Soil protists' niche breadth is more correlated with variations in edaphic gradients in the western Swiss Alps.

Vincent Pignard[1], Antoine Guisan[1], Lucie Malard[1]

[1]. Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

1st step project directed by Antoine Guisan, supervised by Lucie Malard

Master in Behavior, Evolution and Conservation, University of Lausanne, 2020

Abstract

Protists, unicellular eukaryotes, have been studied extensively in recent decades. Their ecology is still largely unknown. Aspects of their ecological niche, notably their niche breadth is still little studied. In this study, we examine the niche breadth of 7 phyla of soil protists in the western Swiss Alps. Using relative abundance of protists (over 136 sampling sites) and 75 environmental gradients, we are able to define the total niche breadth of soil protists in the study area. Here, we show that the total niche breadth of the soil protists is generally more correlated with variations of edaphic gradients than with topo-climatic gradients. We also show that the total niche breadth of the phylum Centrohelida is correlated with variations in topo-climatic gradients. This study is important because it allows to refine the distribution models of the protists. It also highlights the fragility of the phylum Centrohelida in the face of climate change.

Keywords

average niche position, environmental niche, niche width, microorganisms

Introduction

Hutchison (1957) describes the ecological niche as a space governed by environmental gradients, interspecific interactions and environmental barriers. The ecological niche is considered in two ways. The fundamental niche is the range over which the species studied can persist, without any biotic interaction or colonization limit (Kearney & Porter, 2004). When we add to this the biotic interactions and the limits of colonization, we find the niche realized (Leathwick, 1998; Svenning & Skov, 2004). There are several ways to study the realized environmental niche of a species or group of species. One of them is to investigate niche breadth, also called niche width (e.g. Roughgarden, 1972). . The niche breadth is the range of conditions which allow a population, species or clade to have a zero or positive growth rate (Carscadden et al., 2020).

For decades, species ecology has been a major topic of research. Niche breadth investigation is important because it allows to estimate range limits (Morin & Lechowicz, 2012) and to forecast how species may respond to climate change (Fisher-Reid et al., 2012). Unfortunately, microbial niche studies are still lacking, primarily due to methodological difficulties to characterise microbial taxa. Microorganisms play a major role in the ecosystem. They regulate plant diversity and productivity (Hartnett & Wilson, 2002; Van Der Heijden et al., 2007; Fitzsimons & Miller, 2010). They also play an important role in regulating biogeochemical cycles (Rousk & Bengtson, 2014). It is therefore necessary to study their niche to understand the dynamics of an environment.

Protists are a polyphyletic group of eukaryotic organisms. In fact, this clade includes all eukaryotes with the exception of plants, animals and fungi (Geisen et al., 2018). They have a huge diversity of lifestyles. Some are photoautotrophic, some are heterotrophic and some are even capable of both and qualified as mixotrophs (Geisen & Bonkowski, 2018). Many protists are symbiotic and can be parasitic or mutualistic with prokaryotes, plants, animals and fungi (De Vargas et al., 2015). The wide variety of morphologies and lifestyles of protists allows the group to colonize all environments on earth, even the harshest ones (Petz, 1997; De Jonckheere, 2006). The emergence of high-throughput sequencing, as well as electron microscopy, lead to major changes in the classification of protists in the last decades. Before the 21st century, the mode of nutrition determined the taxonomy of protists. Now, scientists increasingly refine the taxonomy of protists, thanks to sequencing, which is more systematic (Foissner, 1999; Geisen et al., 2018). The current classification shows 5 supergroups, or phyla, whose names are SAR (Stramenopila, Alveolata and Rhizaria), Archaeplastida, Excavata, Amoebozoa and Opisthokonta (Burki et al., 2016; Geisen et al., 2018). The scientific community further discusses two other groups: Centrohelida and Cryptophyta (named Cryptophyceae in the study) (Geisen et al., 2018). Soil protists are highly diverse (Grossmann et al., 2016). They play a role in regulating bacterial communities (Flues et al., 2017) but they are also non-negligible producers of organic carbon (Seppey et al., 2017).

In this study, we investigated the environmental niche of soil protists by examining differences in niche breadth among soil protist phyla living in the western Swiss Alps. The western Swiss Alps are a hotspot for research and the large variations in elevation and soil properties make it possible to test the distribution of species over a wide range of topo-climatic and edaphic gradients (Yashiro et al., 2016; Seppey et al., 2019). We calculated the average environmental niche breadth and niche position of each zero-radius Operational Taxonomic Unit (zOTU). These results allowed the comparison of the total niche breadth as well as gradient specific niche breadths of 7 phyla of protists mentioned above and subphyla of two phyla (Amoebozoa and SAR). We hypothesised that the niche breadth of protists would be different for each gradient. We also hypothesised that the niche breadth would differ among the groups studied. Characterising the niche breadth is important because it provides information that could refine the distribution models of protists (Banta et al., 2012). Such models are crucial for predicting hotspots of protist diversity (García, 2006) as well as environments likely to support rare microbial species (De Siqueira et al., 2009). This could serve to better anticipate the effects of climate change on the distribution of protists (Rehfeldt et al., 1999) and may also help predicting the risk of colonization of invasive species (Litmer & Murray, 2019).

Method

Data recovery

Professor Guisan's team provided four databases. The first contained values of 75 environmental gradients based on 136 sampling sites. These included topo-climatic and edaphic variables. The second contained the taxonomy of 3419 protist zero-radius Operational Taxonomic Units (zOTUs) . The third contained the relative abundance of each zOTU at each 136 sampling sites. The last one contained metadata data for each sampling sites, such as the day, month and year of sampling. The database also contained the coordinates and altitude of each sampling site. Sampling took place from July 6th to September 1st, 2013. The list of gradients used to measure the niche breadth is in Appendix 5. For more details on soil sample collection, edaphic and topo-climatic variables, see Yashiro et al. (2016) and Seppey et al. (2019).

