1) Soil structure, nutrient status and water availability shape Uruguayan grassland microbiomes.

2) Garaycochea, Silvia[1]*; Romero, Héctor[2], Beyhaut, Elena[1], Neal, Andrew. L.[3] and Altier, Nora[1].

1-Instituto Nacional de Investigación Agropecuaria (INIA). Andes 1365, 11100, Montevideo, Uruguay.

2-Laboratorio de Organización y Evolución del Genoma/Unidad de Genómica Evolutiva, Departamento de Ecología y Evolución, Facultad de Ciencias/CURE, Universidad de la República, Maldonado, Uruguay.

3-Sustainable Agricultural Systems Department, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UK.

3) Corresponding author: Silvia Garaycochea, Andes 1365, 11100, Montevideo, Uruguay. Telephone number: +598 23657641-1784; fax number: +598 23677641. ORCID: https://orcid.org/0000-0003-1629-0746

4) Soil properties shape soil grassland microbiomes

Abstract: Soil microbial communities play a critical role in the functioning of ecosystems, as they influence important ecosystem processes. They are capable of metabolizing organic nutrients, thus modulating the availability of inorganic N, P, and S in soil. While Uruguayan soils are characterized by a high content of total P, they have low available P, and the organic P fraction is dominant (49-75%). The structure and diversity of bacterial communities in Uruguayan natural grassland were characterized by a 16s rRNA massive sequencing marker. *Proteobacteria, Actinobacteria, Verrucomicrobia, Planctomycetes, Chloroflexi* and *Firmicutes* were the predominant phyla, adding up to 90% of the community. Different multivariate methods detected clear discrimination between the communities from different soils. In particular, supp*Archaea* and bacterial phyla *Firmicutes*, *Acidobacteria*, *Actinobacteria* and *Verrucomicrobia* were different between sandy and clay soils. Clay Content, Available P, Soil Organic Carbon, Porosity, Ca and Total Available Water appear as major driving forces of bacterial community diversity and composition. These results reveal that the soil structure, nutrient status and water availability are strongly associated with bacterial community assemblage.

Phosphorus (P) is the second nutrient needed for agricultural activity following Nitrogen and is generally supplied to soils for agriculture in Latin America. Thus, the search for more efficient production systems in the use of P is of paramount importance. The intensification in the use of phosphate fertilizers in continuous agricultural systems constitutes a problem, because of contamination, particularly in the water streams. The huge environmental impact, the low economic efficiency and the fact that is a finite resource beg the need of using P in a more efficient way. Hence, given the fact that the soil microorganisms play key roles in the P cycle, mediating the availability of this plant nutrient. This work we seek to characterize how these communities are involved in the P mobilization and understanding the links among these

community's assemblage and the soil properties, that all together would be an effect on the P availability.

Keywords: microbial communities, soil nutrients, soil structure.

1 Introduction

Soil microbial communities play a critical role in the functioning of ecosystems, since they influence several important ecosystem processes including nutrient acquisition (Andreote *et al*., 2017; Fierer, 2017), carbon, phosphorus (P) and nitrogen (N) cycling, and soil formation (Van Der Heijden *et al*., 2008). Bacteria and Archaea account for a large proportion of soil microbiome biodiversity and are closely associated with biogeochemical cycles, energy flow and degradation of pollutants (Bardgett *et al*., 2014; Bodelier, 2011). Soil microbiomes are influenced by both biotic and abiotic factors (Griffiths *et al*., 2011; Xue *et al*., 2018) such as edaphic properties, temperature and moisture, as well as vegetation type. Soil pH and the amount of organic carbon, N and P are some of the most influential factors which determine

78 microbial assemblages (Fierer *et al*., 2006; Martiny *et al*., 2006). These factors set the context
79 for microbial interactions to occur, which leads to different assemblages and functions (Fanin *et*
80 *al*., 2016; Kinkel, *et al*., 2011; Garbeva, *et al*., 2004).
81
82 There is no typical soil microbiome; the abundance of bacterial and archaeal taxa may vary
83 considerably depending on soil type, land use and environmental conditions, as described
84 above (Fierer, 2017). However, there are apparent associations between abundant phyla, soil
85 type, and land use. For example, Neal et al, 2017 compared soil microbiome assemblages from
86 three different land uses (arable, bare fallow and grassland) and found *Gemmatimonadetes* and
87 *Armatimonadetes* associated particularly with degraded soil (Neal *et al*. 2017). Furthermore, in a
88 study of soil microbiomes of a pasture–rice rotation, bacterial and archaeal soil communities
89 were dominated by *Firmicutes* and *Proteobacteria* under pasture, whereas *Methanocellales* and
90 *Methanosarcinaceae* dominated under rice (Scavino *et al*. 2013).
91
92 In natural ecosystems, N, P and sulfur (S) are typically bound to organic molecules rendering
93 them unavailable to plants (Jacoby *et al*., 2017). Soil microorganisms are capable of
94 depolymerizing and mineralizing such organic forms, modulating the availability of inorganic N,
95 P, and S in soil, including ionic species such as ammonium, nitrate, phosphate, and sulfate, the
96 preferred nutrient forms for plants (Van Der Heijden *et al*., 2008; Richardson *et al*., 2011).
97 Uruguayan soils have high total P content (150-700 mg kg^{-1}), but most of it is associated with Fe
98 and Al, which renders the bioavailable P portion as relatively low (typically <10 mg kg^{-1}). Such
99 low levels of inorganic P typically found in soils are due to the high reactivity of the
100 orthophosphate (PO_4^{3-}) ion with calcium (Ca) in alkaline soils, and iron (Fe) and aluminium (Al)
101 in acidic soils (Gyaneshwar *et al*., 2002). Organic P is also unavailable for plants and in both
102 cases, enzymes are required to release orthophosphate to be accessible for plants. Organic P

represents a large part of the total P (50 -75%) (Hernández *et al*., 1995). The use of microorganisms as an alternative towards a more efficient production system poses a challenge that requires a deep understanding of the belowground ecosystem.

The parental material, the physicochemical properties and the evolution of Uruguayan soils are well described (Durán *et al*.,1999). In contrast, little is known about the resident microbial communities and how they relate to different soil types or land uses. The aim of this study was to characterize Uruguayan soil microbiomes under natural grasslands. Five representative types of soil were selected, with differential parental material and nutrient status, particularly P form retention and ratio P inorganic/ P organic.

2. Results and Discussion

2.1 Soil properties

General soil physicochemical characteristics are summarized in Tables 2 and 3. According to physicochemical analyses, the 50 samples collected from different soil units differed significantly in Ca, AP, CC, Po, AD and TAW. Ca ranged from 0.90 to 32.82 meq/100-g, AP ranged from <1 to 45 µg-P g^{-1}, CC varied from 12% (sandstone soil) to 47% (basalt soil), TAW ranged from 53 mm (crystalline soils) to 184 mm (basalt soils) (Table 2). As expected, based upon the sampling criteria used in this study, it was possible to identify specific soil properties related to each location.

L	Ca (meq/100g)	Mg (meq/100g)	K (meq/100g)	Na (meq/100g)	TA (meq/100g)	CEC (meq/100g)	TB (meq/100g)	% BT	pH	% N	SOC (%)	OM (%C*1,72)	APR (µg P/g)	APC (µg P/g)

ITA 1	14.53[a]	4.50[a]	0.43[a]	0.47 ab	6.77[a]	26.70 a	19.93[a]	74.66[a] b	5.50 a	0.37[a]	4.78 ab	8.22 ab	2.70 a	4.06 a
ITA 2	13.90[a]	5.90[a]	0.44[a]	0.35 ab	9.35[a]	29.94 a	20.59[a]	68.78[a] b	5.54 a	0.46[a]	6.16 ab	10.60[ab]	4.26 a	3.3 [a]
SPO1	3.45[b]	1.38[b]	0.54[a]	0.30 ab	5.99[ab]	11.66 b	5.67[b]	48.64[cd]	5.37 a	0.28[a]	2.69 ab	4.63 ab	16.08[ab]	21.6 ab
SPO2	3.14[b]	2.02[b]	0.44[a]	0.33 ab	7.18[ab]	13.10 b	5.93[b]	45.21[cd]	5.76 a	0.25[a]	4.57 ab	7.86 ab	0.144[ab]	3.6 ab
TBO1	0.90[b]	0.32[b]	0.11[a]	0.23 a	3.15[b]	4.72[b]	1.57[b]	33.24[c]	5.05 a	0.11[b]	0.79 a	1.35 ab	42.8 c	57.9 c
TBO2	1.0	0.3	0.05[a]	0.32	8.78[b]	4.4	1.7	39.2	5.3	0.35[b]	0.60 a	1.03 a	4.67 c	8.9 [c]
YNG1	32.82[C]	3.19[bc]	1.13[a]	0.33 ab	3.7[ab]	37.48 a	37.48[c]	100.00[b]	6.98 a	0.45[c]	6.88 b	11.83[b]	38.6 bc	31.08 [bc]
YNG2	27.21[C]	3.50[bc]	1.43[a]	0.36 ab	4.6[ab]	37.10 a	32.50[c]	87.61[b]	6.35 a	0.61[c]	8.19 b	14.09[b]	72.5 bc	95.7 bc
TRO1	9.59[a]	4.31[ac]	1.50[a]	0.81 a	9.74[ab]	25.95 a	16.21[a]	62.46[a]	5.40 a	0.35[a]	3.65 ab	6.28 ab	4.00 ab	3.8 ab

								d						
TRO2	13.95[a]	6.28[ac]	0.57[a]	1.18[a]	8.68[ab]	30.66[a]	21.98[a]	71.70[ad]	6.00[a]	0.32[a]	3.33[ab]	5.73[ab]	4.00[ab]	3.5[ab]

