Distinguishing Archaeorhizomycetes as a potential saprobe and endophytic-symbiont along an altitudinal gradient in the Swiss Alps

Emily Seddon[1], Eric Pinto-Figueroa[1] and Antoine Guisan[1,2]

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

[2] Institute of Earth Sciences, University of Lausanne, Lausanne, Switzerland

Abstract

Archaeorhizomycetes is a recently proposed class of ancient soil fungi found in terrestrial ecosystems worldwide. Despite its global distribution and high abundance in the fungal community, its potential ecological role is poorly understood. We used Illumina sequenced fungal internal transcribed spacer (ITS) region, plant cover, and abiotic data collected in the western Swiss Alps to investigate whether the distribution of Archaeorhizomycetes along an altitudinal gradient provides insight into its ecological function. A selection of the most correlated variables was chosen using Spearman's rank correlation analyses and these factors were then fitted with linear and quadratic regression models. We observed that certain topoclimatic factors significantly explained Archaeorhizomycetes abundance. Archaeorhizomycetes is also significantly positively correlated with the plant species *Carex flacca*, which is known to interact with endophytic fungi, while significantly negatively correlated with a known mycorrhizal fungi, indicating potential competition between fungal families. Thus, we proposed that Archaeorhizomycetes acts as an endophytic-symbiont.

Keywords:

Saprobe, endophyte, topoclimatic, edaphic factors, elevational gradient, correlative analyses

Introduction

Understanding of the ecological role of cryptic soil microorganisms in terrestrial ecosystems is a continuous pursuit both for microbial and plant ecologists (van der Heijden *et al*., 2008; Pellissier *et al*., 2013; Buttigieg & Ramette, 2014). Archaeorhizomycetes is one of the most ubiquitous soil fungi classes with a dominant presence in the soil fungal community (Rosling *et al*., 2011; Rosling *et al*., 2013), however, its ecological role has not yet been clearly defined. Thus, there are unanswered questions as to how this fungal class contributes to the terrestrial ecosystem and under which environmental conditions are these taxa best suited to do so. Defining the potential ecological role of Archaeorhizomycetes through its significant relationships with biotic and abiotic factors will improve our understanding of the hidden biodiversity in the soil ecosystem.

Despite the advances for exploring soil biodiversity by non-culturing techniques such as DNA barcoding and the use of operational taxonomic units (OTUs) to describe microorganisms (Barberán *et al*., 2012), the ecological roles of the vast majority of soil bacteria and fungi remain uncharacterized and not accurately interpreted (Rosling *et al*., 2011; Rodriguez *et al*., 2009). Most insights on the ecological functions of bacteria and soil fungal organisms have emerged from experimental studies on agricultural crop plants and insect model species (Philippot *et al*., 2013). These studies have described soil fungi interacting with plant roots through a variety of symbiotic mechanisms, including mutualistic associations as either mycorrhizal or endophytic fungi (Carlile *et al*., 2001) and antagonistic associations as pathogenic fungi (Carlile *et al*., 2001). For example, approximately 80% of land plants engage in mutualisms with arbuscular mycorrhizal fungi (AMF) (van der Heijden *et al*., 1998).

Certain fungi do not require carbon from plants and can obtain substantial nourishment from decaying organic material (Lindahl *et al*., 2007; Carlile *et al*., 2001).

These decomposers are free-living soil saprobes, which break down organic material and release nitrogen from plant litter (Hodge & Fitter, 2010; Lindahl *et al*., 2007). Many soil fungi have the capability to switch from a free-living state to an endophytic life strategy and vice versa (Rodriguez *et al*., 2009). Endophytes are a diverse and ubiquitous fungal functional group, colonizing plant tissues without having an obvious negative impact on the plant host (Jumpponen & Trappe, 1998; Rodriguez *et al*., 2009). Most endophytes belong to the phylum Ascomycota and have been located in every family of sampled terrestrial plants in every type of biome, extending from tropical to arctic ecosystems (Rodriguez *et al*., 2009; Jumpponen & Trappe, 1998). Benefits that different forms of endophyte groups provide to plant hosts range from increasing nutrient uptake and biomass, improving drought tolerance, enhancing herbivore deterrence, and managing habitat-specific stressors as soil pH, temperature, and salinity (Jumpponen & Trappe, 1998; Rodriguez *et al*., 2009). However, the precise function of the majority of fungal endophytes remains to be explored and properly characterized (Rodriguez *et al*., 2009).

Rosling and colleagues (2011) proposed a new fungal class called Archaeorhizomycetes, which is taxonomically located within the phylum Ascomycota. The type species *Archaeorhizomyces finlayi* has been determined by laboratory culturing methods and electron micrograph images (Rosling *et al*., 2011). Previously, Archaeorhizomycetes was included in the Soil Clone Group 1 (SCG1), a highly abundant soil fungi found in a variety of biomes, ranging from tropical to alpine ecosystems (Rosling *et al*., 2011; Rosling *et al*., 2013). It has been suggested that Archaeorhizomycetes may be mycoparasitic on other fungal organisms, but this life strategy remains unexplored (Rosling *et al*., 2013). Archaeorhizomycetes has exhibited plant-host preference in studies performed in Alaska (Rosling *et al*., 2013), however, these preferences may not indicate its behavior towards plant hosts worldwide (Rosling *et al*., 2013). Archaeorhizomycetes

78 has been observed to interact with plant roots, yet these interactions are neither parasitic
79 nor mycorrhizal relationships, and is suggested to be an endophyte (Rosling *et al*., 2013).
80 Due to its broad geographical range and its strong affiliation with soil environments
81 harboring plant roots (Rosling *et al*., 2013), it is important to distinguish the potential
82 relationships Archaeorhizomycetes has with both abiotic factors, such as soil
83 physicochemical properties and topoclimatic variables, and biotic factors, such as plant
84 species and other soil organisms. To determine these relationships, we used a DNA
85 metabarcoded soil fungal dataset from 103 random-stratified sampling sites distributed
86 along an altitudinal gradient in the western Swiss Alps. This study area has been
87 previously and extensively investigated in order to understand plant species and microbial
88 community distributions along an elevational gradient (Pellissier *et al*., 2013). From these
89 studies, a collection of topoclimatic variables and an inventory of the plant species of the
90 region have been recorded (Pellissier *et al*., 2013; Dubuis, 2013). Elevation was chosen
91 as a proxy to which to infer the fungal OTU richness because differences in abiotic and
92 biotic factors are easily observable along an altitudinal gradient and can resemble
93 changes in latitude (Pellissier *et al*., 2014).
94 Here, exploratory and predictive-oriented analyses were used to examine whether
95 Archaeorhizomycetes abundance OTUs are related to specific abiotic and biotic factors.
96 The edaphic factors were first reduced by agglomerative hierarchical clustering. Next, co-
97 abundance analyses were performed between Archaeorhizomycetes and other fungal
98 family OTUs, selected plant species, topoclimatic factors, and edaphic properties to
99 choose specific environmental variables. Correlative principal component analyses (PCA)
100 were then created to visualize the relationships between the abiotic and biotic factors and
101 Archaeorhizomycetes. Lastly, linear and quadratic regression models were fitted to these
102 selected predictors in order to discern which factors have significant correlations to

Archaeorhizomycetes abundance. This approach complements the culture-dependent characterization of these taxa performed by Rosling and colleagues (2011).

The aim of this study is to conclude if the distribution of Archaeorhizomycetes OTUs along an altitudinal gradient can provide insight into its ecological role. Specifically, which underlying environmental factors, abiotic and biotic, significantly explain its distribution and ecological role? Since Archaeorhizomycetes is neither pathogenic nor mycorrhizal fungi but rather is described as either an endophyte or a saprobe (Rosling *et al*., 2011), we hypothesize that Archaeorhizomycetes has the capacity to employ both life strategies as either free-living soil saprobe or occupy plant tissues as an endphytic-symbiont. It is important to note as this is an empirical correlative approach and future experimental studies examining Archaeorhizomycetes OTUs along an altitudinal gradient are needed to validate the evidence presented here.

Materials and Methods

Study area and sampling

This study was conducted along an elevation gradient from 600m to 3,000 m in a 700 km^2 area in the western Swiss Alps, Switzerland (Figure 1). Soil samples were collected from 103 sites across a variety of non-forested areas, during the plant-growing season between July and September 2012. We focused on non-forested areas because trees influence soil fungal composition in a different way than herbaceous plants (Pellissier *et al*., 2014). The plots were randomly chosen from a proportional design (Hirzel & Guisan, 2002) considering elevation, slope, and aspect strata. After removing plant litter, five soil cores (4 corners and 1 central, 0 - 5 cm in depth) were collected from a square 4 m^2 plot. Soil core samples were immediately placed on ice, and were prepared for molecular and soil chemical analyses within 24–36 hours after collection.

