variables	pval	Variables	pval
pH	0.001	d13C	0.016
SCD	0.001	ThickSand	0.027
bio7_tar	0.001	slope25	0.040
ACE_ARCH	0.001	Carbon	0.070
ACE_FUN	0.001	ETPs	0.073
TotalP	0.001	MnO	0.071
SoilWaterC	0.001	Fe2O3	0.152
EC	0.001	ThickSilt	0.108
HI	0.009	clay	0.025
Plant_richness	0.006	MEM5	0.120
Shannon_ARCH	0.004	MEM15	0.193

Variables	pval
pH	0.001
SCD	0.001
EC	0.001
ACE_BAC	0.001
Plant_richness	0.001
d13C	0.001
TotalP	0.002
SoilWaterC	0.002
ACE_FUN	0.012
bio2_dr	0.015
Shannon_BAC	0.021
Na2O	0.024

Variables	pval	Variables	pval
gdd5	0.001	Shannon_BAC	0.020
pH	0.001	ThickSand	0.109
ACE_BAC	0.001	MEM2	0.046
ACE_ARCH	0.001	sRadY	0.003
Calcite	0.001	Phyllosilicates	0.120
TotalP	0.001	MEM4	0.122
SoilWaterC	0.002	MEM6	0.146
Plant_richness	0.001	ThinSand	0.106
EC	0.005	soilTemp	0.068
slope25	0.019		

Variables	pval	Variables	pval
pH	0.001	HI	0.006
bio5_tmaxw	0.001	Calcite	0.023
EC	0.022	MEM12	0.031
ACE_BAC	0.001	slope25	0.063
SoilWaterC	0.001	ACE_FUN	0.037
MEM2	0.002	P2O5	0.036
K2O	0.019	NiO	0.095
MEM3	0.007	Dolomite	0.093
MEM6	0.001	MEM9	0.046
Plant_richness	0.037	MEM11	0.144

Figure 3

The edaphic, topoclimatic, spatial and biotic factors best explaining the beta diversity of bacterial, archaeal, fungal and protists communities with only the significant variables of the final db-RDA model plotted using principal coordinate analysis (PCoA). The significance of variables was tested using ANOVA. The variation partitioning analysis was computed using the significant variables identified within

each category. Significance levels are as follows: *P < 0.05 and ** P < 0.01 and *** P < 0.001. Residuals indicate the remaining unexplained variance.