123

124

125

L	% CC	BD (g/cc)	Po.	PWP (mm/10cm)	TAW (mm)
ITA1	47.00[a]	0.93[a]	63.00[a]	16.90[ab]	184.90[a]
ITA2	26.85[a]	0.91[a]	60.00[a]	14.70[ab]	185.20[a]
SPO1	36.00[ab]	1.33[b]	50.00[b]	9.60[c]	65.50[b]
SPO2	29.00[ab]	1.26[b]	53.00[b]	11.10[c]	53.00[b]
TBO1	12.00[b]	1.42[d]	46.00[c]	4.30[d]	122.70[c]
TBO2	12.05[b]	1.47[d]	50.00[c]	4.80	118.90[c]
TRO1	12.50[b]	1.24[bd]	53.00[bd]	16.40[a]	106.60[c]
TRO2	18.10[b]	1.20[db]	55.00[bd]	12.50[a]	137.90[c]
YNG1	31.00[ab]	1.18[d]	56.00[d]	18.70[d]	178.10[a]
YNG2	25.00[ab]	1.18[d]	56.00[d]	18.70[d]	178.10[a]

2.2 Community analysis

Sequencing of 16S rRNA gene (V3-V4) amplicons resulted in a total of 6,188,081,507 sequences with an average length of 442 bases. High-quality reads from each soil sample were subsampled to 12,496 sequences (the number of sequences associated with the smallest sample). A total of 4,547 OTUs were obtained using a 97% identity threshold across the whole sample set. This set was reduced to 1,160 when considering OTUs with more than 200 sequences across the whole set. A total of 27 phyla were identified across all sites. *Proteobacteria* (26.6%), *Actinobacteria* (18.1%), *Firmicutes* (17%), *Verrucomicrobia* (14.2%), *Acidobacteria* (11.3%), *Planctomycetes* (1.9%) and *Chloroflexi* (1.5%) were the predominant phyla with a combined prevalence over 90%. (Supporting information in Fig. S1)

2.2.1 *Alpha-diversity*

Rarefaction curves showed a similar pattern for all samples from all sites, suggesting that sequencing had captured similar levels of diversity in each sample (Supporting information in Fig. S2). The highest values of alpha diversity were observed for ITA soil (PD: 56.23; S_{Chao1}: 50; *H*': 6.86;). The lowest *H*' values were found in the samples of SPO soil. However, when the PD was analyzed it was observed that the values of SPO were similar to those of ITA. This is indicative of a prokaryotic community with relatively divergent taxa. Different behavior was observed in TRO samples, with high *H*' values but low PD, indicating that the community of TRO is formed by phylogenetically closer taxa. On the other hand, the lowest values in YNG (PD: 38.03; S_{Chao1}: 28 *H*': 6.2) and TBO (PD: 45.38; S_{Chao1}: 16.9; *H*': 6.2) were consistent in the three alpha diversity indexes. The one-factor ANOVA and *post-hoc* Tukey HSD test showed significant differences ($\alpha = 0.01$) among the diversity values of each soil unit. The most

significant differences in the pairwise comparison was recorded between ITA and both TRO and SPO (Fig. 1 and Supporting information in Table S1).

2.2.2 *Beta-diversity*

Downstream analyses were performed using three phylogeny-sensitive distances: weighted UniFrac, KR-o distance and KR-r distance (see methods), in order to compare their power in recovering biologically meaningful patterns. First, weighted-UniFrac distance showed a clear discrimination between bacterial communities according to soil unit (Figure 2a), there was no significant heterogeneity of multivariate dispersion between the soils (*pseudo-F* = 2.3, p_{perm} = 0.195). PERMANOVA indicated a significant effect of soil unit upon the OTU assemblages (*pseudo-F* $_{=}$ 24.5, p_{perm} < 0.0001) and *post-hoc* pairwise comparisons indicated that all assemblages were significantly different from the others (smallest *pseudo-t* = 2.7, p_{perm} < 0.0001). Then, the KR-o distance based on OTUs identification (97%) and their respective abundance was used. Results were highly similar to the previous ones, they showed no significant heterogeneity of multivariate dispersion (*pseudo-F* = 2.8, p_{perm} = 0.1281); and *post-hoc* pairwise comparisons also indicated that all assemblages were significantly different from the others (smallest *pseudo-t* = 2.6, p_{perm} < 0.0001). Finally, was used KR-r distance based on phylogenetic placement of exact sequence variant (Figure 2b). These results showed a better discrimination intra groups compared with the two previous analyses based on distance matrices, even with no significant heterogeneity of multivariate dispersion (*pseudo-F*= 0.533, p_{perm} =0.88). PERMANOVA based on KR-r distance also showed significant differences between groups as an effect of soil unit (*pseudo-F* = 26.8, p_{perm} < 0.0001). The *post-hoc* pairwise comparison showed highest *pseudo-t* values (values between 3.2 to 6.9, p_{perm} < 0.0001). The differences between weighted UniFrac distance based upon grouping of OTUs with > 97% similarity and phylogenetic placement of exact sequence variants in their ability to

describe the assemblage dispersion indicates a greater sensitivity of the latter approach to differences in amplicon assemblages as suggested by Nguyen *et al*. (2016).

However, PCoA based upon KR-r distance accounted for a greater proportion of total variability (72.6%) than that based upon weighted-UniFrac distance using OTUs (60.6%) and based on KRo distance using OTU (63,1%), suggesting the increased power of non-OTU-based approaches. Is interesting to note that the widely used NMDS ordination based on phylogenetically uninformed Bray-Curtis dissimilarities (Figure 2c) with a stress value of 0.009, showed a similar distribution as the PCoA. In this sense, all respective axes of PcoAPCoA-WU, PcoAPCoA-KR-r, PCoA-KR-o, and NMDS-BC displayed high correlations values between them (Supporting information Fig. S3).

Edge-PCA analysis provides insight about which taxa hold differences between soil unit samples. 68% of the total 16S rRNA assemblage variation was explained by the first two edge-PCA axes (Fig. 2d). The microorganism assemblage's variation associated with the first axis separated soil types. Differences observed in the microbe communities between sandy soils and clay soils were associated primarily with a higher contribution of OTUs classified as *Archaea* and bacterial *Firmicutes*, *Acidobacteria, Actinobacteria* and *Verrucomicrobia* phyla (Supporting information in Fig. S4a). In particular, *Firmicutes* and *Acidobacteria* were overrepresented in samples from YGN and ITA soils. These two phyla are sensitive to changes in P content and C/N ratio (Karimi *et al*., 2018; Hermans *et al*. 2017). Karimi *et al* (2018) also reported that *Archaea* are sensitive to soil nutrient content. In fact, *Verrucomicrobia* and *Archaea* were characteristic of TBO soil, which had the lowest N and SOC content (Table 2).

The second edge-PCA axis was related to a higher relative abundance of *Proteobacteria* and a different *Verrucomicrobia* lineage (Supporting information in Fig. S4b), as detected in samples

from TBO and YNG soils. *Planctomycetes* and *Firmicutes* differentiated SPO soils. *Planctomycetes* have also been reported to show negative correlations with SOC and N content (Hermans *et al*., 2017; Lauber *et al*., 2008). *Chloroflexi* were overrepresented in YNG and TBO soils, whereas *Actinobacteria* were differential for SPO soils. SPO soil showed the lowest values in SOC; %N and AP were also relatively low among the five sites studied. The *Actinobacteria* phylum is involved in soil functions such as nutrients cycling and organic matter turnover (Lewin *et al*., 2017; Nasrabadi *et al*., 2013), and SOC and soil moisture affect the composition of *Actinobacteria* communities (Kopecky *et al*., 2011). All the above confirms the notion that the structure of soil microbial communities is strongly associated with soil characteristics (Lauber *et al*., 2008).