129 ***Plant species and abiotic data***

130 The western Swiss Alps study area represents a heterogeneous environment of
131 biotic (as plant species, other soil microbes) and abiotic factors (edaphic variables,
132 topoclimatic factors). For this study, a total of 448 plant species previously recorded as
133 percent cover (Dubuis, 2013) during two projects (Modiplant, 2002 – 2006; Bioassemble,
134 2009) were used for further analyses. A total of 48 abiotic factors were taken into
135 consideration (Supplementary, Table S1). Data for the soil properties, for example nutrient
136 and mineral content, were measured and recorded from the soil samples in the laboratory
137 (Buri *et al*., unpublished). Data of 10 topoclimatic variables calculated at 25 m resolution
138 for the study area were selected (Pradervand *et al*., unpublished).
139
140 ***Fungal community assessment***

141 Fungal community composition was measured as described in Pinto-Figueroa *et*
142 *al*., (unpublished). Briefly, for each of the 103 samples, environmental DNA was extracted
143 from 1 g of soil using the MOBIO PowerSoil DNA isolation kit (MOBIO Laboratories) and
144 PCR analyses were used to amplify the fungal internal spacer transcriber 1 (ITS1) region
145 using customized barcode labelling on both forward ITS1F (Schmidt *et al.*, 2013) and the
146 reverse ITS2 primers. Paired-end sequencing (2 x 100 nt) was performed on the Illumina
147 Hiseq 2500 platform. Raw sequencing data was processed and prepared for QIIME 1.8
148 (Caporaso *et al*., 2011). Forty-nine million of demultiplexed paired-end reads were
149 analysed using open-reference OTU picking (Rideout *et al.*, 2014). Representative fungal
150 OTUs were rarefied at 2,100 reads per sample. For downstream analyses, relative
151 abundance of Archaeorhizomycetes and other fungal OTUs per sample were converted to
152 sequence reads by multiplying the relative abundance per sample by the rarefaction
153 value.
154

155 ***Data analysis***

156 Abiotic and biotic factors that best predict the ecological function of
157 Archaeorhizomycetes in terms of OTU abundance along an elevational gradient were
158 selected using correlative and statistical approaches in R version 3.2.2 (R Development
159 Core Team, 2015). The session information and specific R packages used to conduct
160 these tests are included in the supplementary information (Table S2). These different
161 exploratory and predictive methods were used in conjunction with one another in order to
162 select the most significant ecological factors.

163

164 ***Selection of abiotic factors***

165 A non-parametric Spearman rank correlation coefficient was applied separately to
166 each the topoclimatic factors and and the soil properties to assess how well each variable
167 is related to another in the group. The correlated factor groups determined by the
168 Spearman test were then plotted in a *k*-means graph graph (Supplementary, Figures S1a
169 and S1b) depicting the number of related clusters (*k*). The number of clusters retained
170 was *k*=10 for each the edaphic variables and the topoclimatic factors. The results from the
171 Spearman rank correlation test were then applied to an agglomerative hierarchical
172 clustering method using Ward's linkage criterion. For the hierarchical clustering, Ward's
173 linkage criterion was implemented because the algorithm minimizes the total within cluster
174 variance, hence each factor variable in each cluster is more closely related to the other
175 (Murtagh & Legendre, 2014). These hierarchical clusters determined by Ward's criterion
176 were then depicted in a dendrogram (Supplementary, Figures S2a and S2b), and split into
177 10 partitions as determined from the *k*-means graph. One factor was chosen from each of
178 the 10 clusters. This process allowed us to reduce the edaphic data by selecting one soil
179 explanatory factor per each cluster of similar variables in the dendrogram (Supplementary,
180 Figure S2a).

The relatedness between these selected abiotic factors were then spatially represented in a correlation PCA (Figure 2a). We selected the data to be spatially reduced into 3 principal components (PCs) and calculated the variance explained by the eigenvalues for each PC. We then summed these 3 variance calculations in to one value totaling the variance explained by the correlation PCA. Individual correlation PCAs were also generated for only the topoclimatic factors (Supplementary, Figure S3a) and the reduced edaphic factors (Supplementary, Figure S3b) to spatially interpret the relationships between variables.

Both hierarchical clustering and the PCA approaches were combined to visually maximize the (dis)similarities in the factor variables (Ramette, 2007) in relation to Archaeorhizomycetes, allowing for the reduction from 48 variables to 20. Although both methods are associated, the hierarchical clustering of the soil explanatory variables and the topoclimatic factors provide another perspective to the spatial ordination of these same factors in a PCA, as some of the interactions in the data in a PCA can become distorted (Ramette, 2007).

Selection of biotic factors

The plant species and other fungi factors were greatly reduced from the raw data set using a Spearman's rank correlation to determine the pairwise relationships between these variables and Archaeorhizomycetes abundance. A Spearman's rank test was used because the data is based on abundance rather than presence/absence. A filter was then applied to the results of the Spearman's rank test, with a threshold of ± 0.30 for the plant species and ±0.5 for the fungi families. The correlation coefficient is on a scale from -1 being highly uncorrelated to +1 being highly correlated, a monotonic relationship. The plant species which had a relative abundance of above ± 0.30 were retained as the most correlated with Archaeorhizomycetes, while all other factors below this threshold were

excluded from further analyses. Similarly, for the fungi OTUs, variables which had a correlation with Archaeorhizomycetes above ±0.5 were omitted from the final analyses.

Results obtained after filtering the data were then represented in a second correlation PCA relative to Archaeorhizomycetes (Figure 2b) for these biotic predators. Three PCs were chosen to describe the variance in the factors, and the variance determined by their eigenvalues were calculated and summed.

Co-abundance correlation plotting

The selections made from the Spearman's rank correlation analyses on both the abiotic and biotic factors were then plotted in a correlative co-abundance heat map with Archaeorhizomycetes (Figures 3a and 3b). The heat maps follow a Spearman's correlation range of -1 (highly uncorrelated) to +1 (highly correlated, a value of 1 is the variable paired with itself). The biotic factors heat map (Figure 3b) includes the filtered variables which met or were above the threshold of ± 0.30 for the plant species and ± 0.50 for the fungal OTUs. These heat maps offer another visualization to those of the correlation PCAs of the relationships among the factors and Archaeorhizomycetes and allow us to infer which factors are most correlated with these OTUs. The session information and specific R packages used to conduct the co-abundance correlations are included in the supplementary information (Table S2).

Linear and quadratic regression models

In order to determine which selected abiotic and biotic correlations are most significant in supporting our hypothesis, simple linear and a quadratic regression models were each fit to the data. Both models were used to account for the different distributions in the data relative to the OTUs abundance of Archaeorhizomycetes. These models were constructed using Archaeorhizomycetes as the response variable and the 33 selected

variables as the terms of the model. The Akaike Information Criterion (AIC) stepwise algorithm was applied to estimate the quality of the models under all possible permutations and further reduction of explanatory variables (Akaike, 1974).

To account for false-positives returned from the models, three different *p*-value adjustments were calculated. These techniques included a Bonferroni correction, a Holm (1979) correction, and a Benjamini & Hochberg (1995) (BH) correction. Three correction formulas were used because the Bonferroni correction is a highly conservative correction method for multiple testing (Holm, 1979). The other two methods, BH and the Holm correction, and less stringent in correcting for false discoveries in the data. The BH correction considers the family-wise error rate, in which the mathematical equations regulate the probability of comparing multiple variables in simultaneous modeling giving a false-positive (Benjamini & Hochberg, 1995). The three different correction techniques allow us to make more well-rounded interpretations of the results.

Table 3 and Table 4 present the abiotic and biotic factors that returned a significant *p*-value (≤ 0.05) from the best fit linear regression model and best fit quadratic regression model, respectively. The adjusted *p*-values obtained from the correction formulas are represented in each table as well. It is important to note that a simple linear regression model was used in this study and not a generalized linear model (GLM) and that the significant *p*-values obtained from the linear and quadratic model testing should be

Results

Selection for abiotic factors

After hierarchical clustering, 10 soil properties were chosen (Table 1). These edaphic factors were selected based on the ecological information they provide. Due to the high correlation between each topoclimatic variable, these 10 factors were not

reduced further. Thus, the abiotic factors were reduced from 48 variables to 20 and the results are depicted in Table 1.

The correlation PCA for the selected topoclimatic factors (Supplementary, Figure S3a) explained 87.43% of the variance with 3 PCs. The PCA for edaphic variables selected (Supplementary, Figure S3b) explained 63.79% of the variance with its 3 PCs, indicating a good relationship between these reduced factors. The final PCA with both topoclimatic and edaphic factors demonstrates in 2-dimensional space three independent groupings of correlated factors, and the 3 PCs describe 59.66% of the total variation in the data (Figure 2a). In Figure 2a, Archaeorhizomycetes is located toward the center of the correlation circle and appears to be highly correlated with sodium oxide (Na_2O) and aluminum oxide (Al_2O_3).

Co-abundance patterns with abiotic factors

Results of the Spearman's rank correlation for the abiotic factors show that potassium oxide (K_2O) returned the highest positive correlation with Archaeorhizomycetes (r_S = 0.30), while the minimum monthly average temperature (tmin) was the most negatively correlated (r_S = -0.17). The correlation coefficient values for the remaining 18 abiotic variables are: maximum monthly average temperature (tmax, r_S = -0.09); annual growing degree days with 3°C threshold limits (gdd, r_S = -0.10); minimum precipitation days per growing season (pmin, r_S = -0.01); maximum precipitation days per growing season (pmax, r_S = -0.04); sum precipitation days per growing season (psum, r_S = -0.03); daily average evapotranspiration per month (etp, r_S = -0.12); topographic position (topos, r_S = 0.04); topographic aspect (asp, r_S = 0.01); slope (slp, r_S = 0.14); soil temperature (r_S = 0.02); soil pH (r_S = 0.01); soil nitrogen concentration (r_S = 0.25); Al_2O_3 content (r_S = 0.26); Na_2O content (r_S = 0.02); organic matter (OM, r_S = 0.25); phosphorus pentoxide content (P_2O_5, r_S = -0.06); soluble phosphate (r_S = -0.11); isotopic nitrogen ($d^{15}N$, r_S = -0.14).