Total niche breadth calculation

We used the "rda" function of the "vegan" package (Oksanen et al., 2007) on the R software version 4.0.3 (R Core Team, 2020) to produce a Principal Component Analysis (PCA) thanks to the value of each site for each of the 75 environmental gradients. We retrieved the sites values and species values of Principal Component 1 (PC1) and Principal Component 2 (PC2) to produce two plots (see Figure 2) with the values of sites and gradient vectors thanks to the "ggplot2" package (Wickham et al., 2016).

Total environmental niche breadth was calculated based on the PCA variables PC1 and PC2, used as proxy variables for the 75 environmental gradients. Specifically, niche position was calculated as the average of the PCA scores of sites where the zOTUs were present, weighted by their abundance. Similarly, niche breadth was calculated as the weighted square root of the PCA scores where the zOTUs were present. We used these values to investigate correlations between niche breadth and niche position along PC1 and PC2. We used the formula "y~poly(x,2)" in the "geom_smooth" function of the "ggplot2" package (Wickham et al., 2016) to make polynomial second-degree regressions and calculate the adjusted R-squared value and its p-value between niche breadth along PC1 and PC2. We used the same process, coupled with the "facet_wrap" function of the "ggplot2" package (Wickham et al., 2016) to calculate the adjusted R-squared value and its p-value between the total niche breadth along PC1 and PC2, for each phylum (Amoebozoa, Archaeplastida, Centrohelida, Cryptophyceae, Excavata, Opisthokonta, SAR, + zOTUs with undefined phylum).

Gradient specific niche breadth calculation

We selected 12 gradients that have been shown to influence soil protist communities. Seppey et al. (2019) identified "slope steepness" and "mean summer temperature" while Oliveiro et al. (2020) identified "mean annual temperature", "pH" and "mean annual precipitations" as key variables influencing soil protists distribution. Although topo-climatic gradients have been shown to primarily influence protist communities, edaphic properties often influence other taxonomic groups such as bacteria (Delgado-baquerizo et al., 2018) or fungi (Tedersoo et al., 2014). Here, we selected "carbon", "CaO", "SiO_2" and "soil water content" as Yashiro et al.

(2016) and Pellissier et al. (2014) identified them as important for bacterial and fungal communities within the same study region.

We also had an interest in testing “snow cover duration” and “freezing degree days” as they are key characteristics of mountainous environments such as the Alps (Chen et al., 2008; Sakai & Weiser, 1973); as well as “total phosphorus” because it is often a limiting factor related to plants (Chen et al., 1998). Topo-climatic gradients that are related to climate in the region included: snow cover duration, mean annual precipitation, freezing degree days, mean annual temperature and mean summer temperature. Edaphic gradients included: CaO, carbon, pH, SiO_2, soil water content, slope steepness, total phosphorus.

We used the same process as for the total niche breadth calculation, but we replaced the PC values by the gradient values on each site, in the calculations. We repeated the process for each of the 12 gradients of interest. We calculated the adjusted R-squared value as well as its p-value between the niche breadth and the average niche position for all zOTUs together, then differentiated by phylum. We repeated the process for each subphylum of SAR and Amoebozoa as they contained several hundred zOTUs.

Results

Total niche breadth analysis

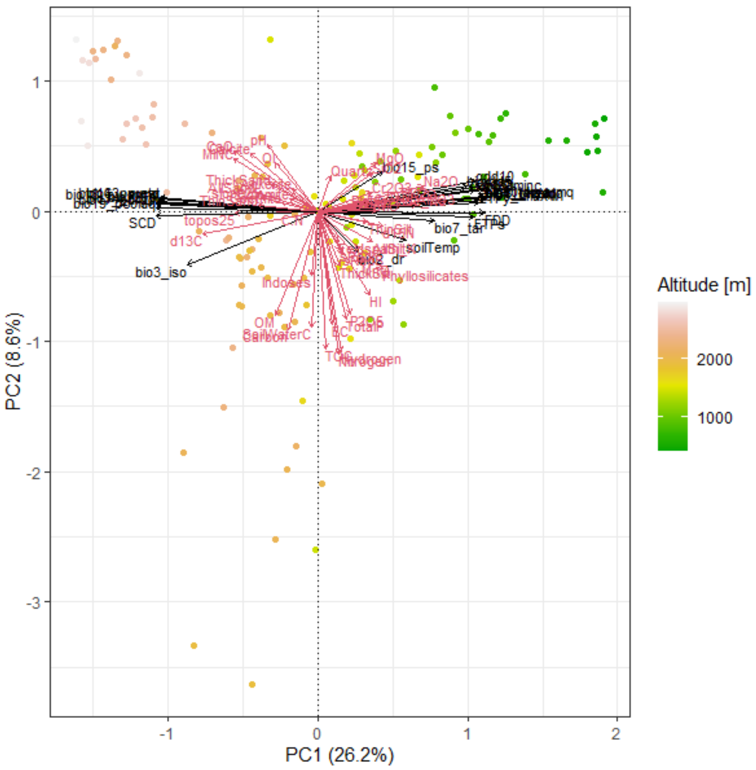

Figure 1: Principal Component Analysis of 75 environmental gradients. Each dot corresponds to one sampling site. Topo-climatic gradients are shown in black while edaphic gradients are shown in red. Principal Component 1 (PC1) explains 26.2% of the total variance. Principal Component 2 (PC2) explains 8.6% of the total variance. These two axes are the ones that explain the most of variance.

To measure the niche breadth relative all 75 gradients (total niche breadth), we first performed a Principal Component Analysis. Figure 1 shows that the elevation of the sampling sites is gradual along the PC1 axis, with sites at higher altitude on the left and sites at lower altitude on the right. In addition, we see that the topo-climatic gradients are concentrated along the PC1 axis. These vectors have a strong PC1 compound but a weak PC2 component. Edaphic gradients are more scattered.