2.3 Relationship between microbial community phylogenetic structure and soil properties

Canonical Analysis of Principal Coordinates (CAP) based on weighted-UniFrac distance between microbial assemblages was chosen to perform the study of the relationship between phylogenetic composition and soil properties, given the widespread nature of this metric and the similar results using different b-diversity metrics (see above). Initially, all five soil units were included, regardless of their history and management (Figure 3a). Eight of the twenty-two environmental variables had correlation coefficients (r) > |0.20| with at least one of the first two CAP axes. CAP1 axis (canonical correlation [δ^2] = 0.999) was characterized by associations with available P as APR (r = 0.362) and APC (r = 0.534), CC (r =0.521), SOC (r = 0.354), Po (r = -0.207) and FC (r = -0.251), effectively separating TRO soils from the other sites on the basis of reduced SOC, CEC, CC, Ca, and AP. The highest correlation values with the CAP2 axis (δ^2 = 0.994) were Ca (r = -0.205), FC (r = -0.341) and TAW (r = -0.858) (Figure 3a), separating wetter soils (YNG and ITA) from the others. These results reveal that the nutrient status, water

availability and soil texture are associated with differences in bacterial community assemblages as previously reported (Brockett *et al*., 2018; Karimi *et al*., 2018; Delgado-Baquerizo *et al*. 2018).

The analysis showed that the TBO soil community composition was strongly associated with low CC, AP and SOC. The soil overlays sandstone and is characterized by low CC and low nutrient content (SOC and AP). *Verrucomicrobia* and *Archaea* were most associated with TBO soils. TBO soil properties are favorable for the proliferation of taxa that can adapt to limiting growth conditions of nutrients and water, such as *Verrucomicrobia*, which have highly flexible metabolism (Balmonte *et al*., 2016). Samples from YNG were associated with the highest nutrient values and a high capacity for nutrient exchange as suggested by CEC (Table 2). *Firmicutes* and *Acidobacteria* were the characteristic phyla of this soil. The main family of *Firmicutes* identified in the OTU taxonomic classification was *Baciliaceae*. While members of this family are widely distributed, they are more abundant where organic matter is plentiful (Mandic-Mulec *et al*., 2015). YNG soils are characterized by high CEC, and typically have high amounts of clay minerals and soil organic matter. In addition, they are soils with more capacity to retain nutrients with negative charge like N and P, thus decreasing their mobility and uptake. The land-use nearby the YNG unit has been predominantly agricultural over the last century. Although the sampling sites were not in a field with frequent fertilization, the nutrient content figures were consistent with those of fertilized soils. This nutrient contribution appears to have caused a significant change in the soil prokaryotic community.

Variation in TAW was associated with the CAP2 axis and was related to the community composition of ITA and YNG soil units. In addition, ITA soils have high Po, suggesting that these soils have more pore space associated with air and water that facilitates nutrient diffusion/advection and cell-cell communication. Recently, Borer et al (2018) showed though a

mathematical model how the pore network influences the spatial organization of soil microbes by considering nutrient and oxygen counter-gradients and cell motility. In particular, they showed that total bacterial abundance decreased with a reduction of pore network connectivity. The dynamics, composition, and distribution of soil microbes are shaped by heterogeneous water and resource distribution, and by their ability to rapidly adapt to dynamic changes in local conditions (Tecon et al., 2017). However, deeper analyses are necessary to understand how porosity, pore size distribution and pore connectivity influence environmental bacterial community assemblages (Borer *et al.*, 2018; Rabbi *et al*. 2016).

As mentioned above, ITA had two differentially phyla, *Firmicutes* and *Acidobacteria.* Although both phyla are widely distributed, little is known about the ecophysiology of *Acidobacteria*. Recent studies have characterized some of the families most frequently found in soil, *Solibacteraceae* and *Koribacteraceae* (Kielak *et al*., 2016). Although the OTUs classified as *Acidobacteria* were predominant in the samples from the TBO soil unit, this phylum was differentially detected in the ITA samples, particularly the *Koribacteracea* and *Solibacteraceae* families. Environmental factors influencing *Acidobateria* communities are thought to be pH and nutrient availability. In this work, the samples did not show significant differences in pH (Table 2), but nutrient content varied significantly. Eichorst *et al* (2018) reported that certain strains of the *Koribacteraceae* have a broad genome-coded metabolic potential with a high proportion of genes related to carbohydrate metabolism, which allows them to use a wide variety of carbohydrates as nutrient and energy sources. In addition, putative genes have been identified for these strains that encode for assimilatory nitrate reductase (NaR), nitrite reductase (NiR) and nitrate/nitrite transporter (NNP, TC: 2.A.1.8). New efforts are needed to understand how the *Acidobacteria* metabolism functions under the environmental conditions of this work, in order to characterize the soil properties, chemical forms of nutrients in the soil and the adaptability of this phylum's metabolism to such conditions.

CAP analysis demonstrated a strong influence of AP and CEC upon the TBO prokaryotic assemblage. This result matched our expectations based on the relations between nutrient availability and CEC in TBO sandy soils. Higher TAW values influenced ITA communities, where the differential phyla were *Acidobacteria* and *Firmicutes.* In addition, ITA soils have high Po, suggesting that these soils have more pore space associated with air and water, which facilitates nutrient diffusion/advection and cell-cell communication. However, deeper analysis is needed to understand how porosity, pore size distribution and pore connectivity influence environmental bacterial community assemblages (Borer *et al.*, 2018; Rabbi *et al*. 2016).

A second CAP analysis using weighted-UniFrac distance was performed removing YNG, due to the impact that may cause anthropogenic fertilization. Eight physical and chemical soil properties were identified with r > |0.20|. Most of these variables were consistent with the previous analysis including YNG (Supporting information Table S2 for details). TAW stood out with the highest correlation value with axis CAP1 (r = -0.967). Axis CAP2 show correlations with AP (r = 0.602), Po (r = -0.446), CEC (r = 0.339) (Figure 3b). A clear-cut separation was observed between sites differing in the type of soil. TBO soil communities, developed over sandstone parent material, are particularly different from the communities belonging to other soils. When YNG samples were removed from the analysis, a clearer association with high values of nutrient availability and CEC (Table 2 and 3) was apparent in TBO sandy soils. In this line, higher TAW values were associated with the ITA communities, as in the previous CAP analyses. When the CAP results using KRr and KRo dissimilarity matrices were analyzed, they were in general agreement with those of weighted-UniFrac. Most of the selected variables were the same, albeit some inconsistencies were found (Supporting information table S2 for details). The most significant observations were that both, TAW and Po were detected as the variables with the largest correlation coefficients. Conversely, AP diminished its r values, whilst CEC and

CC increased them. Despite these changes, ITA is still associated with high TAW values and the trend of TBO with nutrient availability and CEC holds. Nevertheless, TRO present and opposite behavior when using either KR distance. These results show that these methods are generally robust to different approaches (OTU clustering or global sequence data) and distance metrics (weighted-UniFrac or KR) between communities. Nevertheless, at some point some inconsistencies may appear and it is important to explore them, with respect to the biological interpretations and verify if these are altered.

OTUs with differential abundance in sites under natural grassland (ITA, SPO, TBO and TRO, excluding YNG) were identified using the DESeq2 R-package (Love *et al*., 2014). Twenty-nine OTUs were reported to have significantly differing abundance between the four sites using TAW and AP as factors ($p_{adj} < 0.05$). OTUs were classified taxonomically using SILVA (Supporting information in Table S2) and most of them were also detected by edge-PCA analysis, namely *Verrucomicrobia*, *Firmicutes*, *Chloroflexi*, *Actinobacteria* and *Planctomycetes*. The OTUs of three orders of *Actinobacteria* are of interest: *Acidimicrobiales*, *Gaiellales* and *Solirubrobacteriales*. Free-living *Actinobacteria* are especially abundant in alkaline and organic matter-rich soil (Barka *et al*., 2016) and play a key role in enhancing soil health. Some members of this phylum have the ability to degrade organic compounds from different sources, such as decaying plant material, chitin and hydrocarbons (Lewin *et al.*, 2017; Sharma *et al.*, 2014), thus contributing to carbon cycling. It has also been reported that some members of *Actinobacteria* have variable responses in the production of acid and alkaline phosphatases, which may release P from organic sources (Nasrabadi *et al*., 2013). It is interesting to note that these OTUs are not detected as differentially abundant when including P-enriched YNG samples as a result of fertilization. In this sense, it has been shown how fertilization impacts on the composition of soil microbial communities (Jangid *et al*., 2008; Wang *et al*., 2019).