These results are visually depicted in Figure 3a. The 10 selected topoclimatic variables and 10 selected edaphic factors are listed in Table 1.

Co-abundance patterns with biotic factors

In the biotic PCA (Figure 2b), Archaeorhizomycetes is displayed in the lower left quadrant of the correlation circle, correlating closely with the non-mycorrhizal plant species *Carex flacca* and the fungal family Mortierellaceae. The cumulative variance percentage explained by the 3 PCs of the final biotic PCA 32.20%. This PCA includes only the reduced plant and fungal factors that are highly correlated with Archaeorhizomycetes.

From the Spearman's rank correlation between Archaeorhizomycetes and the 7 plant species, *Carex flacca* had the most positive correlation with Archaeorhizomycetes abundance, with a r_S = 0.39. This result is supported by the spatial ordination of *C. flacca* in the correlation PCA. Other plant species which displayed positive co-abundance with the taxa are: *Leucanthemum vulgare* (r_S = 0.34), *Alchemilla coriacea* (r_S = 0.32), *Anthyllis vulneraria* (r_S = 0.30), and *Silene vulgaris* (r_S = 0.27). There were two plants that were negatively correlated with Archaeorhizomycetes: *Stellaria graminea* (r_S = -0.32) and *Rumex acetosa* (r_S = -0.26) (Figure 3b).

For the 6 fungal families, the co-abundance analyses showed that the fungal family Mortierellaceae has the most positive relationship with Archaeorhizomycetes (r_S = 0.44). This is congruent with the spatial ordination given in the correlation PCA. Two other fungal families positively co-abunding with Archaeorhizomycetes: Sebacinaceae (r_S = 0.43) and Thelephoraceae (r_S = 0.18). The three fungal families that were most negatively correlated are: Archaeosporaceae (r_S = -0.26), Tubeufiaceae (r_S = -0.20), and Suillaceae (r_S = -0.20) (Figure 3b). The 7 selected plant species, including their families and proposed interaction with fungi, and 6 selected fungal OTUs, with their phyla and presumed function, are represented in Table 2.

Linear and quadratic regression models

The results from the linear regression and quadratic regression models are displayed in Table 3 and Table 4. Only predictors that were statistically significant with a *p*-value of ≤ 0.05 are listed, along with the adjusted false-positive *p*-value results. The non-mycorrhizal plant *C. flacca* is the only plant species with the most significant positive relationship to Archaeorhizomycetes (*p*-value = $3.53e^{-4}$, corrected *p*-value = $3.20e^{-3}$ for all 3 techniques). Squaring the terms in a quadratic model decreased the significance of the *p*-value for *C. flacca* (*p*-value = 0.042, corrected *p*-value = NS for all 3 techniques). Only one fungal family, Archaeosporaceae, had a significant negative relationship with Archaeorhizomycetes in the linear regression model and not found significant in the quadratic model (*p*-value = 0.028, corrected *p*-value = 0.050 BH; other 2 methods = NS). No fungal families were significant predictors of Archaeorhizomycetes abundance in the quadratic model. Although there was a strong correlation between Mortierellaceae and Archaeorhizomycetes as seen in the biotic PCA (Figure 2b), this factor was not determined to be a significant predictor from the linear and quadratic regression modeling.

From the linear model for the topoclimatic factors, only maximum monthly average temperature (tmax) had a significant positive correlation, with a *p*-value of $3.44e^{-3}$ (corrected *p*-value = 0.010 BH; 0.031 Bonferroni; 0.024 Holm), while annual growing degree days (gdd) had a significant negative correlation (*p*-value = $4.44e10^{-3}$, corrected *p*-value = 0.010 BH; 0.040 Bonferroni; 0.027 Holm). In the quadratic regression, maximum precipitation days per growing season (pmax) (*p*-value = $5.32e10^{-3}$, corrected *p*-value = 0.012 BH; 0.037 Bonferroni; 0.027 Holm) and daily average evapotranspiration per month (etp) (*p*-value = $2.48e10^{-3}$, corrected *p*-value = $8.70e10^{-3}$ BH; 0.017 Bonferroni; 0.015 Holm) each had significant negative correlations with Archaeorhizomycetes abundance.

Two edaphic variables were returned as significant factors from the linear model. Soil pH has a significant positive relationship (*p*-value = 0.048, corrected *p*-value = NS for

all 3 techniques), as does Al_2O_3 (p-value = 1.15e10^{-3}, corrected p-value = 5.20e10^{-3} BH; 0.010 Bonferroni; 9.20e10^{-3} Holm). From the quadratic model, only the soil concentration of Al_2O_3 had a significant positive relationship to Archaeorhizomycetes (p-value = 1.13e10^{-4}, corrected p-value = 8e10^{-4} for all 3 techniques). Despite the high correlation of Na_2O with Archaeorhizomycetes as seen in the abiotic PCA (Figure 2a), this factor was not determined to be a significant predictor from the linear and quadratic regression modeling.

Discussion

Soil microbes comprise an important and necessary role in the functioning and diversity of the terrestrial ecosystem. Studying ecological interactions is dynamic and complex due to the intricate relationship of multiple variables that influence specific responses in the organisms in question (Buttigieg & Ramette, 2014). Archaeorhizomycetes is a ubiquitous soil fungus yet much about the putative ecological interactions of these taxa remains to be studied. Specific fungal life strategies of Archaeorhizomycetes being non pathogenic and non mycorrhizal have already been determined (Rosling *et al*., 2011; Rosling *et al*., 2013). Thus, it is of great interest to further examine the potential influence and interactions these taxa have in the soil ecosystem.

To confirm our hypothesis that Archaeorhizomycetes is an endophyte with potential mutualistic symbiotic and saprotrophic characteristics, we identified the significant abiotic and biotic factors related with these taxa by an empirical correlative approach. Through correlative analyses, a large amount of abiotic and biotic data collected from the western Swiss Alps study area was reduced to a selection of factors that drive OTU abundance of Archaeorhizomycetes and predict a prospective ecological function.

From the correlation tests, 20 abiotic factors were chosen, composed of 10 topoclimatic factors and 10 edaphic properties (Table 1). A total of 4 topoclimatic factors, maximum monthly average temperature (tmax), annual growing degree days (gdd) with

362 3°C threshold limits, maximum precipitation days per growing season (pmax), and daily
363 average evapotranspiration per month (etp), had significant correlations with affecting the
364 OTU richness of Archaeorhizomycetes. From these results, it can be inferred that the
365 abundance of these taxa is influenced by warmer temperature conditions as indicated by
366 the positive correlation with maximum monthly average temperature. The negative
367 correlation with annual growing degree days was an unexpected result and we cannot
368 interpret how this variable is affecting the abundance of Archaeorhizomycetes. Both
369 predictors associated with moisture conditions, maximum precipitation days per growing
370 season and daily average evapotranspiration per month, are significantly negatively
371 related to the OTU abundance, denoting that Archaeorhizomycetes may be more rich in
372 drier environments. The remaining 6 topoclimatic factors do not have a significant impact
373 on Archaeorhizomycetes OTUs.
374 Two edaphic variables including increasing soil pH and low to increasing Al_2O_3
375 content were significantly related with Archaeorhizomycetes abundance. Soil pH values
376 were positively related to Archaeorhizomycetes, suggesting that these taxa can tolerate
377 more basic soil settings. This result is congruent with the strong correlation with soil pH
378 and OTU abundance found by Rosling and colleagues (2013). However, we cannot
379 assume that Archaeorhizomycetes displays a preference towards higher pH units because
380 there may be other environmental factors that drive abundance which overlap with a soil
381 pH gradient (Rosling *et al*., 2013). Low to slightly increasing levels of Al_2O_3 positively
382 drives Archaeorhizomycetes abundance, potentially meaning that these fungi are tolerant
383 to low levels of Al_2O_3 and thus may be able to outcompete other soil microbes that are
384 more sensitive to Al_2O_3 because at high concentrations, Al_2O_3 is toxic to plants and other
385 soil microbes (Huang *et al*., 2002).
386 For the biotic factors, one plant species was determined to have significant positive
387 co-abundance relative to Archaeorhizomycetes OTUs. *Carex flacca* is a known non-

mycorrhizal plant, and has demonstrated a robust affiliation with endophytic fungi (van der Heijden *et al*., 1998; Jumpponen & Trappe, 1998; and Johnson *et al*., 2003). The fact that this correlation was positive concludes that perhaps Archaeorhizomycetes interacts with living plant tissues as an endophyte. Thus, these taxa have the ability to form mutualistic symbioses as an endophyte, rather than being limited to only decomposing organic material as a saprobe.