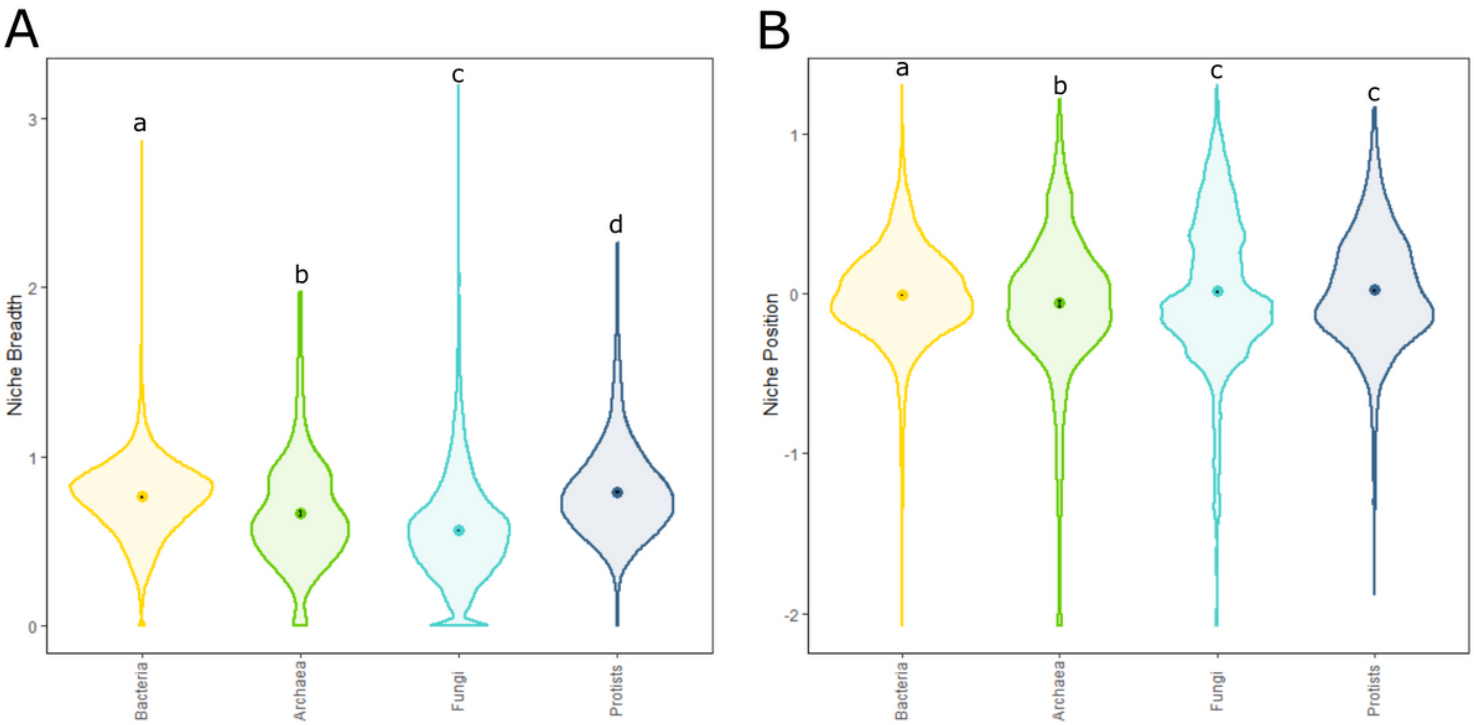


Figure 4

The average niche breadth (A) and niche position (B) of each ASV with a microbial group along the PCA axes. The mean (point) and standard error (black bars) are represented. Subscript characters with the same letters (a,b,c,d) indicate no statistically significant difference (p > 0.05) with the Tukey HSD test.

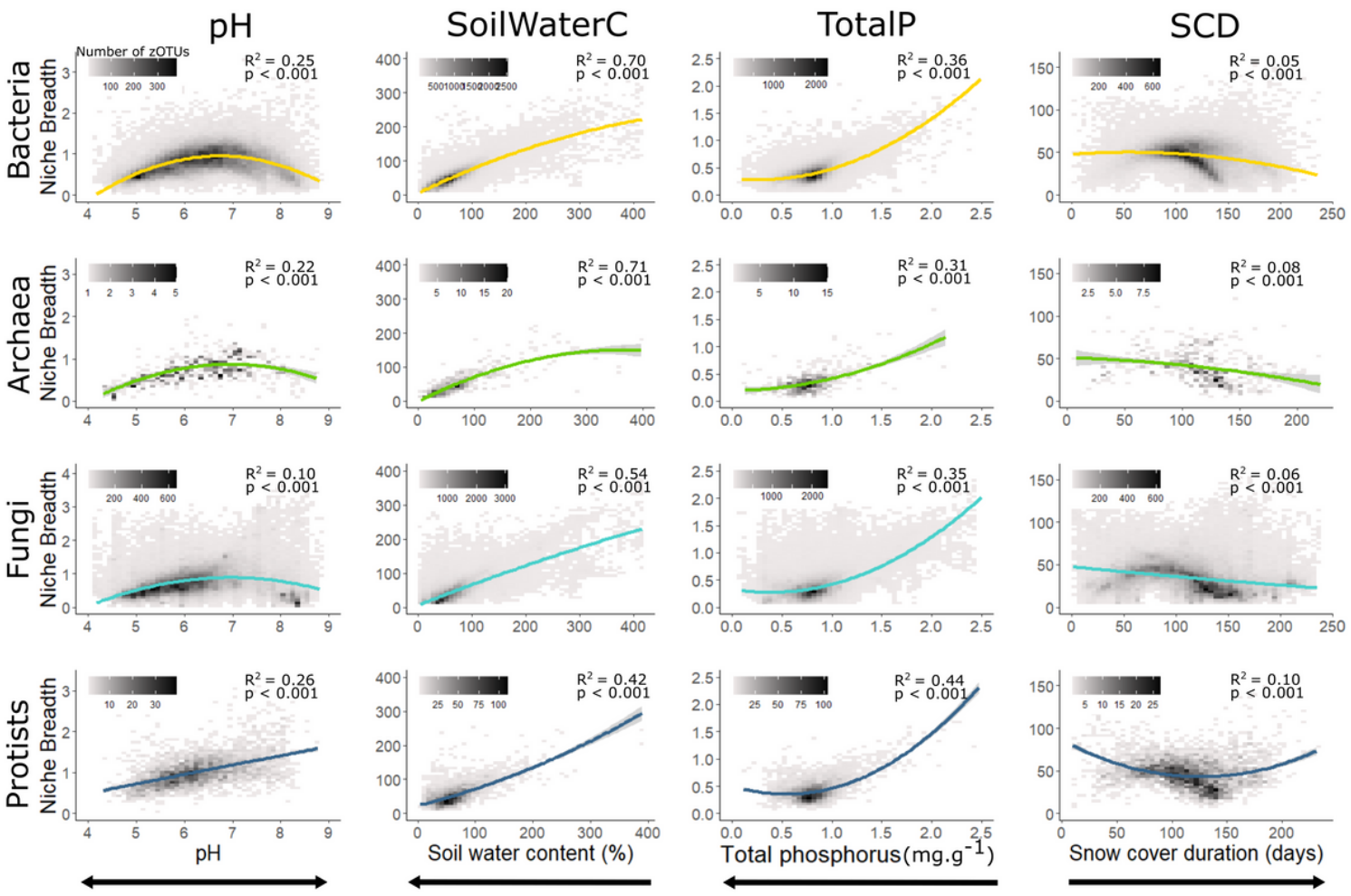


Figure 5

The average niche breadth (range) of each ASV along environmental gradients (niche position) and quadratic regressions. The colour gradient indicates the number of zOTU per bin (density). Each bin is constructed in ggplot2 using geom_bin2d. The arrows indicate the directionality of the environmental extreme of the gradients. In most cases within this study area, and especially for climatic variables, only part of the gradient is covered but the direction is still indicated.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigsandtable1.docx