In order to compare the results of DESeq2 analysis with the above edge-PCA analysis, a second analysis without YNG samples was performed. On such occasion, *Actinobacteria* appears in the edge-PCA analysis as one of the differential phyla between sites. These results were consistent with other works that report how chemical fertilization changes the structure and functions of bacterial communities by altering nutrients balance, organic matter content and edaphic properties such as pH (Jangid *et al*., 2008; Kopecky et al., 2011; Lauber *et al*., 2008; Wang *et al*., 2019).

In summary, Using the 16S rRNA gene were described archaeal and bacterial communities in five soil units with characteristically different physical and chemical properties. These properties influenced the assemblages of the communities making differential profiles for each soil. As described above, *Proteobacteria*, *Actinobacteria*, *Verrucomicrobia*, *Planctomycetes*, *Chloroflexi* and *Firmicutes* were the predominant phyla, accounting for as much as 90% of the total community. *Solirubrobacteriales*, *Roseiflexaceae*, *Rhodomicrobiaceae*, *Haliangiaceae* were among the most relevant taxa which contribute to the differences between the five soils. Soil structure, nutrient status and available water were responsible for the differential prokaryotic community assemblage observed for each soil. Such differences may be linked to differential sets of metabolic functions from each community responding to the different conditions of nutrients, water and porosity.

3. Experimental procedures

3.1 Soil collection

Five Uruguayan soil units were selected as representative of different agroecological regions (Hernández *et al*., 1995; Hernández *et al*., 1998). The main criterion for soil unit classification

was based on parental materials: basalt for Itapebí Tres Árboles (ITA), crystalline basement for Sierra de Polanco (SPO), sandstone for Tacuarembó soils (TBO), and tertiary silt for both Tala Rodríguez (TRO) and Young (YNG). Selected soils have different ratios of organic-P and inorganic-P, as well as different mechanisms for inorganic-P retention, associated with Fe, Al or Ca. Four soil units consisted of natural grassland ecosystems (ITA, SPO, TBO, TRO), whereas YNG was close to an agricultural management region. A description of the five soils is presented in Table 1.

Soil Unit	**CODE**	**Parental Material**	**Soil Type (USDA)**	**Land-use**
Itapebí Tres Arboles	ITA	Basalt	Argiudoll Pachic, smectitic, fine, thermic.	Natural grassland
Sierra de Polanco	SPO	Crystalline	Argiudoll Typic (shallow), Fine-loamy, superactive, mixed, thermic	Natural grassland
Tacuarembó	TBO	Sandstone	Hapludalf Typic, Fine-loamy (coarse), siliceous, active, thermic	Natural grassland
Tala-Rodriguez	TRO	Tertiary silt	Natralbaquolls Glossic, superactive, mixed, fine, thermic.	Natural grassland
Young	YGN	Tertiary silt	Argiudoll Pachic, fine,	Agricultural

	superactive, mixed, thermic	ecosystem

369

370

371

372 *3.2 Sampling methodology*

373 Five replicates of each soil were collected and geo-referenced in two locations during autumn
374 2015. Each replicate was the aggregated soil from 15 samples taken with a 3cm diameter core
375 to a depth of 10cm (effectively the A Horizon). Replicates were spaced 3m apart. Soil samples
376 were transported to the laboratory at 4 °C and sieved through a 2-mm mesh to remove roots
377 and plant detritus (within three days of sampling). Sieved soils were stored at -20 °C until
378 nucleic acid extraction.

379

380 *3.3 Soil properties*

381 Ten soil locations were characterized by their physicochemical properties. Soil total nitrogen (N)
382 was determined by combustion at 900 ºC and subsequent N_2 thermal conductivity detection;
383 available phosphorus (APR) was determined by the resin membrane technique (Sharpley *et al*.,
384 1994) and Citric Acid by colourimetric method (APC) (Murphy and Rilley, 1962); available
385 potassium (K) and available sodium (Na) were determined by ammonium acetate (pH 7)
386 extraction followed by atomic emission spectrometry; Ca and Mg were determined by
387 ammonium acetate (pH 7) extraction followed by atomic absorption spectrometry. Soil pH was
388 measured by potentiometric determination in water. Soil organic carbon (SOC) was determined
389 by combustion at 900 ºC and subsequent CO_2 infrared detection. The cation exchange capacity
390 (CEC) was determined by acid-base titration. Soil granulometric composition was determined
391 and physical parameters were calculated, including aeration (Po) and available water retention
392 capacity (TAW). Clay content (CC) was determined by the hydrometric method (Gee *et al*.,

1986). Analysis of variance (ANOVA) and Tukey HSD test was applied to pairwise comparison. All basic statistical procedures were performed using R-base (R-core Team, 2018).

3.4 DNA extraction and metagenome sequencing

Metagenomic DNA was extracted from 0.25 g aliquots of soil using the Power Soil DNA Isolation kit (Qiagen) following the manufacturer's protocols. The V3 – V4 region of the 16S rRNA gene was amplified by PCR with the following primers: forward 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG and reverse 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC selected from Klindworth et al. (2013). Amplicons were sequenced on an Illumina MySeq platform, generating 300-base paired-end reads.

3.5 Sequence and statistical analysis

Raw Illumina sequence data was pre-processed with the Microbiome Helper pipeline (Comeau *et al*., 2017), and the operational taxonomic unit (OTU) identification and taxonomic assignation were performed using the Quantitative Insights into Microbial Ecology pipeline (QIIME v. 1.9.0) (Caporaso *et al*., 2011). Sequences were quality-filtered. The only sequences included in subsequent analyses were those with a minimum quality score of 30 in at least 75% of the sequence length, containing no ambiguous bases, and with no more than 10 consecutive low-quality base pairs and one base mismatch, Paired-end reads were joined, generating an average read length of 590 bases. Paired reads and reads with length under 400 bp were removed. Chimeric sequences were also removed. Quality filtered reads were used to perform downstream analyses. Operational taxonomic units (OTU) were assigned using the open-reference method (Navas-Molina *et al*, 2015). Sequences were clustered into OTUs using a

97% sequence similarity based on the UCLUST classifier (Edgar, 2010). Representative sequences were aligned to the Greengenes 13-8 reference database (DeSantis *et al*., 2006) with PyNAST (Caporaso *et al*., 2010). A maximum-likelihood 16S rRNA phylogenetic tree was constructed with RAxML 7.0.4 software using default settings and the GTR-model (Stamatakis, 2006). This was manually edited and plotted with iTOL (Letunic and Bork 2019). Taxonomy was assigned using the USEARCH v. 7.0 (Edgar, 2013) based upon a 90% confidence threshold and the Greengenes phylogeny. The resulting OTU table was filtered using a minimum cluster size of 0.1% of the total reads (Bokulich *et al*., 2013).

Rarefaction curves and observed species were calculated using QIIME (Caporaso *et al*., 2011). The Chao1 abundance-based estimator of species richness (S_{Chao1}) and Shannon entropy (*H*') were calculated with the Vegan R package (Oksanen *et al*., 2008). Phylogenetic Diversity (PD) was calculated using the Picante R package (Kembel, 2010). *Post-hoc* Tukey's HSD tests were carried out for each diversity index to identify significant differences in alpha-diversity estimates between treatments.

Different beta-diversity measures were computed for comparison. First, based on the identified OTUs, Bray-Curtis distance, weighted UniFrac (Lozupone *et al*., 2007), and Kantorovich-Rubinstein (here named KR-o) (Evans and Matsen, 2012) were computed. Second, pplacer (Matsen *et al.*, 2010) was used for the phylogenetic placement of exact sequence variants (reads) and these results were used to compute the Kantorovich-Rubinstein distance (KR-r). In addition, an edge-PCA (Matsen and Evans, 2013) analysis was performed also on these results. KR distances and edge-PCA were performed using guppy (Matsen and Gallagher, 2011).