One fungal family was found to have a negative relationship with Archaeorhizomycetes OTUs. Archaeosporaceae is in the phylum Glomeromycota and is characterized as an arbuscular mycorrhizal fungi (AMF) (Morton *et al.*, 2001). This negative correlation can infer that perhaps there is competition between free-living saprotrophic Archaeorhizomycetes and other fungal OTUs with different life strategies. We postulate that in the presence of mycorrhizal fungi such as Archaeosporaceae, these taxa and potential saprotrophic forms of Archaeorhizomycetes compete for nutrients.

Since Archaeorhizomycetes is a ubiquitous yet cryptic fungal class, our findings in this study provide novel and important information on its potential function. However, future laboratory experiments are necessary to supplement this correlative approach in order to investigate how Archaeorhizomycetes interacts with plant roots and other fungi, and how climate and soil characteristics shape abundance patterns. For example, by performing experiments on the significant predictors found in this study, we can better define how these taxa are influencing and interacting with their environment. Is Archaeorhizomycetes forming endophytic structures inside the plant tissues of *C*. *flacca*? One can answer this question by designing plant microcosm experiments similar to those carried out by van der Heijden and colleagues (1998) or Johnson and colleagues (2003), where *C*. *flacca* is first grown in soil inoculated with Archaeorhizomycetes OTUs, and then after growing, examine the plant tissues for Archaeorhizomycetes fungal structures.

The results from the linear and quadratic regressions did not return significant

414 factors that predict Archaeorhizomycetes' ability to employ a free-living saprotrophic
415 lifestyle, yet experiments conducted by Rosling and colleagues (2011 and 2013) suggest
416 that these taxa could be saprobes. Therefore, we propose further research to deliberately
417 study if Archaeorhizomycetes has saprotrophic qualities, and if there is demonstrated
418 competition for resources between other fungal OTUs and Archaeorhizomycetes. To test
419 this theory, an experiment similar to Hodge & Fitter's (2010) study can be conducted,
420 where only Archaeorhizomycetes is exposed to a patch of organic material labeled with
421 isotopic nitrogen (^{15}N), and then one can detect if the previously labeled ^{15}N moved inside
422 the fungal hyphae. Introducing Archaeosporaceae OTUs into the experiment along with
423 Archaeorhizomycetes and the investigating the amount of ^{15}N is present in either fungal
424 family can allow one to measure competition rates. There is also room to consider
425 whether the observed Archaeorhizomycetes OTU abundance is related to environmental
426 gradients of the abiotic factors by performing experiments manipulating the abiotic factors
427 and then measure how these taxa react.
428 In conclusion, our data support part of the initial hypothesis that
429 Archaeorhizomycetes has potential endophytic-symbiotic capabilities due to the significate
430 relationship to the endophytic plant species *C*. *flacca*. However, the results did not validate
431 the prediction that Archaeorhizomycetes is free-living saprobe, but further research may
432 confirm this postulation. Much consideration was taken when choosing which correlative
433 and exploratory analyses to use to properly explain the function Archaeorhizomycetes,
434 and we tried many additional approaches before arriving at those described in this study.
435 Due to the large amount of raw data collected from field sampling, the tactic of reducing
436 these variables, and then determining a putative ecological role of Archaeorhizomycetes
437 from these selected factors proved to be a formidable process, as it is difficult to isolate
438 one single factor that can indicate the ecology of an organism. Despite these challenges,
439 we arrived at similar conclusions to the potential ecological role of these taxa as did

Rosling and colleagues (2011), yet we took a correlative analytical approach on data from field sampling in the western Swiss Alps study area rather than performing specific laboratory and field experiments. We advise that in order to gain a richer understanding of Archaeorhizomycetes life strategies, further exploration is necessary. Determining the ecological role of Archaeorhizomycetes is of principal importance because these taxa are abundant worldwide and thus can have a large impact on the interworking of the terrestrial ecosystem.

Acknowledgments

I would like to thank Eric Pinto-Figueroa, my direct supervisor for this First Steps project. Without his insight, organization, enthusiasm, direction, and encouragement, this project would not have been the challenging yet enjoyable experience that it was. Eric was also responsible of the soil sampling and *in situ* data gathering at the study area. He developed the molecular analyses to assess the fungal content by Illumina sequencing. Aline Buri, who was responsible of the processing and gathering the soil physicochemical properties. Jean-Nicolas Pradervand, for providing the topoclimatic variables of the study area.

I also am grateful for the scholarly and practical contributions of Prof. Antoine Guisan, the Principal Investigator for this project. I would like to acknowledge his entire research group as well for their feedback on my work and for the previous collection of the plant species data. Lastly, a thank you to Prof. Claus Wedekind and the Department of Ecology and Evolution for organizing the initial Frist Steps presentations at the start of the semester.

References

Akaike H (1974) A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control* **19**: 716–723.

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**: 343–351.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* **57**: 289–300.

Buttigieg PL, Ramette A (2014) A Guide to Statistical Analysis in Microbial Ecology: a community-focused, living review of multivariate data analyses. *FEMS Microbiol Ecol*. **90**: 543–550.

Carlile MJ, Watkinson SC, Gooday GW (2001) *The Fungi*, 2nd edn, Academic Press, London, UK.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *PNAS* **108**: 4516–4522.

Derelle D, Declerck S, Genet P, Dajoz I, van Aarle IM (2012) Association of highly and weakly mycorrhizal seedlings can promote the extra- and intraradical development of a common mycorrhizal network. *FEMS Microbiol Ecol*. **79**: 251–259.

Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* **22**: 1-20.

Dray S, Siberchicot A, with contributions from Thioulouse J. Based on earlier work by Alice Julien-Laferrière (2015) adegraphics: An S4 Lattice-Based Package for the Representation of Multivariate Data. R package version 1.0-3. <http://CRAN.R-project.org/package=adegraphics>.

491 Dubuis A (2013) Predicting spatial patterns of plant biodiversity: From species to
492 communities. PhD thesis, University of Lauanne.
493 Forta F, Cruza P, Catriceb O, Delbruta A, Luzarretaa M, Stroia C, Jouanya C (2014) Root
494 functional trait syndromes and plasticity drive the ability of grassland *Fabaceae* to
495 tolerate water and phosphorus shortage. *Environmental and Experimental Botany*
496 **110**: 62–72.
497 Haug I, Weiß M, Homeier J, Oberwinkler F, Kottke I (2005) Russulaceae and
498 Thelephoraceae form ectomycorrhizas with members of the Nyctaginaceae
499 (Caryophyllales) in the tropical mountain rain forest of southern Ecuador. *New*
500 *Phytologist* **165**: 923–936.
501 Hirzel A, Guisan A (2002) Which is the optimal sampling strategy for habitat suitability
502 modelling. *Ecological Modelling* **15**: 331–341.
503 Hodge A, Fitter AH (2010) Substantial nitrogen acquisition by arbuscular mycorrhizal fungi
504 from organic material has implications for N cycling. *PNAS* **107**: 13,754–13,759.
505 Hohenstein S, Kliegl R (2015) remef: Remove Partial Effects. R
506 package version 1.0.6.9000. <https://github.com/hohenstein/remef/>.
507 Holm S (1979) A simple sequentially rejective multiple test procedure. *Scandinavian*
508 *Journal of Statistics* **6**: 65–70.
509 Huang PM, Bollag J-M, Senesi N (2002) *Interactions between soil particles and*
510 *microorganisms: impact on the terrestrial ecosystem*. Chichester, West Sussex,
511 England: John Wiley & Sons.
512 InfoFlora, *Centre national de données et d'informations sur la flore de Suisse*. Available
513 from < https://www.infoflora.ch/fr/flore/>. [17 Nov 2015].
514 Johnson D, Vandenkoornhuyse PJ, Leake JR, Gilbert L, Booth RE, Grime JP, Young
515 JPW, Read DJ (2003) Plant communities affect arbuscular mycorrhizal fungal
516 diversity and community composition in grassland microcosms. *New Phytologist*

517 **161**: 503–515.
518 Jumpponen A, Trappe JM (1998) Dark septate endophytes: a review of facultative
519 biotrophic root-colonizing fungi. *New Phytologist* **140**: 295–310.
520 Kassambara A (2015) factoextra: Extract and visualize the results of
521 multivariate data analyses. R package version 1.0.3. <http://www.sthda.com>
522 Kodsueb R, Jeewon R, Vijaykrishna D, McKenzie EHC, Lumyong P, Lumyong S, Hyde
523 KD (2006) Systematic revision of Tubeufiaceae based on morphological and
524 molecular data. *Fungal Diversity* **21**: 105–130.
525 *Leucanthemum vulgare* (oxeye daisy), *Invasive Species Compendium*. Available from <
526 http://www.cabi.org/isc/datasheet/13357>. [10 Dec 2015].
527 Lindahl BD, Ihrmark K, Boberg J, Trumbore SE, Högberg P, Stenlid J, Finlay RD (2007)
528 Spatial separation of litter decomposition and mycorrhizal nitrogen \nuptake in a
529 boreal forest. *New Phytologist* **173**: 611–620.
530 Morton JB, Redecker D (2001) Two new families of Glomales, Archaeosporaceae and
531 Paraglomaceae, with two new genera *Archaeospora* and *Paraglomus*, based on
532 concordant molecular and morphological characters. *Mycologia* **93**: 181–195.
533 Murtagh F, Legendre P (2014) Ward's Hierarchical Agglomerative Clustering Method:
534 Which Algorithms Implement Ward's Criterion? *Journal of Classification* **31**: 274–
535 295.
536 Oksanen JF, Blanchet G, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL,
537 Solymos P, Stevens MHH, Wagner H (2015) vegan: Community Ecology Package.
538 R package version <2.3-1. http://CRAN.R-project.org/package=vegan>.
539 Olsson PA, Tyler G (2004) Occurrence of non-mycorrhizal plant species in south Swedish
540 rocky habitats is related to exchangeable soil phosphate. *Journal of Ecology* **92**:
541 808–815.
542 Pellissier L, Niculita-Hirzel H, Dubuis A, Pagni M, Guex N, Ndiribe C, Salamin N, Xenarios

I, Goudet J, Sanders IR, Guisan A (2014) Soil fungal communities of grasslands are environmentally structured at a regional scale in the Alps. *Mol Ecol*. **23**: 4274–4290.