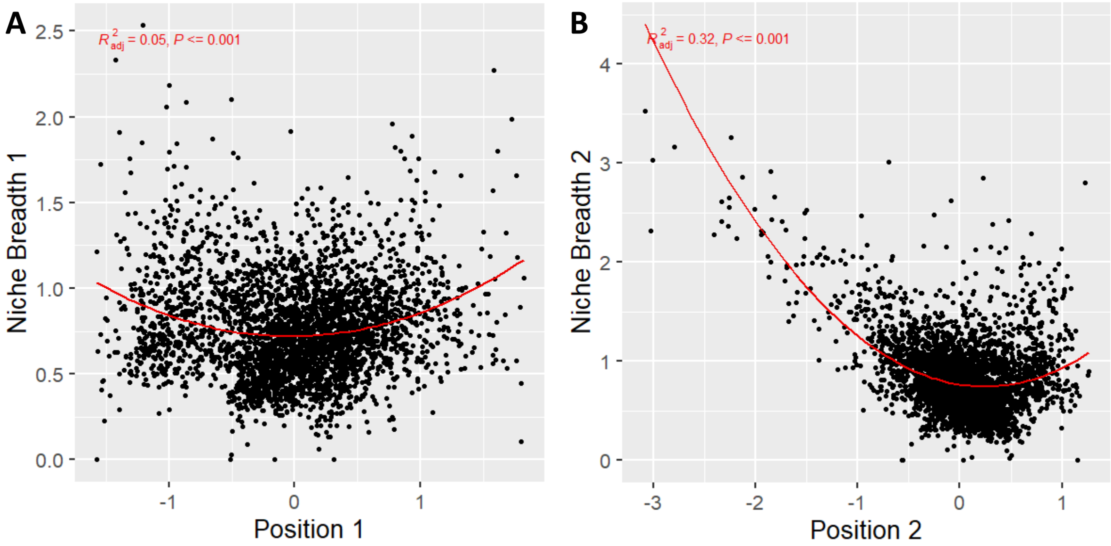

Figure 2. Second-degree polynomial regression between total niche breadth and total average niche position for each individual zOTU. A. "Position 1" and "Niche Breadth 1" are calculated on the base of sites' PC1 coordinate in the PCA plot. B. "Position 2" and "Niche Breadth 2" are calculated on the base of sites' PC2 coordinate in the PCA plot. "$R^2_{adj.}$" is the adjusted R-squared value. "P" is the p-value.

Thanks to the PCA site values, we were able to calculate the correlation between the total niche breadth and the total average niche position for each zOTUs and then clustered by phylum. Figure 2A shows the correlation between all protists average niche position on PC1 and their total niche breadth related to PC1. The adjusted R-squared value is equal to 0.05 with a p-value smaller than 0.001. This mean that the total niche breadth is weakly correlated with the average niche position of PC1. Figure 2B shows the total niche breadth of all our protists in function of their average niche position of PC2. The correlation value is equal to 0.32 with a p-value smaller than 0.001 between the total niche breadth and the total average niche position. This shows a high correlation between total niche breadth and average niche position for PC2. Table 1 shows (for the niche breadth relative to PC1, which corresponds to topo-climatic gradients) weak correlation values smaller than 0.08 for all zOTUs together, as well as for phyla Amoebozoa, Opisthokonta and SAR. The phyla Archaeplastida, Excavata and the group of zOTUs with unknown phylum do not show a correlation between the total niche breadth and the average niche position, for values taken from PC1. The phylum Centrohelida has a high correlation value of 0.75. It is the only phylum with a strong correlation of total niche breadth with PC1, suggesting high correlation between niche breadth and topo-climatic gradients. Table 1 shows higher PC2 correlation values for phyla that had a low PC1 correlation value. PC2 is more related to edaphic gradients. This suggests, for these phyla, a high correlation between the niche breadth and the edaphic gradients. As a result, all zOTUs, as well as phyla Amoebozoa, Opisthokonta and SAR have a correlation value greater than 0.3. This is also the case for the phylum Archaeplastida, which has no significant correlation relative to PC1. The phylum Excavata has a correlation value of 0.94 between the total niche breadth and the total average niche position relative to PC2. The phylum Centrohelida and the group of non-classified zOTUs have no significant correlation value relative to PC2.

Table 1: **Adjusted R-squared values and their significance according to all zOTUs together and by phyla, along PC1 and PC2.** The red gradient is relative to the correlation values: the higher the correlation value, the redder the cell. Cryptophyceae phylum is not in the table because it contains only 2 zOTUs from the sampling area, which is not enough to proceed to second-degree polynomial regressions. Significance levels: ***: p-value < 0.001; **: p-value < 0.01; *: p-value ≤ 0.05; NS: p-value > 0.05. For exact p-values, see Appendix 1.

	All zOTUs		Amoebozoa		Archaeplastida		Centrohelida		Excavata		Opisthokonta		SAR		Unknown	
Environmental Gradients	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.
PC1	0.05	***	0.059	***		NS	0.75	**		NS	0.079	*	0.048	***		NS
PC2	0.32	***	0.35	***	0.39	***		NS	0.94	*	0.31	***	0.32	***		NS
Number of zOTUs	3419		531		175		9		5		57		2620		20	

Gradient specific niche breadth analysis

Table 2: **Adjusted R-squared values and their significance according to all zOTUs together and phyla, for 12 specific gradients.** The red gradient is relative to the correlation values: the higher the correlation value, the redder the cell. Gradients are classified according to the decreasing correlation values for all zOTUs. Under "Environmental Gradients", E stands for "Edaphic" and TC stands for "Topo-Climatic". Cryptophyceae phylum is not in the table because it contains only 2 zOTUs from the sampling area, which is not enough to proceed to second-degree polynomial regressions. Significance levels: ***: p-value < 0.001; **: p-value < 0.01; *: p-value ≤ 0.05; NS: p-value > 0.05. For exact p-values, see Appendix 2.