Comparisons of community assemblages using the different distance metrics were first tested for heteroscedasticity using PERMDISP (Anderson, 2006). Permutational multivariate analysis

of variance (PERMANOVA) was used to test assemblage differences between different soils, and pairwise comparisons were performed in those cases where significant treatment effect was identified *post hoc*. NMDS was performed using the Bray Curtis metric as implemented in the Vegan R package (Oksanen *et al*., 2008). Principal coordinates analysis (PCoA) and canonical analysis of principal coordinates (CAP) were performed using weighted-UniFrac and KR metrics to calculate the correlation between physicochemical properties and microbial communities. PERMDISP, PERMANOVA, PCoA and CAP were performed using PRIMER PERMANOVA+ ver7.0.13 (PRIMER-*e*, Auckland, New Zealand) and 99,999 permutations where required. Graphics were produced with the R package ggplot2 (Wickham, 2016) and Archaeopteryx tree viewer (Zmasek, 2012).

Estimation of differentially abundant OTUs was performed using DESeq2 (Love *et a*l., 2014), using a two–factor model, TAW and AP, with no interaction term. These two factors were selected based on CAP results. Differential OTUs were classified using the SILVA 132 16s rRNA database (Quast *et al*., 2013). R-base (R-core Team) was used to determine the correlation between TAW, AP and differential OTUs abundances. All basic statistical procedures were done using R-base (R-core Team, 2018).

Acknowledgments

This work was funded by INIA (Project SA_26 and SA_24), Agencia Nacional de Investigación e Innovación (POS_NAC_2015_1_110075). ALN is supported by Rothamsted Research's Soil to Nutrition strategic programme which is funded by the U.K.'s Biotechnology and Biological Science Research Council (BBS/E/C/000I0310) and jointly by the Natural Environment

Research Council and BBSRC as part of the Achieving Sustainable Agricultural Systems research programme (NE/N018125/1 LTS-M). We acknowledge Carolina Leoni, Victoria Bonnecarrère, and Gastón Quero for their contribution to the discussion. Marco Dalla Rizza for his continuous support for this work. We thank the soil microbiology laboratory group and the field team of INIA. We also thank the anonymous reviewers for their comments that substantially improved the manuscript

References

Anderson, M. J. (2006). Distance-Based Tests for Homogeneity of Multivariate Dispersions, (March), 245–253. https://doi.org/10.1111/j.1541-0420.2005.00440.x

Andreote, F. D., & Pereira E Silva, M. de C. (2017). Microbial communities associated with

plants: learning from nature to apply it in agriculture. *Current Opinion in Microbiology*, *37*, 29–34. https://doi.org/10.1016/j.mib.2017.03.011

Balmonte, J. P., Arnosti, C., Underwood, S., & Mckee, B. A. (2016). Riverine Bacterial Communities Reveal Environmental Disturbance Signatures within the Betaproteobacteria and Verrucomicrobia, *7*(September), 1–14. https://doi.org/10.3389/fmicb.2016.01441

Bardgett, R. D., & Putten, W. H. Van Der. (2014). Belowground biodiversity and ecosystem functioning. *Nature*, *515*(7528), 505–511. https://doi.org/10.1038/nature13855

Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-vaillant, N., Jacquard, C., Klenk, H., Wezel, P. Van. (2016). Taxonomy , Physiology , and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews : MMBR*, *80*(1), 1–44. https://doi.org/10.1128/MMBR.00019-15.Address

Borer, B., Tecon, R., & Or, D. (2018). Spatial organization of bacterial populations in response to oxygen and carbon counter-gradients in pore networks. *Nature Communications*, https://doi.org/10.1038/s41467-018-03187-y

Brockett, B., Brockett, B. F. T., Prescott, C. E., & Grayston, S. J. (2018). Patterns in forest soil microbial community composition across a range of regional climates in Western Canada Soil Biology & Biochemistry Soil moisture is the major factor in fl uencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. *Soil Biology and Biochemistry*, *44*(1), 9–20. https://doi.org/10.1016/j.soilbio.2011.09.003

Caporaso, J. G., Bittinger, K., Bushman, F. D., Desantis, T. Z., Andersen, G. L., & Knight, R. (2010). PyNAST: A flexible tool for aligning sequences to a template alignment. *Bioinformatics*, *26*(2), 266–267. https://doi.org/10.1093/bioinformatics/btp636

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Gonzalez-Peña A, Goodrich JK, G. J. et. (2011). QIIME allows analysis of high-throughput community sequencing data

Comeau, A. M., Douglas, G. M., & Langille, M. G. I. (2017). Microbiome Helper: a Custom and Streamlined Workflow for Microbiome Research. *MSystems*, *2*(1), 1–11. https://doi.org/10.1128/mSystems.00127-16

Dequiedt, S., Wincker, P., Jolivet, C., Cruaud, C., Karimi, B., Arrouays, D., Terrat, S. (2018). Biogeography of soil bacteria and archaea across France. *Science Advances*, *4*(7), eaat1808. https://doi.org/10.1126/sciadv.aat1808

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G.L. (2006)G.reengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB.Appl Environ Microbiol.72(7):5069-72.

Durán, A., Caligra, A., & Molfino, J. H. (1999). Suelos del uruguay según soil taxonomy (1999), 1–14. Carta de Reconocimiento de Suelos del Uruguay - Escala 1:1000000, MGAP - http://www.mgap.gub.uy/unidad-organizativa/descarga/carta-de-reconocimiento-de-suelos-del-uruguay-11000000;

Bodelier, P. L. (2011). Toward understanding, managing, and protecting microbial ecosystems. https://doi.org/10.3389/fmicb.2011.00080

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*(10), 996–998. https://doi.org/10.1038/nmeth.2604.

Eichorst, S., Trojan, S., Roux, S., Herbold, C., Rattei,T., Woebken, D. (2018).Genomic insights into the Acidobacteria reveal strategies for their success in terrestrial environments.Environ Microbiol.20(3): 1041–1063. doi: 10.1111/1462-2920.14043

Evans, S.N. and Matsen, F.A(2012). The phylogenetic Kantorovich–Rubinstein metric for environmental sequence samples. J R Stat Soc Series B Stat Methodol.1; 74(3): 569–592. doi: 10.1111/j.1467-9868.2011.01018.x

Fanin, N., & Bertrand, I. (2016). Soil Biology & Biochemistry Aboveground litter quality is a

better predictor than belowground microbial communities when estimating carbon mineralization along a land-use gradient. *Soil Biology and Biochemistry*, *94*, 48–60. https://doi.org/10.1016/j.soilbio.2015.11.007

Fierer, N. (2017). Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, *15*(10), 579–590. https://doi.org/10.1038/nrmicro.2017.87

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 626–631. https://doi.org/10.1073/pnas.0507535103

Garbeva, P., van Veen, J. A., & van Elsas, J. D. (2004). MICROBIAL DIVERSITY IN SOIL: Selection of Microbial Populations by Plant and Soil Type and Implications for Disease Suppressiveness. *Annual Review of Phytopathology*, *42*(1), 243–270. https://doi.org/10.1146/annurev.phyto.42.012604.135455

Gee, G.W., and Bauder, J.W. (1986). Particle size analysis. In A. Klute, Ed., Methods of Soil Analysis, 2nd ed, part 1, Madison, WI: American Society of Agronomy, p. 383

Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., Whiteley, A. S., Gifford, C. (2011). The bacterial biogeography of British soils, *13*, 1642–1654. https://doi.org/10.1111/j.1462-2920.2011.02480.x

Gyaneshwar, P., Kumar, G. N., Parekh, L. J., & Poole, P. S. (2002). Role of soil microorganisms in improving P nutrition of plants Author (s): P . Gyaneshwar , G . Naresh Kumar , L . J . Parekh and P . S . Poole Source : Plant and Soil , Vol . 245 , No . 1 , FOOD SECURITY IN NUTRIENT-STRESSED ENVIRONMENTS : EXPLOITING P. *Plant and Soil*, *245*(1), 83–93.