R Development Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Ramette A (2007) Multivariate analyses in microbial ecology. *FEMS Microbiol Ecol*. **62**: 142–160.

Rideout JR, He Y, Navas-Molina JA, Walters WA, Ursell LK, Gibbons SM, Chase J, McDonald D, Gonzalez A, Robbins-Pianka A, Clemente JC, Gilbert JA, Huse SM, Zhou H, Knight R, Caporaso JG (2014) Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ* **2**:e545 <https://doi.org/10.7717/peerj.545>.

Rodriguez RJ, White Jr. JF, Arnold AE, Redman RS (2009) Fungal endophytes: diversity and functional roles. *New Phytologist* **182**: 314–330.

Rosling A, Cox F, Cruz-Martinez K, Ihrmark K, Grelet GA, Lindahl BD, Menkis A, James TY (2011) Archaeorhizomycetes – a new class of ancient, widespread soil fungi. *Science* **333**: 876–879.

Rosling A, Timling I, Taylor DL (2013) Archaeorhizomycetes: Patterns of Distribution and Abundance in Soil. *Soil Biology* **36**: 333–349.

Schmidt PA, Bálint M, Greshake B, Bandow C, Römbke J, Schmitt I (2013) Illumina metabarcoding of a soil fungal community. *Soil Biol. Biochem*. **65**:128–132.

Wei T (2013) corrplot: Visualization of a correlation matrix. R package version 0.73. <http://CRAN.R-project.org/package=corrplot>.

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer New York.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T, Wiemken A, Sanders IR (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**: 69–72.

van der Heijden MGA, Bardgett RD, van Straalen NM (2008) The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters* **11**: 296–310.

Figure Legends

Figure 1: A digital elevation model (DEM) of the the western Swiss Alps study area in 2012 depicting the distribution of the 103 random-stratified soil sampling sites, as indicated with red circles. The resolution of the model is 5m. This image is courtesy of Eric Pinto-Figueroa and Jean-Nicolas Pradervand.

Figure 2a: The PCA correlation circle includes the 20 selected abiotic variables with Archaeorhizomycetaceae is explained with 3 axes, however only PC1 and PC2 are represented in the figure. The percentage of variance that each PC explains is included as well. The diameter of the correlation circle ranges from -1 to +1.

Figure 2b: The PCA correlation circle for the 13 selected biotic variables with Archaeorhizomycetaceae is explained with 3 axes, however only PC1 and PC2 are represented in the figure. The percentage of variance that each PC explains is included as well. The diameter of the correlation circle ranges from -1 to +1.

Figure 3a: The co-abundance heat map portrays the correlative relationship, either positive or negative, of each of the 20 selected abiotic factors with Archaeorhizomycetaceae (outlined in the purple box). Topoclimatic factors are represented with orange triangles and edaphic variables with brown stars. The shading of

the blue or red color in each box indicates the strength of the correlation, on a scale of -1 to +1.

Figure 3b: This co-abundance heat map shows the abundance relationship, either positive or negatively correlated, of each of the 13 selected biotic factors with Archaeorhizomycetaceae (outlined in the purple box). A filter threshold of ± 0.50 and ± 0.30 has been applied to the fungal families and plant species, respectively. Plant species are indicated with a green leaf and fungal families with a pink diamond. The shading of the blue or red color in each box indicates the strength of the correlation, on a scale of -1 to +1.

Table 1: The 20 selected abiotic factors based on *k*-means clustering, agglomerative hierarchical clustering, and PCA methods. With these analyses, the edaphic variables were reduced from 38 factors to 10, all given with their respective units of measurement.

Table 2: The 13 selected biotic factors based on Spearman's rank correlation with filtering thresholds. With these analyses, the plant species were reduced from 448 species to 7 factors and the other fungal OTUs were reduced from 96 families to 6. The families for each plant species and the phyla information and primary function of each fungal family are provided.

Table 3: Statistical results from the best fit linear regression model determined from an AIC test for Archaeorhizomycetaceae explained by the selected abiotic and biotic factors based on correlative co-abundances analyses. Predictor coefficients and their respective *p*-values are represented. The corrected *p*-values are also given, using three different correction techniques, Benjamini & Hochberg (BH), Bonferroni, Holm. The predictors are

grouped according to type, with plant species names colored green, fungal family OTUs in pink, topoclimatic factors in orange, and the edaphic variables in brown.

Table 4: Statistical results from the best fit quadratic regression model determined from an AIC test for Archaeorhizomycetaceae explained by the selected abiotic and biotic factors based on correlative co-abundances analyses. Predictor coefficients and their respective *p*-values are represented. The corrected *p*-values are also given, using three different correction techniques, Benjamini & Hochberg (BH), Bonferroni, Holm. The predictors are grouped according to type, with plant species names colored green, topoclimatic factors in orange, and the edaphic variables in brown.

Tables

Table 1: The 20 selected abiotic factors

Topoclimatic variables		
Minimum monthly average temperature [tmin,°C]	Minimum precipitation days per growing season [pmin, days]	Daily average evapotranspiration per month [etp, (1/10mm) / (day)]
Maximum monthly average temperature [tmax, °C]	Maximum precipitation days per growing season [pmax, days]	Topographic position [topos, unitless]
Annual growing degree days [gdd, day*deg] with 3°C threshold limits	Sum precipitation days per growing season [psum, days]	Topographic aspect [asp, (deg/percent)]
	Slope [slp, °]	
Edaphic properties		
Soil temperature [µS*cm-1]	Soil pH	Nitrogen [%]
Al_2O_3 [wt-%]	Na_2O [wt-%]	K_2O [wt-%]
P_2O_5 [wt-%]	Organic matter [OM, wt-%]	$d^{15}N$ [Isotopic nitrogen, per mil Air-N2]
Soluble P [mgP/kg DW]		

Table 2: The 13 selected biotic factors

Plant species	*Family*	*Mycorrizal ecology*	*References*
Carex flacca	Cyperaceae	Non-mycorrhizal	Johnson *et al*. (2003)
Leucanthemum vulgare	Asteraceae	Mycorrhizal	Invasive Species Compendium (2015)
Silene vulgaris	Caryophyllaceae	Weakly mycorrhizal	Derelle *et al.* (2012)
Alchemilla coriacea	Rosaceae	Unknown	InfoFlora (2015)
Anthyllis vulneraria	Fabaceae	Mycorrhizal, but plant biomass is negatively affected	Fort *et al*. (2014)
Rumex acetosa	Polgonaceae	Non-mycorrhizal	Olsson & Tyler (2004)

Stellaria graminea	Caryophyllaceae	Non-mycorrhizal	Olsson & Tyler (2004)

Fungal families	*Fungal Phyla*	*Primary Ecological Function*	*References*
Tubeufiaceae	Ascomycota	Saprobes	Kodsueb *et al.* (2006)
Suillaceae	Basidiomycota	Ectomycorrhizal (EM)	Carlile *et al*. (2001)
Archaeosporaceae	Glomeromycota	Arbuscular mycorrhizae (AM)	Morton *et al.* (2001)
Thelephoraceae	Basidiomycota	Ectomycorrhizal (EM)	Haug *et al. (2005)*
Mortierellaceae	Zygomycota	Saprobes	Carlile *et al*. (2001)
Sebacinaceae	Basidiomycota	Mycorrhizae	Carlile *et al*. (2001)

Table 3: Results for the best fit linear regression model

Archaeorhizomycetaceae as the response variable

				Corrected *p*-values					
Predictors	Estimate	*p*-value		BH		Bonferroni		Holm	
Carex flacca	74.7458	0.000353	***	0.0032	**	0.0032	**	0.0032	**
Archaeosporaceae	-15.7189	0.028058	*	0.0505	*	0.2525	NS	0.1403	NS
Miaximum monthly avg temperature (tmax)	5.0881	0.003436	**	0.0100	*	0.0309	*	0.0241	*
Annual growing degree days (gdd)	-2.792	0.004443	**	0.0100	*	0.0400	*	0.0267	*
Soil pH	109.8982	0.047486	*	0.0712	NS	0.4274	NS	0.1899	NS
Al_2O_3	126.8227	0.001154	**	0.0052	**	0.0104	*	0.0092	**

Signif. codes: '***' 0.001; '**' 0.01; '*' 0.05; NS - not significant

Residual standard error: 527.9 on 93 degrees of freedom (DF)