		All zOTUs		Amoebozoa		Archaeplastida		Centrohelida		Excavata		Opisthokonta		SAR		Unknown	
	Environmental Gradients	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.	adj. R^2	p-val.
E	CaO [wt%]	0.8	***	0.77	***	0.8	***		NS		NS	0.88	***	0.8	***	0.8	***
E	Carbon [%]	0.46	***	0.51	***	0.46	***	0.64	*		NS	0.53	***	0.46	***	0.51	***
E	SiO_2 [wt%]	0.46	***	0.44	***	0.49	***		NS		NS	0.44	***	0.48	***	0.37	**
E	Soil Water content [%]	0.42	***	0.46	***	0.34	***		NS		NS	0.56	***	0.42	***	0.44	**
E	Total Phosphorus [mg/g]	0.36	***	0.34	***	0.4	***	0.81	**	0.91	*	0.076	*	0.37	***	0.25	*
E	Slope Steepness [°]	0.33	***	0.31	***	0.42	***	0.84	**		NS	0.22	***	0.32	***		NS
E	pH	0.25	***	0.31	***	0.31	***		NS		NS	0.46	***	0.25	***		NS
TC	Mean Annual Precipitations [mm/m^2]	0.12	***	0.084	***	0.21	***	0.95	***		NS		NS	0.11	***		NS
TC	Snow Cover Duration [Days]	0.098	***	0.12	***		NS	0.7	*		NS		NS	0.098	***		NS
TC	Freezing Degree Days [°C]	0.079	***	0.079	***	0.11	***	0.8	**		NS		NS	0.081	***		NS
TC	Mean Annual Temperature [°C]	0.048	***	0.064	***	0.034	*	0.91	***		NS		NS	0.047	***		NS
TC	Mean Summer Temperature [°C]	0.042	***	0.053	***	0.055	**	0.92	***		NS		NS	0.041	***		NS
	Number of zOTUs	3419		531		175		9		5		57		2620		20	

We selected 12 specific gradients to investigate changes in niche breadth and niche position along these gradients. Table 2 shows that the edaphic gradients have greater correlation values than the topo-climatic gradients, for all zOTUs together, as well as for the 4 phyla with more than 50 zOTUs, i.e.: Amoebozoa, Archaeplastida, Opisthokonta and SAR. We see that phyla with the least zOTUs (Centrohelida, Excavata and Opisthokonta) as well as the group of non-classified zOTUs have four or more gradients for which the correlation value is non-significant. Table 2 also shows that the phylum Centrohelida has high correlation values (0.7 minimum) for all topo-climatic gradients. This same phylum has different values for the edaphic gradients, some values are high while others are non-significant. The phylum Excavata had only one significant correlation value. This is the relative value of the Total Phosphorus gradient. Overall, niche breadth is more correlated to edaphic gradients than to topo-climatic. The "CaO" gradient shows the largest correlation values with the niche breadth, for the six groups with the most zOTUs. The "mean annual precipitations" gradient is most correlated with the niche breadth of the phylum Centrohelida.

Subphylas' gradient specific niche breadth Analysis

Table 3: **Adjusted R-squared values and their significance according to Amoebozoa phylum and Amoebozoa's subphyla, for 12 specific gradients.** The red gradient is relative to the correlation values: the higher the correlation value, the redder the cell. Gradients are classified according to the decreasing correlation values for the Amoebozoa phylum. Under "Environmental Gradients", E stands for "Edaphic" and TC stands for "Topo-Climatic". Significance levels: ***: p-value < 0.001; **: p-value < 0.01; *: p-value ≤ 0.05; NS: p-value > 0.05. For exact p-values, see Appendix 3.

		Amoebozoa		Cavosteliida		Discosea		Gracilipodida		Protosteliida		Schizoplasmodiida		Tubulinea		Unknown	
	Environmental Gradients	adj. R^2	p-value	adj. R^2	p-value	adj. R^2	p-value	adj. R^2	p-value	adj. R^2	p-value	adj. R^2	p-value	adj. R^2	p-value	adj. R^2	p-value
E	CaO [wt%]	0.77	***	0.68	***	0.91	***	0.91	***	0.58	*	0.81	***	0.75	***	0.9	***
E	Carbon [%]	0.51	***	0.29	***		NS	0.53	***	0.44	*	0.76	***	0.59	***	0.61	***
E	Soil Water content [%]	0.46	***	0.56	***		NS		NS		NS	0.77	***	0.47	***	0.66	***
E	SiO_2 [wt%]	0.44	***	0.35	***	0.16	*	0.6	***	0.59	*	0.4	***	0.53	***	0.62	***
E	Total Phosphorus [mg/g]	0.34	***	0.6	***	0.28	**	0.66	***	0.51	*		NS	0.35	***	0.33	***
E	Slope Steepness [°]	0.31	***	0.31	***	0.75	***	0.79	***	0.79	***		NS	0.33	***	0.085	*
E	pH	0.31	***		NS	0.35	**		NS		NS		NS	0.15	***	0.42	***
TC	Snow Cover Duration [Days]	0.12	***	0.48	***		NS		NS		NS	0.34	***	0.094	***	0.1	*
TC	Mean Annual Precipitations [mm/m^2]	0.084	***		NS	0.32	**		NS		NS		NS	0.12	***		NS
TC	Freezing Degree Days [°C]	0.079	***	0.14	**	0.35	**	0.42	***		NS		NS	0.074	***		NS
TC	Mean Annual Temperature [°C]	0.064	***	0.25	***		NS		NS		NS	0.19	*	0.057	***		NS
TC	Mean Summer Temperature [°C]	0.053	***	0.22	***		NS		NS		NS	0.19	*	0.047	***		NS
	Number of zOTUs	531		53		27		37		11		34		317		52	

Tubulinea and the group of zOTUs whose subphylum is not determined, have non-significant correlation values for certain edaphic gradients and topo-climatic gradients. Correlation values for "snow cover duration", "mean annual temperature" and "mean summer temperature" gradients that are stronger than those for "pH" for Cavosteliida and Schizoplasmodiida subphyla. Correlation values of "freezing degree days" gradient are stronger than some of edaphic gradients for Discosea and Gracilipodida subphyla. Tubulinea is the only subphylum of Amoebozoa for which the values of correlation relative to the edaphic gradients are greater than that of the topo-climatic gradients. SAR subphyla all have higher correlation values for edaphic gradients than for topo-climatic gradients (see Table 4). Overall, not all edaphic gradients are more correlated with the niche breadth than topo-climatic gradients. The CaO gradient remains the most correlated for most groups. Only the slope steepness gradient is more correlated for the subphylum Protosteliida.