Hernández, J.; Otegui, O.; Zamalvide, J.P. 1995. Formas y contenidos de fósforo en algunos suelos del Uruguay. Boletín de Investigaciones No.43. Facultad de Agronomía, Universidad de la República. 32p

Hermans, S. M., Buckley, H. L., Case, B. S., Curran-Cournane, F., Taylor, M., & Lear, G. (2017). *Bacteria as emerging indicators of soil condition*. *Applied and Environmental Microbiology* (Vol. 83). https://doi.org/10.1128/AEM.02826-16

Jacoby, R., Peukert, M., Succurro, A., Koprivova, A., & Kopriva, S. (2017). The Role of Soil Microorganisms in Plant Mineral Nutrition—Current Knowledge and Future Directions. *Frontiers in Plant Science*, *8*(September), 1–19. https://doi.org/10.3389/fpls.2017.01617

Jangid, K., Williams, M. A., Franzluebbers, A. J., Sanderlin, J. S., Reeves, J. H., Jenkins, M. B., Whitman, W. B. (2008). Soil Biology & Biochemistry Relative impacts of land-use , management intensity and fertilization upon soil microbial community structure in agricultural systems. *Soil Biology and Biochemistry*, *40*(11), 2843–2853. https://doi.org/10.1016/j.soilbio.2008.07.030

Navas-Molina; Peralta-Sánchez, J., González, A., McMurdie, P., Vázquez-Baeza, Y., Xu, Z., Ursell, L., Lauber, C., Zhou, Z., Song, S., Huntley, J., Ackermann, G., Berg-Lyons, D., Holmes, J. Gregory , J. G. C. and Knigh R. (2015). *HHS Public Access*. https://doi.org/10.1016/B978-0-12-407863-5.00019-8.Advancing

Kembel, S. (2010). An introduction to the picante package. *R Project*, (April), 1–16. https://doi.org/10.1093/bioinformatics/btq166

Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., Veen, J. A. Van, & Kuramae, E. E. (2016). The Ecology of Acidobacteria : Moving beyond Genes and Genomes. *Frontiers in Microbioloy*, *7*(May), 1–16. https://doi.org/10.3389/fmicb.2016.00744

Kinkel, L. L., Bakker, M. G., & Schlatter, D. C. (2011). A Coevolutionary Framework for Managing Disease-Suppressive Soils. https://doi.org/10.1146/annurev-phyto-072910-095232

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, *41*(1), e1.

https://doi.org/10.1093/nar/gks808

Kopecky, J., Kyselkova, M., Omelka, M., Ladislav, c., Jitka, N., Grundmann, G., & Moe, Yvan, Moenne-Loccoz, Y., & Sagova-Mareckova, M. (2011). Actinobacterial community dominated by a distinct clade in acidic soil of a waterlogged deciduous forest. *Federation of European Microbiological Societies*. https://doi.org/10.1111/j.1574-6941.2011.01173.x

Lauber, C. L., Strickland, M. S., Bradford, M. A., & Fierer, N. (2008). The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biology and Biochemistry*, *40*(9), 2407–2415. https://doi.org/10.1016/j.soilbio.2008.05.021

Letunic I and Bork P (2006) Bioinformatics 23(1):127-8 Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation

Lewin, G., Carlos, C, Chevrette, M., Horn, H., McDonald, B., Stankey, R., Fox, B., Currie, B. (2017). Evolution and Ecology of Actinobacteria and Their Bioenergy Applications. *ANnual Rev. Microbiol.*, (58), 235–254. https://doi.org/10.1146/annurev-micro-102215-095748.Evolution.

Love, M.I., Huber, W., Anders, S. (2014). “Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.” *Genome Biology*, 15, 550. doi: 10.1186/s13059-014-0550-8.

Lozupone, C. A., Hamady, M., Kelley, S. T., & Knight, R. (2007). Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology*, *73*(5), 1576–1585. https://doi.org/10.1128/AEM.01996-06

Delgado-Baquerizo,M., Oliverio, A., Brewer T., Benavent-González, A. Eldridge, D., Bardgett, R., Maestre, F., Singh, B., 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

Mandic-Mulec, I., Stefanic, P., & Elsas, (2015). Ecology of Bacillaceae. *American Society for Microbiology Press*, (April). https://doi.org/10.1128/microbiolspec.TBS-0017-2013

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Kane, M., Krumins, J. A., Kuske, C. R., Smith, V. H. (2006). Microbial biogeography : putting microorganisms on the map, *4*, 102–112. https://doi.org/10.1038/nrmicro1341

Matsen, F. A., Kodner, R. B., & Armbrust, E. V. (2010). pplacer: linear time maximum-likelihood and Bayesian phy- logenetic placement of sequences onto a fixed reference tree, 1–28. Retrieved from papers2://publication/uuid/C99CC177-956E-406D-9FEB-4A88AFD24F63

Matsen, F.A., Hoffman,N.G., Gallagher, A., Stamatakis, A. (2012). A Format for Phylogenetic Placements. PLoS ONE 7(2): e31009. doi:10.1371/journal.pone.0031009

Matsen, F.A. and Evans, S.N. (2013). Edge Principal Components and Squash Clustering: Using the Special Structure of Phylogenetic Placement Data for Sample Comparison. PLoS One. 2013; 8(3): e56859. doi: 10.1371/journal.pone.0056859

Murphy, J. and J.P Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. Anal. Chim. Acta. 27: 31-36.

Nasrabadi, R. G., Resources, N., Greiner, R., & Hamedi, J. (2013). Distribution of actinomycetes in different soil ecosystems and effect of media composition on extracellular phosphatase activity Distribution of actinomycetes in different soil ecosystems and effect of media composition on extracellular phosphatase activity. https://doi.org/10.4067/S0718-95162013005000020

Neal, A. L., Rossmann, M., Brearley, C., Akkari, E., Guyomar, C., Clark, I. M., … Hirsch, P. R. (2017). Land-use influences phosphatase gene microdiversity in soils. *Environmental Microbiology*, *19*(7), 2740–2753. https://doi.org/10.1111/1462-2920.13778

Nicholas A. Bokulich, Sathish Subramanian, Jeremiah J. Faith, Dirk Gevers, Jeffrey I. Gordon, Rob Knight, David A. Mills, and J. G. C. (2013). HHS Public Access Patients, *10*(1), 1–16. https://doi.org/10.3109/21678421.2014.980428.Depression

Nguyen, N., Warnow, T., Pop, M. White, B. (2016) A perspective on 16S rRNA operational taxonomic unit clustering using sequence similarity. npj Biofilms Microbiomes 2, 16004

653 doi:10.1038/npjbiofilms.2016.4
654 Oksanen, A. J., Kindt, R., Legendre, P., Hara, B. O., Simpson, G. L., Stevens, M. H. H.,
655 Solymos, P. M. (2008). The vegan package. *The Vegan Package*, (May 2014), 190.
656 Retrieved from http://cran.r-
657 project.org/,%0Ahttps://bcrc.bio.umass.edu/biometry/images/8/85/Vegan.pdf
658 Quast,C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner,
659 F.O. (2013) The SILVA ribosomal RNA gene database project: improved data processing
660 and web-based tools. Nucl.
661 R Core Team. 2018. R: A language and environment for statistical computing. The R
662 Foundation. https://www.R-project.org (accessed October 2019).
663
664 Rabbi, S. M. F., Daniel, H., Lockwood, P. V, Macdonald, C., Pereg, L., & Tighe, M. (2016).
665 Physical soil architectural traits are functionally linked to carbon decomposition and
666 bacterial diversity. *Nature Scientific Reports*, (September), 1–9.
667 https://doi.org/10.1038/srep33012
668 Richardson, A. E., & Simpson, R. J. (2011). Soil Microorganisms Mediating Phosphorus
669 Availability, *156*(July), 989–996. https://doi.org/10.1104/pp.111.175448
670 Scavino, A. F., Ji, Y., Pump, J., Klose, M., Claus, P., & Conrad, R. (2013). Structure and
671 function of the methanogenic microbial communities in Uruguayan soils shifted between
672 pasture and irrigated rice fields. *Environmental Microbiology*, *15*(9), 2588–2602.
673 https://doi.org/10.1111/1462-2920.12161
674 Sharma, M., Dangi, P., & Choudhary, M. (2014). Actinomycetes : Source , Identification , and
675 Their Applications. *International Journal of Current Microbiology and Applied Sciences*,
676 *3*(2), 801–832.
677
678 Sharpley AN, Sims JT, Pierzynski GM (1994). Innovative soil phosphorus indices:

Assessing inorganic phosphorus. In: J. Havlin et al. (ed) Soil testing: Prospects for improving nutrient recommendations. Soil Sci Soc. Am Spec. Pub. 40, ASA, Madison. WI. pp 115 – 142

Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, *22*(21), 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Tecon, R., & Or, D. (2017). Biophysical processes supporting the diversity of microbial life in soil. *FEMS Microbiology Reviews*, (November). https://doi.org/10.1093/femsre/fux039

Van Der Heijden, M. G. A., 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)). *Ecology Letters*, *11*(6), 651. https://doi.org/10.1111/j.1461-0248.2008.01199.x

Wang, Z., Liu, Y., Zhao, L., Zhang, W., & Liu, L. (2019). Change of soil microbial community under long-term fertilization in a reclaimed sandy agricultural ecosystem, 1–21. https://doi.org/10.7717/peerj.6497

Wickham, H. (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Xue, P., Carrillo, Y., Pino, V., Minasny, B., & Mcbratney, A. B. (2018). Soil Properties Drive Microbial Community Structure in a Large Scale Transect in South Eastern. *Scientific Reports*, 1–11. https://doi.org/10.1038/s41598-018-30005-8.