Multiple R^2: 0.3719

Adjusted R^2: 0.3112

F-statistic: 6.12 on 9 and 93 DF

p-value: 1.002e-06

Table 4: Results for the best fit quadratic regression model

Archaeorhizomycetaceae as the response variable

				Corrected *p*-values					
Predictors	Estimate	*p*-value		BH		Bonferroni		Holm	
Carex flacca^2	2.91E+00	0.042076	*	0.0651	NS	0.2945	NS	0.1683	NS
Maximum precipitation days per growing season (pmax^2)	-2.16E-02	0.005324	**	0.0124	*	0.0373	*	0.0266	*
Daily avg evapotransporation/month (etp^2)	-1.33E-01	0.002484	**	0.0087	**	0.0174	*	0.0149	*
$Al_2O_3{}^2$	3.87E+00	0.000113	***	0.0008	***	0.0008	***	0.0008	***

Signif. codes: '***' 0.001; '**' 0.01 ;'*' 0.05; NS - not significant

Residual standard error: 551.2 on 95 degrees of freedom (DF)

Multiple R^2: 0.3007

Adjusted R^2: 0.2492

F-statistic: 5.836 on 7 and 95 DF

p-value: 1.197e-05

Figures

Figure 1: Western Swiss Alps study area

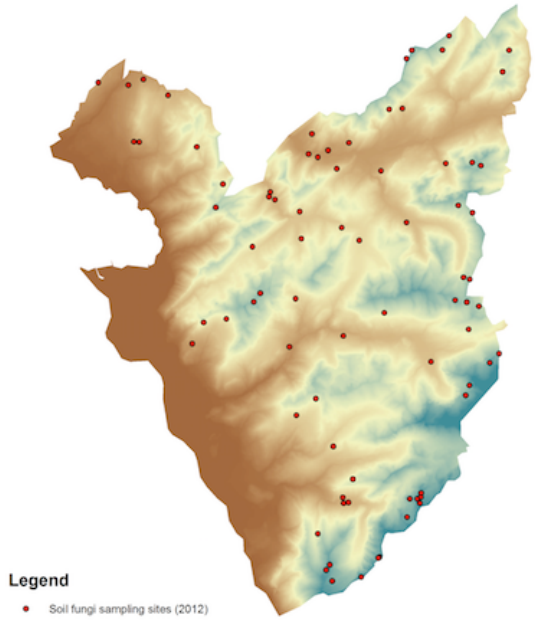


Figure 2a: PCA correlation circle of abiotic factors

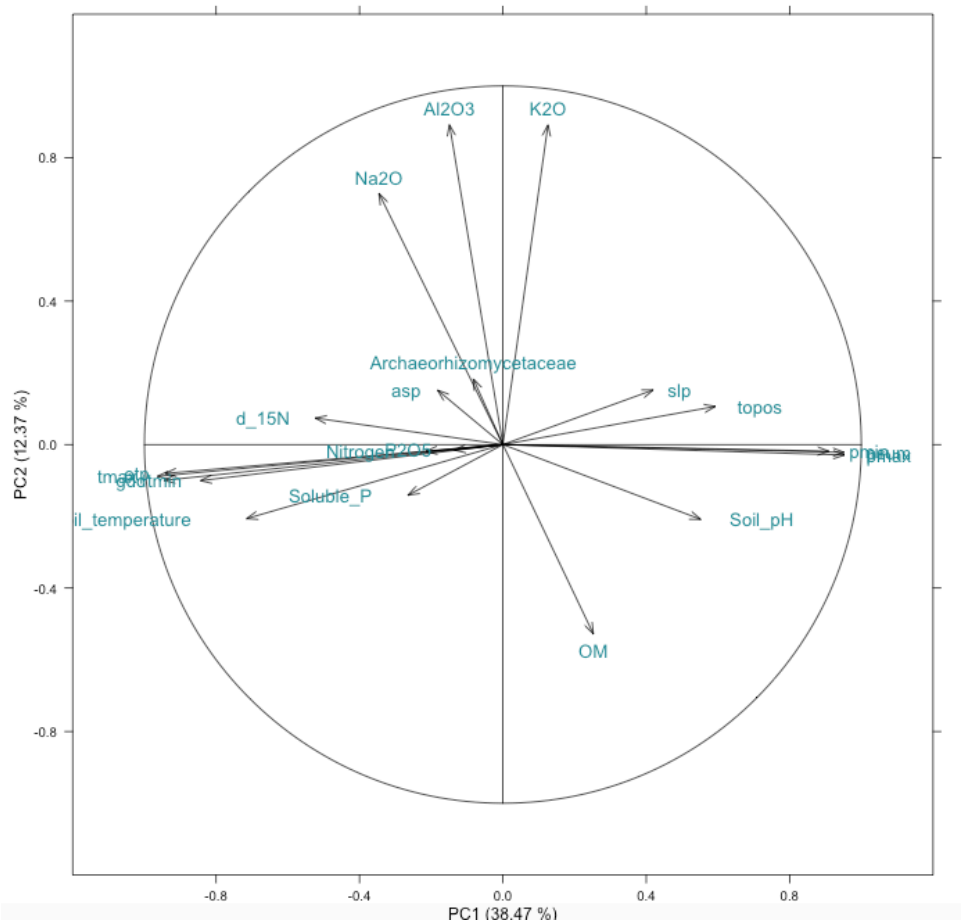


Figure 2b: PCA correlation circle of biotic factors

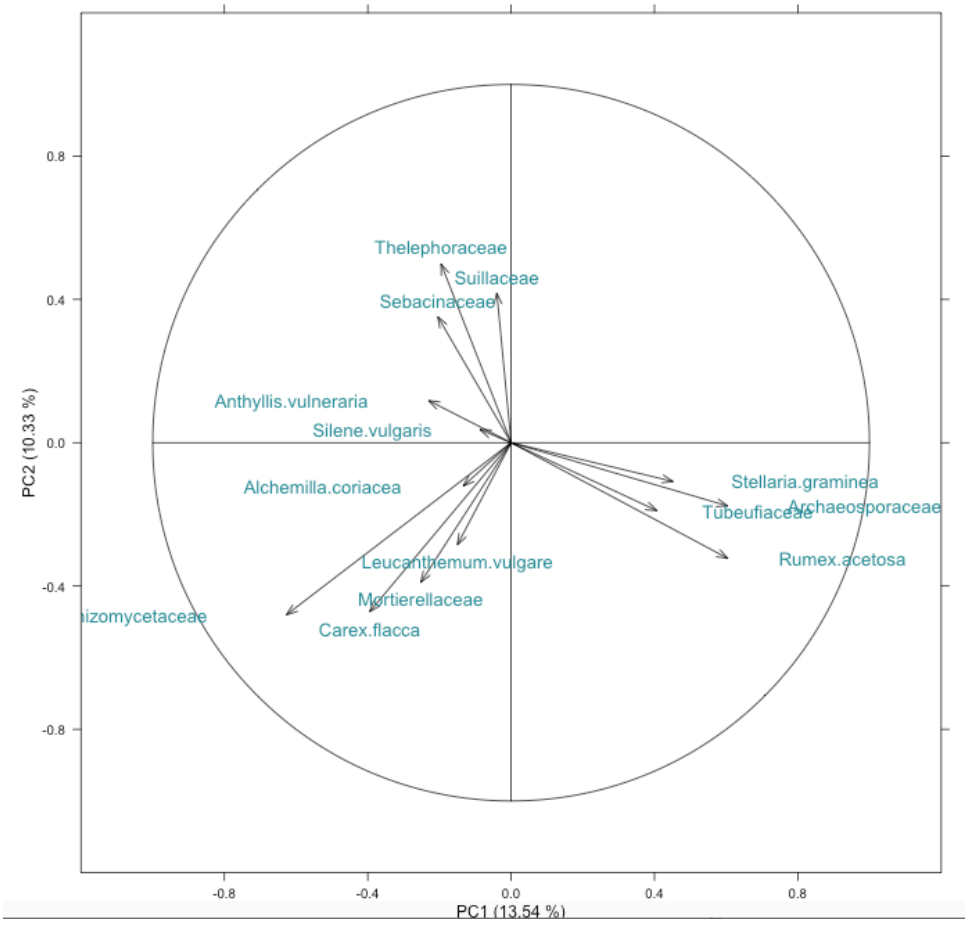


Figure 3a: Co-abundance heat map for abiotic factors

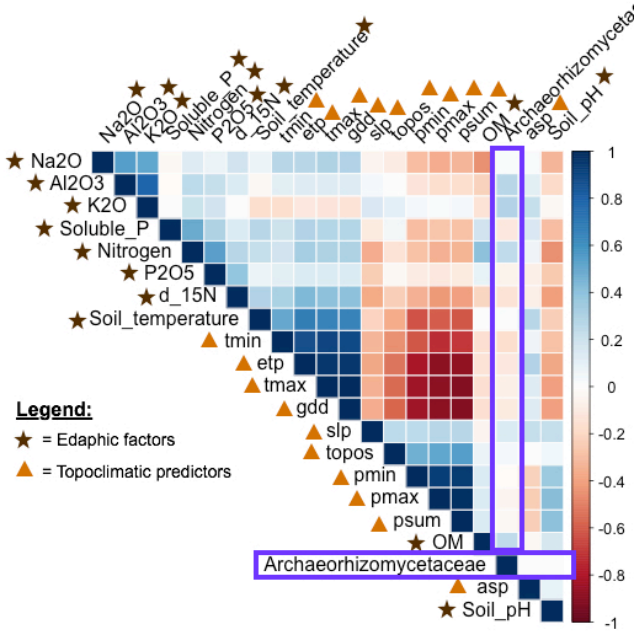
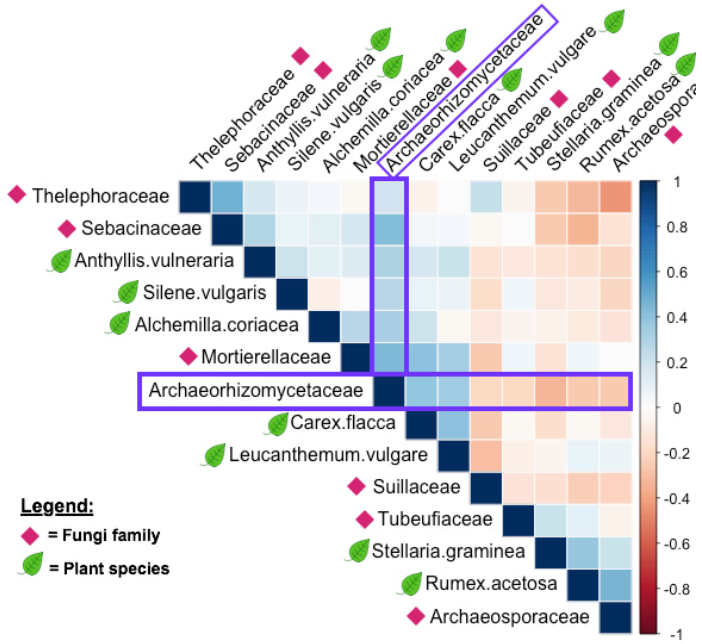


Figure 3b: Co-abundance heat map for biotic factors



Supplementary information

Supplementary Figure Legend:

Figure S1a: The number of *K*-means clusters determined from a Spearman's rank coefficient analysis on soil property factors and topographic and climatic factors (*Figure S1b*). From these graphs, 10 clusters (*k*=10) of soil explanatory variables and topo-climatic factors were selected and then plotted using an agglomerative hierarchical clustering method on a dendrogram tree (*Fig. S2a and S2b*) in order to infer their relatedness and to reduce the amount of variables for the final analysis.

Figure S2a: Soil properties and topographic and climatic factors (*Figure S2b*) are grouped according to their relatedness concluded from a Spearman's rank coefficient analysis. The dendograms were generated using the Ward hierarchical clustering method, which is constructed from a measure of a sum of squares of the data matrix and reduces the variation within a cluster of variables (Murtagh and Legendre, 2014). By choosing *k*=10, the nodes of both dendrograms are partitioned into 10 clusters of similar factor groups based on Euclidian distances between variables.