Table 4: Adjusted R-squared values and their significance according to SAR phylum and SAR's subphyla, for 12 specific gradients. The red gradient is relative to the correlation values: the higher the correlation value, the redder the cell. Gradients are classified according to the decreasing correlation values for the SAR phylum. Under "Environmental Gradients", E stands for "Edaphic" and Tc stands for "Topo-Climatic". 1 zOTU is considered as Holozoa but we cannot perform regressions on 1 sample, so Holozoa is not in the table. Significance levels: ***: p-value < 0.001; **: p-value < 0.01; *: p-value ≤ 0.05; NS: p-value > 0.05. For exact p-values, see Appendix 4.

Environmental Gradients		SAR adj. R^2	SAR p-value	Alveolata adj. R^2	Alveolata p-value	Rhizaria adj. R^2	Rhizaria p-value	Stramenopiles adj. R^2	Stramenopiles p-value
E	CaO [wt%]	0.8	***	0.77	***	0.84	***	0.78	***
E	SiO_2 [wt%]	0.48	***	0.39	***	0.56	***	0.46	***
E	Carbon [%]	0.46	***	0.53	***	0.41	***	0.35	***
E	Soil Water content [%]	0.42	***	0.48	***	0.38	***	0.36	***
E	Total Phosphorus [mg/g]	0.37	***	0.36	***	0.37	***	0.45	***
E	Slope Steepness [°]	0.32	***	0.24	***	0.39	***	0.39	***
E	pH	0.25	***	0.16	***	0.33	***	0.25	***
TC	Mean Annual Precipitations [mm/m^2]	0.11	***	0.085	***	0.13	***	0.15	***
TC	Snow Cover Duration [Days]	0.098	***	0.12	***	0.1	***	0.082	***
TC	Freezing Degree Days [°C]	0.081	***	0.043	***	0.085	***	0.2	***
TC	Mean Annual Temperature [°C]	0.047	***	0.073	***	0.042	***	0.056	***
TC	Mean Summer Temperature [°C]	0.041	***	0.068	***	0.038	***	0.04	***
Number of zOTUs		2620		1034		1166		419	

Discussion

The PCA plot (see Figure 1) showed the topo-climatic gradients spread along the PC1 axis and the edaphic gradients scattered on both axes. This coupled with the correlation values in Figure 2 suggested a higher correlation between niche breadth and edaphic gradients. When we investigated each phylum separately (see Table 1), we saw that the total niche breadth of Phyla Amoebozoa, Archaeplastida, Excavata, Opisthokonta and SAR were more correlated with gradients that have a large component on PC2 than gradients with a large PC1 component. Therefore, we expected to record greater correlations between niche breadth and average niche position along the selected edaphic gradients. On the contrary, Centrohelida phylum was the only one for which the correlation relative to PC1 was stronger than that relative to PC2. We therefore expected this phylum to have correlations values equally high for edaphic and topo-climatic gradients, because both groups of gradients have high components on PC1 (see Figure 1). The Unknown group had non-significant values of correlation between the total niche breadth and the average niche position for each component (PC1 and PC2). This group probably included zOTUs of various phyla. As a result, it is possible that the different memberships (especially from the phylum Centrohelida) mitigated the correlation values. This coupled with the small amount of zOTUs present in the group probably render correlation values relative to PC1 and PC2 non-significant.

Analysis of the 12 specific gradients (see Table 1) showed that the niche breadth of phyla Amoebozoa, Archaeplastida, (Excavata), Opisthokonta and SAR was more correlated with edaphic than topo-climatic gradients. This supported our reflexion concerning the correlation values relative to PC1 and PC2, as these were the same phyla which had higher correlations with PC2 This shows that the edaphic gradients drive the niche breadth of protists, just as they drive the bacterial communities (Yashiro et al., 2016). The phylum Excavata had only a significant correlation value (with total phosphorus gradient) on the 12 gradients tested (see Table 2). However, the low significance of this correlation and the low number of zOTUs in this phylum meant results have to be taken with caution. As predicted with the PCA analysis, the correlations values of the phylum Centrohelida were high for topo-climatic gradients as well as for some edaphic gradients. However, as for Excavata, the low number of zOTUs in this phylum could have act as a confounding factor and results should be taken with caution.

The results of the Amoebozoa subphyla (see Table 3) showed that the more we go into the lower taxonomic ranks, the more the data diverges between the groups. This was not expected because the lower the taxonomic level studied, the more niche conservatism should exist between groups, as shown by Prinzing et al. (2001) for higher plants and by Da Silva et al. (2020) for protists. This therefore suggested that the divergence of evolutionary history between the different subphyla of Amoebozoa is so great that there is only little niche conservatism between the different subphyla. On the other hand, the decrease in the number of individuals in each group limits the statistical power of the analyses. SAR subphyla, for their part, showed less differentiation between them than Amoebozoa subphyla, despite high number of zOTUs (high statistical power), suggesting that they might live in similar niches. Thus, the niche conservatism within the SAR phylum would be more important than within the Amoebozoa phylum.