Tables and Figures legends.

Table 1: Soil unit characteristics.

Table 2. Chemical properties of soils analyzed in this study: Available P by resin method (APR), Available P by Citric acid method (APC), Locations (L), Soil organic carbon (SOC), Organic Material (OM), %Base Saturation (%BS), Titratable Acid (TA).

Table 3. Physical properties of soils analyzed in this study: Bulk Density (BD), Clay Content (CC), Permanent Wilt Point (PWP), Porosity (Po), Total Available Water (TAW).

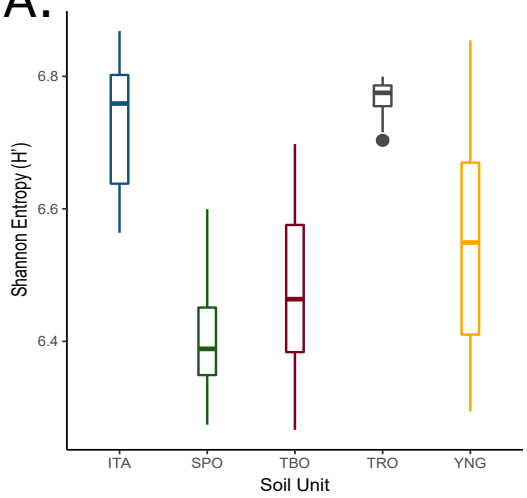
Figure 1: Alpha diversity indices box plot: A) *H*'; B) S_{Chao1} species richness; C) Phylogenetic Diversity

Figure 2: A) PCoA plot based on weighted UniFrac distances between prokaryotic communities of each soil unit: Each point represents replicates from ITA (blue and light blue), SPO (green and light green), TBO (red and light red), TRO (grey and light grey), and YNG (orange and light orange).

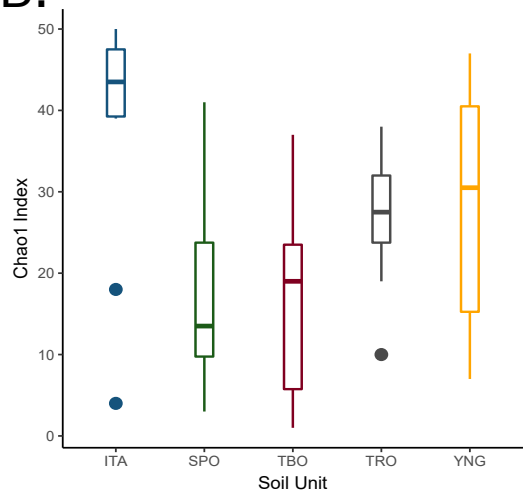
– The same color code was used for all figures (n = 10 for all soils). The percent variation explained by each principal coordinate is indicated on the axes. There was no significant difference in assemblage dispersion (PERMDISP; *pseudo-F*= 2.3; p_{perm} = 0.195), but there was significant difference in OTU assemblage (PERMANOVA; *pseudo-F*=24.5; p_{perm} < 0.0001) between soils; B) PCoA plot based on KR-r distance between prokaryotic communities of soils. In this case, there were significant differences between assemblage dispersion (*pseudo-F*= 3.6; p_{perm} = 0.029) and in community assemblages (*pseudo-F* = 4.4; p_{perm} < 0.0001) in different soils; C) Nonmetric multidimensional scaling (NMDS) plot of pairwise Bray-Curtis dissimilarities

731 between microbial communities of the studied soil units. Ellipses represent 95% confidence
732 intervals; non-overlapping centroids are significant at α = 0.05. Stress = 0.009 Respective
733 PERMDISP (*F*-value = 8.07; $p = 4.6x10^{-5}$) and PERMANOVA ($R^2 = 0.7098$; $p < 0.001$) D)
734 Ordination of soil units based on edge-PCA analysis.
735
736 Figure 3: Canonical Analysis of Principal Coordinates (CAP) A) based on weighted-UniFrac
737 distance for bacterial communities in the different soils. PERMANOVA analysis with 99,999
738 permutations was performed to determine the significance between microbial communities of
739 five soil units with two locations in each one (*n* = 10) and soil physicochemical properties. B)
740 based on KR-r distance. C) based on weighted-UniFrac distance without YNG and D) based on
741 KR-r without YNG. PERMANOVA analysis with 99,999 permutation were performed in each
742 case A and B) ITA (blue and light blue), SPO (green and light green), TBO (red and light red),
743 TRO (grey and light grey), and YNG (orange and light orange). C and D) ITA (blue and light
744 blue), SPO (green and light green), TBO (red and light red), TRO (grey and light grey).
745
746
747

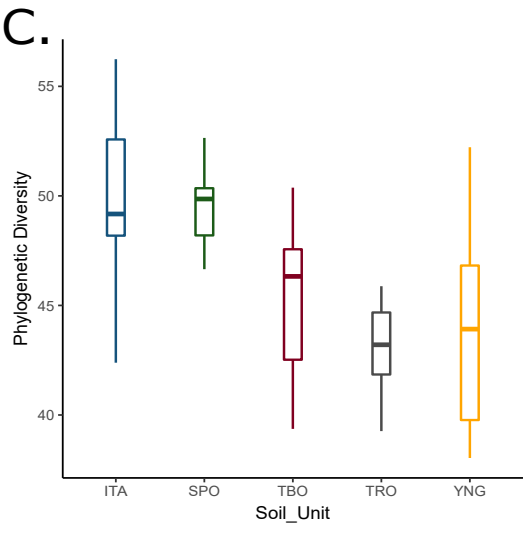
A.