Figure S3a: The PCA correlation circle with only the topoclimatic variables is explained with 3 axes, however only PC1 and PC2 are represented in the figure. The percentage of variance that each PC explains is included as well. The diameter of the correlation circle ranges from -1 to +1.

Figure S3b: The PCA correlation circle for only the reduced edaphic variables which were selected by the hierarchical clustering method is explained with 3 axes, however only PC1 and PC2 are represented in the figure. The percentage of variance that each PC explains is included as well. The diameter of the correlation circle ranges from -1 to +1.

Table S1: The 48 total abiotic factors before reduction via agglomerative hierarchical clustering methods and Spearman's rank correlation analyses. There are 10 topoclimatic variables and 38 edaphic factors, all given with their respective units of measurement.

Table S2: The R session information detailing the packages used to perform the correlative and statistical analyses.

Tables

Table S1: The 48 abiotic factors

Topographic and Climatic variables		
Minimum monthly average temperature [tmin, °C]	Minimum precipitation days per growing season [pmin, days]	Daily average evapotranspiration per month [etp, (1/10mm) / (day)]
Maximum monthly average temperature [tmax, °C]	Maximum precipitation days per growing season [pmax, days]	Topographic position [topos, unitless]
Annual growing degree days [gdd, day * deg] with 3°C threshold limits	Sum precipitation days per growing season [psum, days]	Topographic aspect [asp, (deg/percent)]
Slope [slp, °]		

Edaphic properties		
Altitude [m]	Soil temperature [°C]	Soil water content [%]
Soil pH	EC 1:1 soil [µS*cm-1]	Phosphorous (P) [mg/g]
Total organic carbon (TOC) [%]	Mineral carbon content (MINC) [%]	Hydrogen index (HI) [mg HC/g TOC]
Oxygen index (OI) [mg CO_2/g TOC]	Nitrogen [%]	Carbon [%]
Hydrogen [%]	Phyllosilicates [%]	Quartz [%]
Feldspath K [%]	Plagioclase Na [%]	Calcite [%]
Dolomite [%]	Goethite [%]	Ankérite [%]
Indosés [%]	SiO_2 [wt-%]	TiO_2 [wt-%]
Al_2O_3 [wt-%]	Fe_2O_3 [wt-%]	MnO [wt-%]
MgO [wt-%]	CaO [wt-%]	Na_2O [wt-%]
K_2O [wt-%]	P_2O_5 [wt-%]	Organic matter (OM) [wt-%]
Cr_2O_3 [wt-%]	NiO [wt-%]	$d^{15}N$ [per mil Air-N2]
$d^{13}C$ [per mil V-PDB]	Soluble P [mgP/kg DW]	

Table S2: R session information

R version 3.2.2 (2015-08-14)

Platform: x86_64-apple-darwin13.4.0 (64-bit)

Running under: OS X 10.11.2 (El Capitan)

locale:

en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:

grid	stats	graphics	grDevices
utils	datasets	methods	base

other attached packages:

caret_6.0-58	corrplot_0.73	factoextra_1.0.3	vegan_2.3-1
lattice_0.20-33	permute_0.8-4	adegraphics_1.0-3	ade4_1.7-2
ggplot2_1.0.1	remef_1.0.6.9000		