The results of this study showed that at the phylum level, edaphic gradients tended to be more correlated with the niche breadth. This was not consistent with our hypothesis that each group would have phylum-specific correlations. On the other hand, correlation values were different for each gradient. Our second hypothesis was therefore validated at the phylum level. At the subphylum level, our two hypotheses were validated for Amoebozoa. The correlation values were different between the subphyla of Amoebozoa and between the gradients. However, for SAR, subphyla still show similar niche probability.

Overall, our study showed that edaphic gradients were more correlated with the niche breadth than the topo-climatic gradients. On the other hand results of Seppey et al. (2019) showed that the beta diversity of protists was more correlated with topo-climatic gradients. This divergence of results is surprising, although not inconsistent. They focused on beta diversity while we investigated the environmental niche (different aspects of protists communities). The niche truncation for topo-climatic gradients could explain their low correlation values with the niche breadth. While edaphic gradients in the study area covered a grade part of the total gradient (pH 3 to 9 out of a total of 0 to 14, slope steepness 0 to 60[°], for example), topo-climatic gradients covered a limited part of their global range (for example, mean annual temperatures from -5 to 10 [°C] that mean annual temperatures from -50 to 50 [°C] are possible). Because we only investigated a tiny portion of topo-climatic gradients, we may not see any trend. If we had the whole gradient, we might. This could also explain why the global study of the protists of Oliveiro et al. (2020) showed that some topo-climatic gradient, such as the mean annual precipitations shape the protistan communities.

This duality of results shows the importance of this study to improve the predictions of the protist distribution models. This study also highlights the fragility of certain groups in the face of climate change. The phylum Centrohelida is a major concern because it is particularly correlated with topo-climatic gradients. Research on the niche breadth of protists in other parts of the world would be a good thing to learn more about the dispersal of some phyla. More precise research on the phylum Centrohelida would also make it possible to distinguish which lower ranks within this phylum are more at critical risk in the face of brutal climate changes.

Acknowledgment

We would like to thank Lucie Malard (Department of Ecology and Evolution at the University of Lausanne, Switzerland) for her great involvement and good indications in this project. We would also like to thank the rest of the Guisan Group (Department of Ecology and Evolution at the University of Lausanne, Switzerland) for providing the data of the study and for their feedback on the experience. Finally, we thank Luca Soutter (University of Lausanne) for the help he gave us for the statistical analysis on the R software.

References

Banta, J. A., Ehrenreich, I. M., Gerard, S., Chou, L., Wilczek, A., Schmitt, J., ... & Purugganan, M. D. (2012). Climate envelope modelling reveals intraspecific relationships among flowering phenology, niche breadth and potential range size in Arabidopsis thaliana. *Ecology letters*, *15*(8), 769-777. https://doi.org/10.1111/j.1461-0248.2012.01796.x

Burki, F., Kaplan, M., Tikhonenkov, D. V., Zlatogursky, V., Minh, B. Q., Radaykina, L. V., ... & Keeling, P. J. (2016). Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1823), 20152802. https://doi.org/10.1098/rspb.2015.2802

Carscadden, K. A., Emery, N. C., Arnillas, C. A., Cadotte, M. W., Afkhami, M. E., Gravel, D., ... & Wiens, J. J. (2020). Niche breadth: Causes and consequences for ecology, evolution, and conservation. *The Quarterly Review of Biology*, *95*(3), 179-214. https://doi.org/10.1086/710388

Chen, H., Zheng, C., & Zhu, Y. (1998). Phosphorus: a limiting factor for restoration of soil fertility in a newly reclaimed coal mined site in Xuzhou, China. *Land Degradation & Development*, *9*(2), 115-121. https://doi.org/10.1002/(SICI)1099-145X(199803/04)9:2<115::AID-LDR281>3.0.CO;2-M

Chen, W., Wu, Y., Wu, N., & Luo, P. (2008). Effect of snow-cover duration on plant species diversity of alpine meadows on the eastern Qinghai-Tibetan Plateau. Journal of Mountain Science, 5(4), 327-339. https://doi.org/10.1007/s11629-008-0182-0

Da Silva, D., Aires, A. E., Zurano, J. P., Olalla-Tárraga, M. A., & Martinez, P. A. (2020). Changing Only Slowly: The Role of Phylogenetic Niche Conservatism in Caviidae (Rodentia) Speciation. *Journal of Mammalian Evolution*, 1-9. https://doi.org/10.1007/s10914-020-09501-0

De Jonckheere, J. F. (2006). Isolation and molecular identification of free-living amoebae of the genus Naegleria from Arctic and sub-Antarctic regions. *European Journal of Protistology, 42*(2), 115-123. https://doi.org/10.1016/j.ejop.2006.02.001

De Siqueira, M. F., Durigan, G., de Marco Júnior, P., & Peterson, A. T. (2009). Something from nothing: using landscape similarity and ecological niche modeling to find rare plant species. *Journal for Nature Conservation*, *17*(1), 25-32. https://doi.org/10.1016/j.jnc.2008.11.001

De Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., ... & Carmichael, M. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, *348*(6237). https://doi.org/10.1126/science.1261605

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., ... & Fierer, N. (2018). A global atlas of the dominant bacteria found in soil. *Science*, *359*(6373), 320-325. https://doi.org/10.1126/science.aap9516

Fisher-Reid, M. C., Kozak, K. H., & Wiens, J. J. (2012). How is the rate of climatic-niche evolution related to climatic-niche breadth?. *Evolution: International Journal of Organic Evolution*, *66*(12), 3836-3851. https://doi.org/10.1111/j.1558-5646.2012.01729.x

Fitzsimons, M. S., & Miller, R. M. (2010). The importance of soil microorganisms for maintaining diverse plant communities in tallgrass prairie. *American Journal of Botany*, *97*(12), 1937-1943. https://doi.org/10.3732/ajb.0900237

Flues, S., Bass, D., & Bonkowski, M. (2017). Grazing of leaf-associated Cercomonads (Protists: Rhizaria: Cercozoa) structures bacterial community composition and function. Environmental microbiology, 19(8), 3297-3309. https://doi.org/10.1111/1462-2920.13824

Foissner, W. (1999). Protist diversity: estimates of the near-imponderable. *Protist*, *150*(4), 363-368. https://doi.org/10.1016/S1434-4610(99)70037-4

García, A. (2006). Using ecological niche modelling to identify diversity hotspots for the herpetofauna of Pacific lowlands and adjacent interior valleys of Mexico. *Biological Conservation*, *130*(1), 25-46. https://doi.org/10.1016/j.biocon.2005.11.030

Geisen, S., Mitchell, E. A., Adl, S., Bonkowski, M., Dunthorn, M., Ekelund, F., ... & Spiegel, F. W. (2018). Soil protists: a fertile frontier in soil biology research. *FEMS Microbiology Reviews*, *42*(3), 293-323. https://doi.org/10.1093/femsre/fuy006

Geisen, S., & Bonkowski, M. (2018). Methodological advances to study the diversity of soil protists and their functioning in soil food webs. *Applied soil ecology*, *123*, 328-333. https://doi.org/10.1016/j.apsoil.2017.05.021

Grossmann, L., Jensen, M., Heider, D., Jost, S., Glücksman, E., Hartikainen, H., ... & Boenigk, J. (2016). Protistan community analysis: key findings of a large-scale molecular sampling. *The ISME journal*, *10*(9), 2269-2279. https://doi.org/10.1038/ismej.2016.10

Hartnett, D. C., & Wilson, G. W. (2002). The role of mycorrhizas in plant community structure and dynamics: lessons from grasslands. Plant and soil, 244(1-2), 319-331. https://doi.org/10.1023/A:1020287726382

Hutchinson, G. E. (1957). Concluding remarks. Cold Springs Harbor Symp. Quant. Biol. 22: 415-427. 1959. Homage to Santa Rosalia, or why are there so many kinds of animals? Amer. *Nature*, *93*, 145-159. http://dx.doi.org/10.1101/SQB.1957.022.01.039

Kearney, M., & Porter, W. P. (2004). Mapping the fundamental niche: physiology, climate, and the distribution of a nocturnal lizard. *Ecology*, *85*(11), 3119-3131. https://doi.org/10.1890/03-0820

Leathwick, J. R. (1998). Are New Zealand's Nothofagus species in equilibrium with their environment?. *Journal of Vegetation Science*, *9*(5), 719-732. https://doi.org/10.2307/3237290

Litmer, A. R., & Murray, C. M. (2019). Critical thermal tolerance of invasion: Comparative niche breadth of two invasive lizards. *Journal of Thermal Biology*, *86*, 102432. https://doi.org/10.1016/j.jtherbio.2019.102432

Morin, X., & Lechowicz, M. J. (2013). Niche breadth and range area in North American trees. *Ecography*, *36*(3), 300-312. https://doi.org/10.1111/j.1600-0587.2012.07340.x

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., & Suggests, M. A. S. S. (2007). The vegan package. *Community ecology package*, *10*(631-637), 719. http://cran.r-project.org/ , http://vegan.r-forge.r-project.org/

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

Pellissier, L., Niculita-Hirzel, H., Dubuis, A., Pagni, M., Guex, N., Ndiribe, C., ... & Guisan, A. (2014). Soil fungal communities of grasslands are environmentally structured at a regional scale in the Alps. *Molecular ecology*, *23*(17), 4274-4290. https://doi.org/10.1111/mec.12854

Petz, W. (1997). Ecology of the active soil microfauna (Protozoa, Metazoa) of Wilkes Land, east Antarctica. Polar Biology, 18(1), 33-44. https://doi.org/10.1007/s003000050156

Prinzing, A., Durka, W., Klotz, S., & Brandl, R. (2001). The niche of higher plants: evidence for phylogenetic conservatism. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1483), 2383-2389. https://doi.org/10.1098/rspb.2001.1801

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.r-project.org/.

Rehfeldt, G. E., Ying, C. C., Spittlehouse, D. L., & Hamilton Jr, D. A. (1999). Genetic responses to climate in Pinus contorta: niche breadth, climate change, and reforestation. Ecological monographs, 69(3), 375-407. https://doi.org/10.1890/0012-9615(1999)069[0375:GRTCIP]2.0.CO;2

Roughgarden, J. (1972). Evolution of niche width. *The American Naturalist*, *106*(952), 683-718. https://doi.org/10.1086/282807

Rousk, J., & Bengtson, P. (2014). Microbial regulation of global biogeochemical cycles. *Frontiers in microbiology*, *5*, 103. https://doi.org/10.3389/fmicb.2014.00103

Sakai, A., & Weiser, C. J. (1973). Freezing resistance of trees in North America with reference to tree regions. Ecology, 54(1), 118-126. https://doi.org/10.2307/1934380

Seppey, C. V., Broennimann, O., Buri, A., Yashiro, E., Pinto-Figueroa, E., Singer, D., ... & Lara, E. (2019). Soil protist diversity in the Swiss western Alps is better predicted by topo-climatic than by edaphic variables. *Journal of Biogeography*, *47*(4), 866-878. https://doi.org/10.1111/jbi.13755

Seppey, C. V., Singer, D., Dumack, K., Fournier, B., Belbahri, L., Mitchell, E. A., & 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. https://doi.org/10.1016/j.soilbio.2017.05.002

Svenning, J. C., & Skov, F. (2004). Limited filling of the potential range in European tree species. *Ecology Letters*, *7*(7), 565-573. https://doi.org/10.1111/j.1461-0248.2004.00614.x

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., ... & Smith, M. E. (2014). Global diversity and geography of soil fungi. *science*, *346*(6213). https://doi.org/10.1126/science.1256688

Van Der Heijden, M. G., Bardgett, R. D., & Van Straalen, N. M. (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology letters*, *11*(3), 296-310. https://doi.org/10.1111/j.1461-0248.2007.01139.x

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package ‘ggplot2’. *Create Elegant Data Visualisations Using the Grammar of Graphics. Version*, *2*(1), 1-189. http://ggplot2.org , https://github.com/hadley/ggplot2

Yashiro, E., Pinto-Figueroa, E., Buri, A., Spangenberg, J. E., Adatte, T., Niculita-Hirzel, H., ... & van der Meer, J. R. (2016). Local environmental factors drive divergent grassland soil bacterial communities in the western Swiss Alps. *Applied and environmental microbiology*, *82*(21), 6303-6316. http://dx.doi.org/10.1128/AEM.01170-16

Appendix

Appendix 1. P-values of Correlation values from Table 1. The values marked 0.001 are smaller or equal to 0.001.