Figures

Figure S1a: *K*-means cluster graph for soil properties

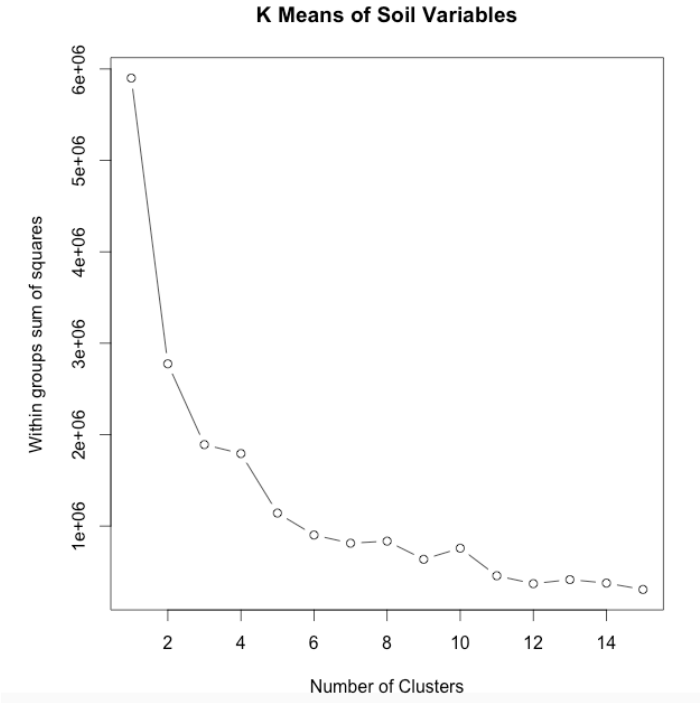


Figure S1b: *K*-means cluster graph for topographic and climatic variables

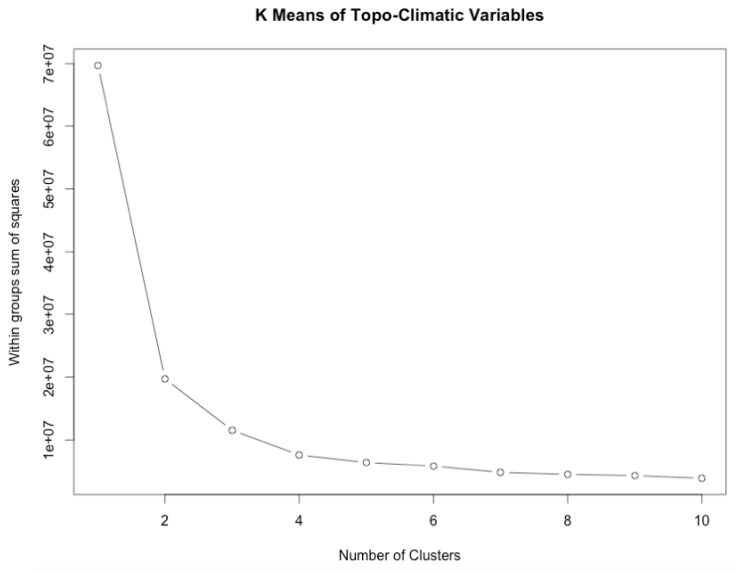


Figure S2a: Hierarchical cluster dendrogram of edaphic properties

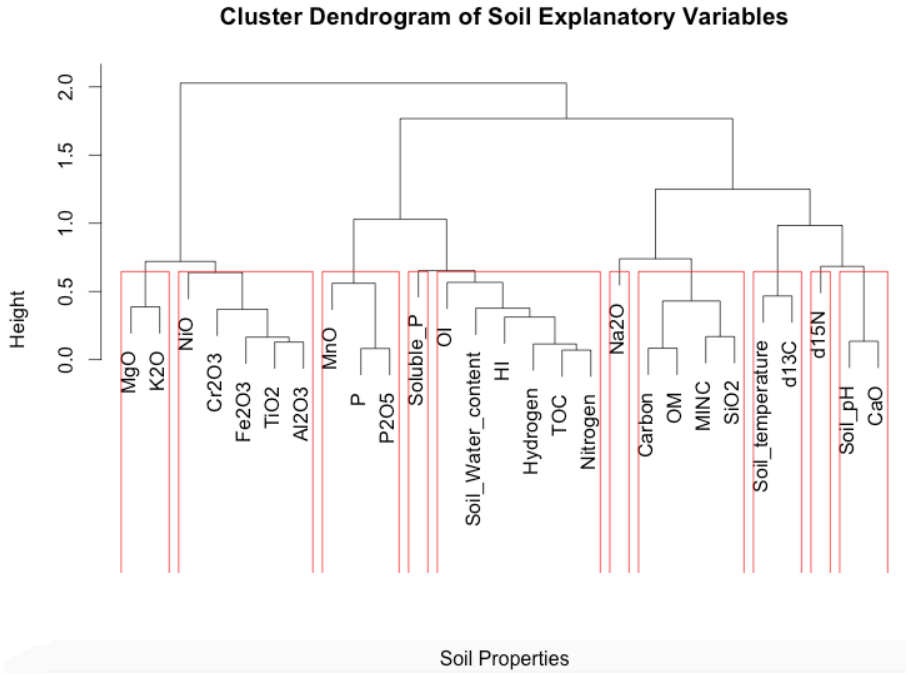


Figure S2b: Hierarchical cluster dendrogram of topoclimatic variables

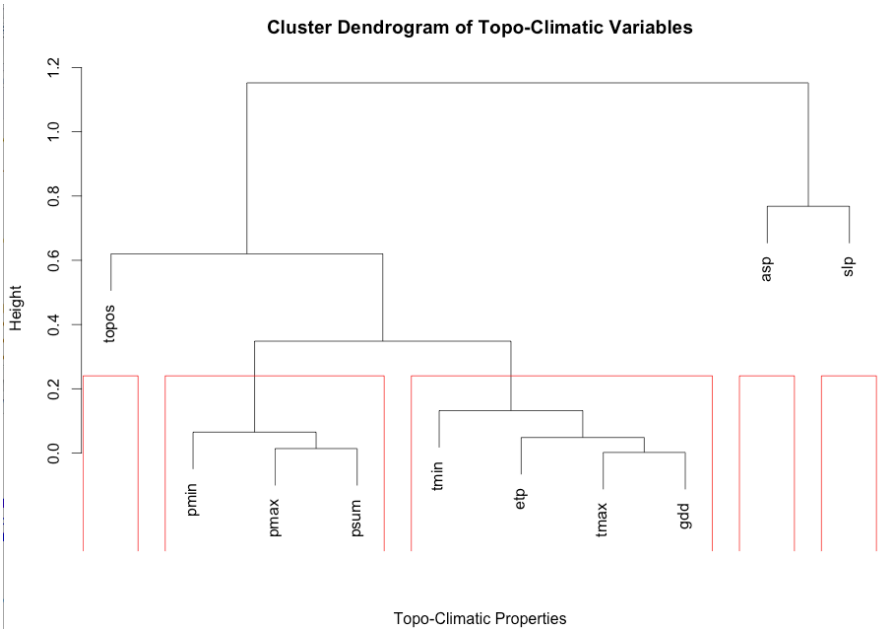


Figure S3a: PCA correlation circle of topoclimatic variables

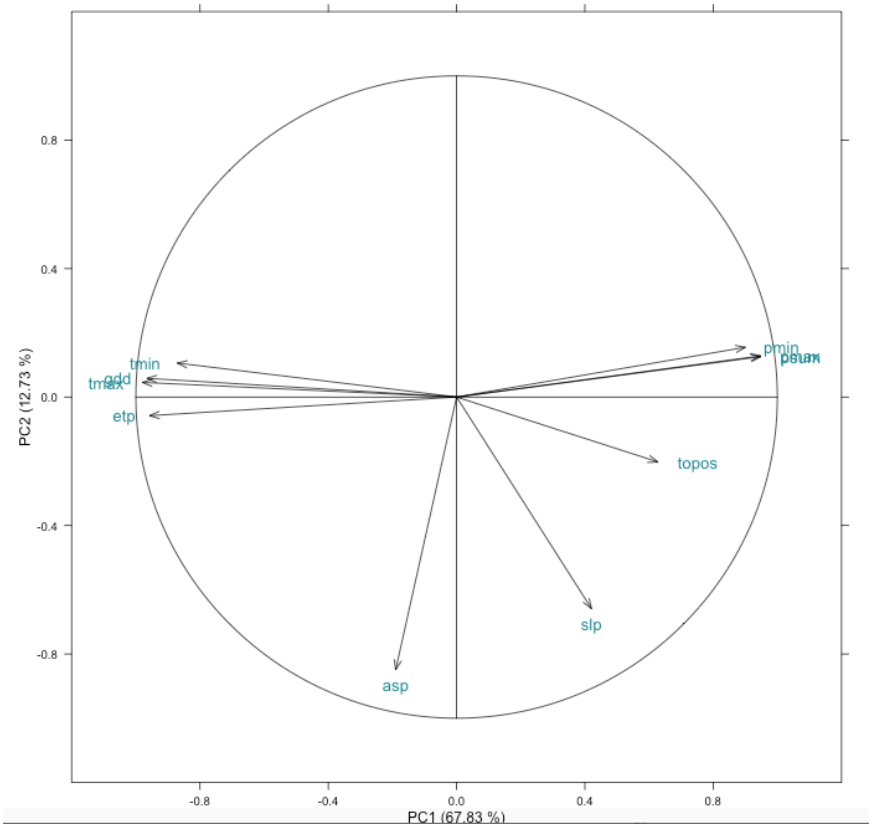
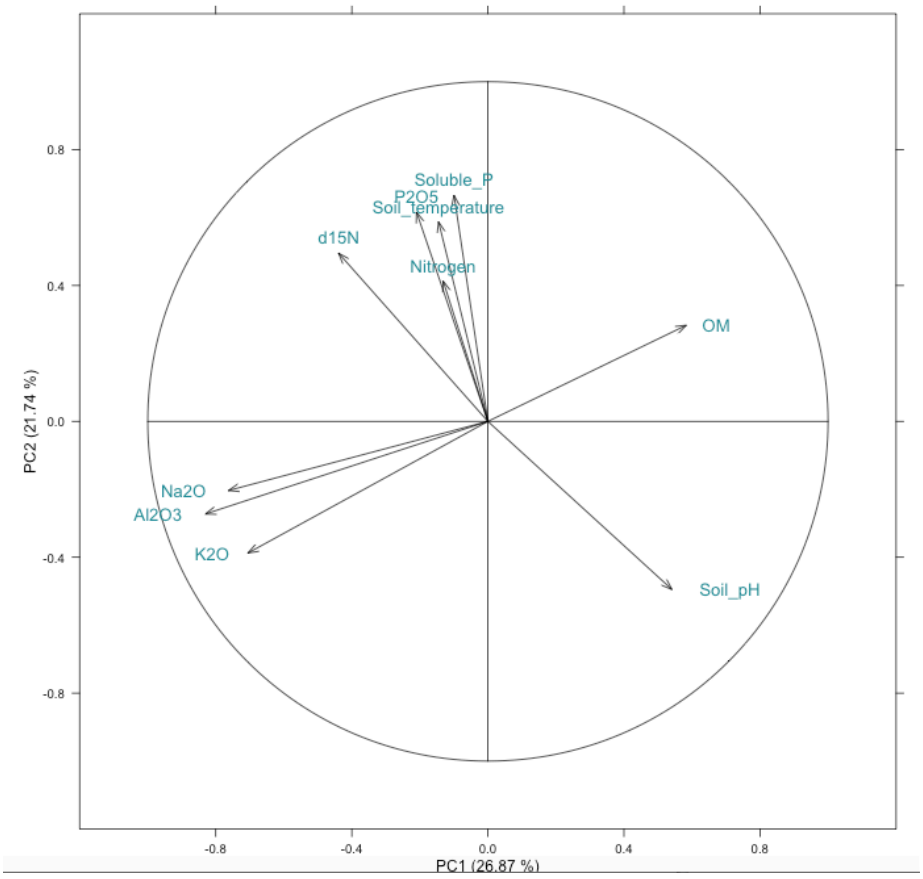


Figure S3b: PCA correlation circle of soil properties


Soluble_P
P2O5
Soil_temperature
d15N
Nitrogen
OM
Na2O
Al2O3
K2O
Soil_pH
PC2 (21.74 %)
PC1 (26.87 %)
0.8
0.4
0.0
-0.4
-0.8
-0.8
-0.4
0.0
0.4
0.8

727