	All zOTUs	Amoebozoa	Archaeplastida	Centrohelida	Excavata	Opisthokonta	SAR	Unknown
PC1	0.001	0.001	0.167	0.00647	0.779	0.0402	0.001	0.27
PC2	0.001	0.001	0.001	0.412	0.0286	0.001	0.001	0.262

Appendix 2. P-values of Correlation values from Table 2. The values marked 0.001 are smaller or equal to 0.001.

Environmental Gradients	All zOTUs	Amoebozoa	Archaeplastida	Centrohelida	Excavata	Opisthokonta	SAR	Unknown
CaO [wt%]	0.001	0.001	0.001	0.431	0.694	0.001	0.001	0.001
Carbon [%]	0.001	0.001	0.001	0.0194	0.0763	0.001	0.001	0.001
SiO_2 [wt%]	0.001	0.001	0.001	0.825	0.188	0.001	0.001	0.00774
Soil Water content [%]	0.001	0.001	0.001	0.414	0.76	0.001	0.001	0.00262
Total Phosphorus [mg/g]	0.001	0.001	0.001	0.00293	0.0468	0.0448	0.001	0.0329
Slope Steepness [°]	0.001	0.001	0.001	0.00163	0.439	0.001	0.001	0.394
pH	0.001	0.001	0.001	0.605	0.8	0.001	0.001	0.105
Mean Annual Precipitations [mm/m^2]	0.001	0.001	0.001	0.001	0.985	0.068	0.001	0.111
Snow Cover Duration [Days]	0.001	0.001	0.0975	0.011	0.388	0.287	0.001	0.0907
Freezing Degree Days [°C]	0.001	0.001	0.001	0.0032	0.683	0.284	0.001	0.532
Mean Annual Temperature [°C]	0.001	0.001	0.0189	0.001	0.344	0.153	0.001	0.0643
Mean Summer Temperature [°C]	0.001	0.001	0.0029	0.001	0.228	0.206	0.001	0.0886

Appendix 3. P-values of Correlation values from Table 3. The values marked 0.001 are smaller or equal to 0.001.

Environmental Gradients	Amoebozoa	Cavosteliida	Discosea	Gracilipodida	Protosteliida	Schizoplasmodiida	Tubulinea	Unknown
CaO [wt%]	0.001	0.001	0.001	0.001	0.0127	0.001	0.001	0.001
Carbon [%]	0.001	0.001	0.895	0.001	0.0404	0.001	0.001	0.001
Soil Water content [%]	0.001	0.001	0.616	0.0666	0.143	0.001	0.001	0.001
SiO_2 [wt%]	0.001	0.001	0.0441	0.001	0.0119	0.001	0.001	0.001
Total Phosphorus [mg/g]	0.001	0.001	0.0074	0.001	0.0238	0.126	0.001	0.001
Slope Steepness [°]	0.001	0.001	0.001	0.001	0.001	0.561	0.001	0.0428
pH	0.001	0.001	0.001	0.001	0.835	0.001	0.001	0.001
Snow Cover Duration [Days]	0.001	0.001	0.587	0.105	0.319	0.001	0.001	0.0283
Mean Annual Precipitations [mm/m^2]	0.001	0.0837	0.00388	0.285	0.383	0.515	0.001	0.76
Freezing Degree Days [°C]	0.001	0.00755	0.00224	0.001	0.264	0.861	0.001	0.0926
Mean Annual Temperature [°C]	0.001	0.001	0.221	0.301	0.462	0.0132	0.001	0.326
Mean Summer Temperature [°C]	0.001	0.001	0.328	0.475	0.483	0.0149	0.001	0.278

Appendix 4. P-values of Correlation values from Table 4. The values marked 0.001 are smaller or equal to 0.001.

Environmental Gradients	SAR	Alveolata	Rhizaria	Stramenopiles
CaO [wt%]	0.001	0.001	0.001	0.001
SiO_2 [wt%]	0.001	0.001	0.001	0.001
Carbon [%]	0.001	0.001	0.001	0.001
Soil Water content [%]	0.001	0.001	0.001	0.001
Total Phosphorus [mg/g]	0.001	0.001	0.001	0.001
Slope Steepness [°]	0.001	0.001	0.001	0.001
pH	0.001	0.001	0.001	0.001
Mean Annual Precipitations [mm/m^2]	0.001	0.001	0.001	0.001
Snow Cover Duration [Days]	0.001	0.001	0.001	0.001
Freezing Degree Days [°C]	0.001	0.001	0.001	0.001
Mean Annual Temperature [°C]	0.001	0.001	0.001	0.001
Mean Summer Temperature [°C]	0.001	0.001	0.001	0.001

Appendix 5. List of gradients used to measure the niche breadth .

Environmental gradients	Units
Edaphic gradients	
CaO	[wt%]
Carbon	[%]
pH	
SiO_2	[wt%]
Slope Steepness	[°]
Soil Water content	[%]
Total Phosphorus	[mg/g]
Topo-climatic Gradients	
Freezing Degree Days	[°C]
Mean Annual Precipitations	[mm/m^2]
Mean Annual Temperature	[°C]
Mean Summer Temperature	[°C]
Snow Cover Duration	[Days]