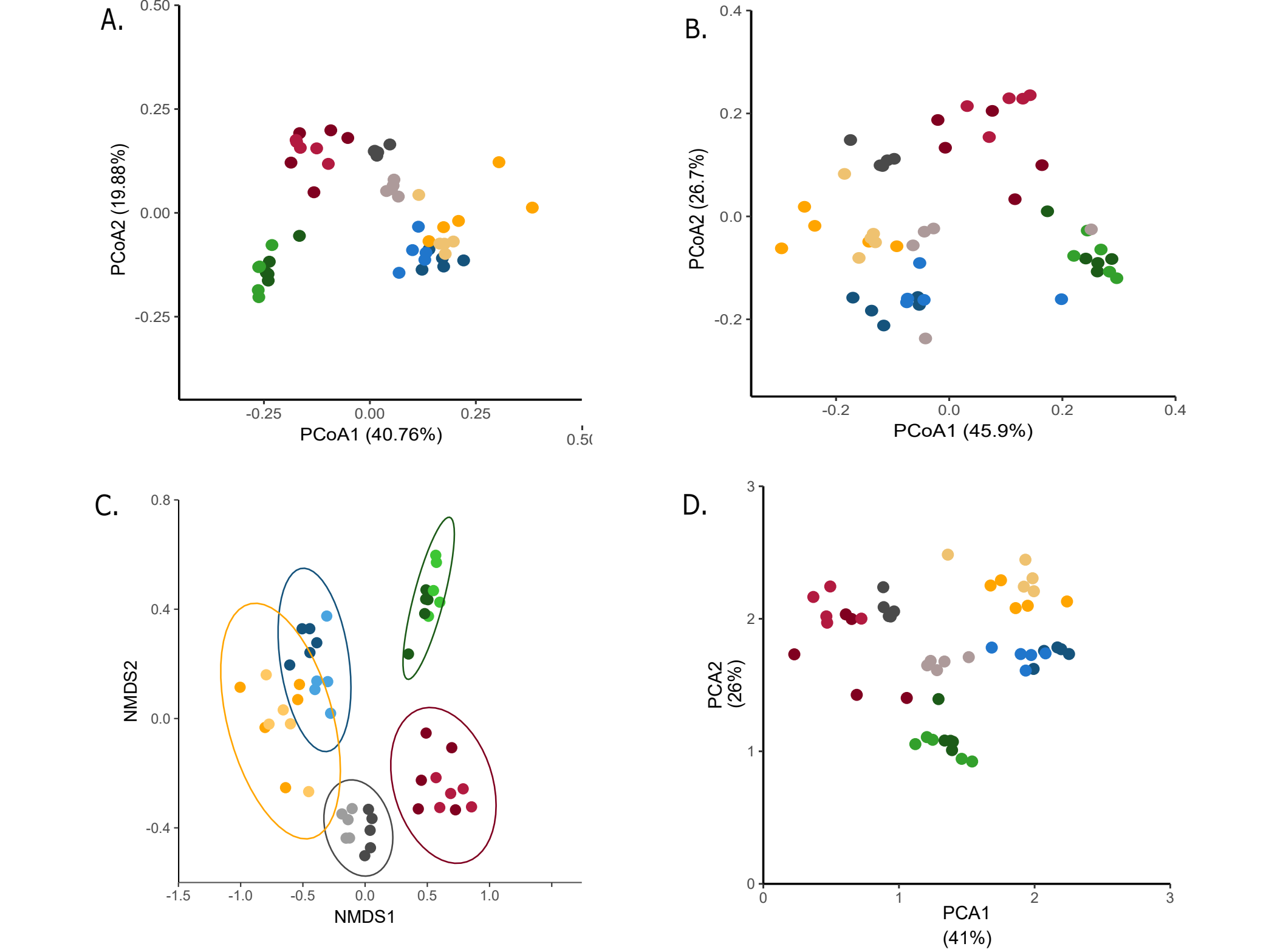

Shannon Entropy (H')
6.8
6.6
6.4
ITA
SPO
TBO
TRO
YNG
Soil Unit

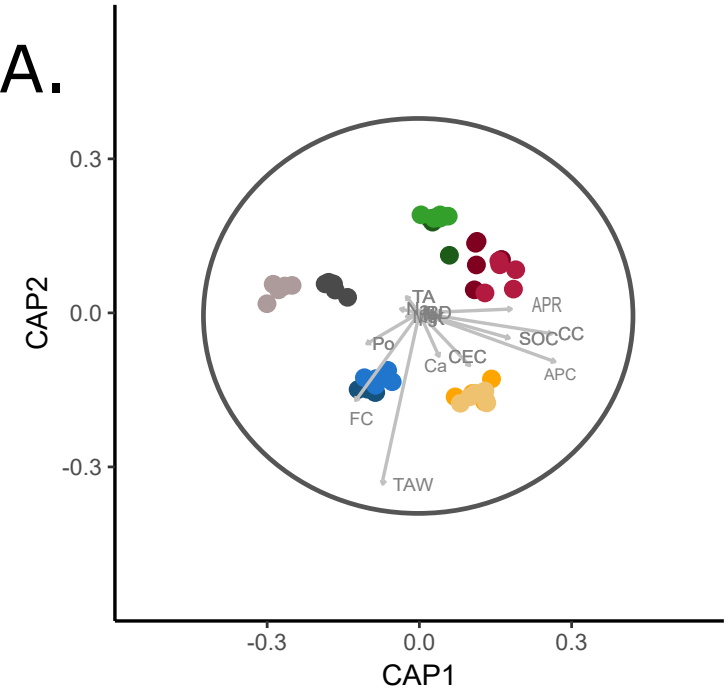
B.

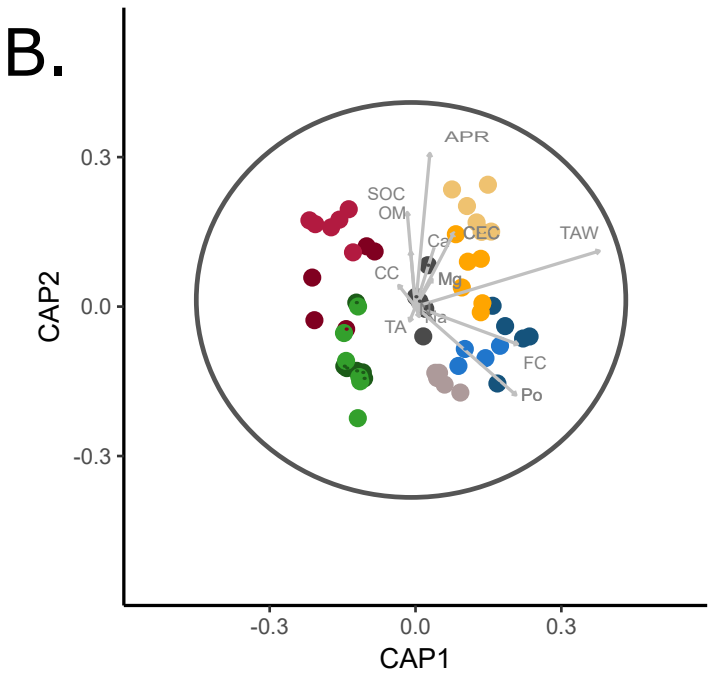

Chao1 Index
50
40
30
20
10
0
ITA
SPO
TBO
TRO
YNG
Soil Unit

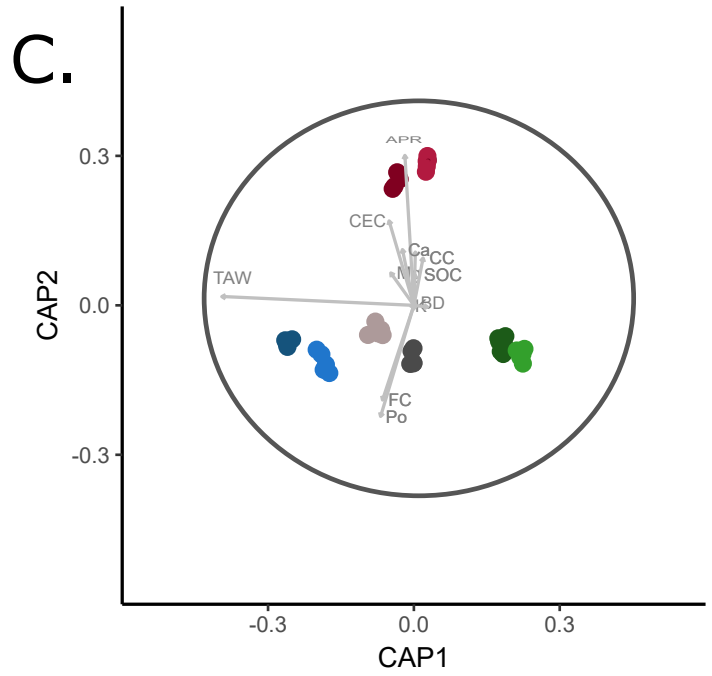
C.

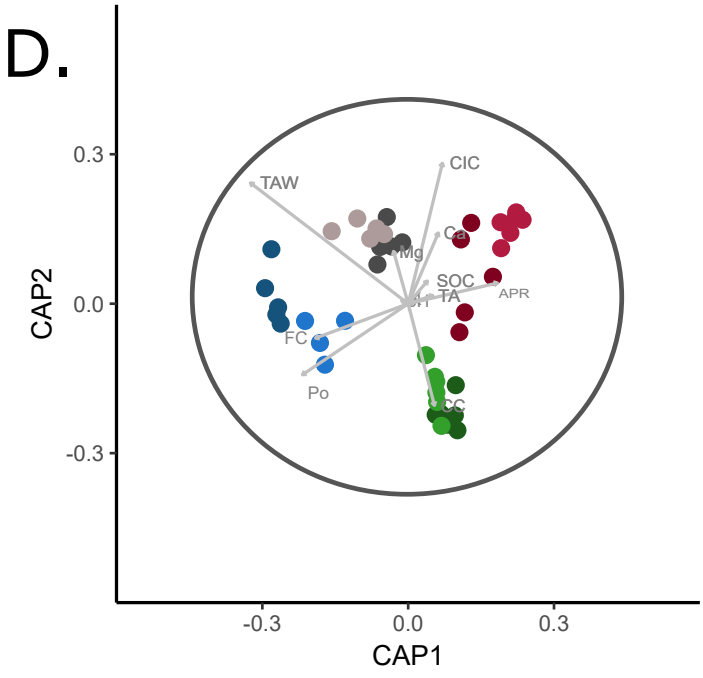

Phylogenetic Diversity
55
50
45
40
ITA
SPO
TBO
TRO
YNG
Soil_Unit


A.
PCoA2 (19.88%)
PCoA1 (40.76%)
B.
PCoA2 (26.7%)
PCoA1 (45.9%)
C.
NMDS2
NMDS1
D.
PCA2
(26%)
PCA1
(41%)


A.
0.3
0.0
-0.3
CAP2
-0.3
0.0
0.3
CAP1
TA
APR
Po
SOC
CC
Ca
CEC
APC
FC
TAW


B.
0.3
0.0
-0.3
CAP2
-0.3
0.0
0.3
CAP1
APR
SOC
OM
CEC
TAW
Ca
CC
Mg
TA
FC
Po


C.
0.3
0.0
-0.3
CAP2
-0.3
0.0
0.3
CAP1
APR
CEC
Ca
CC
SOC
TAW
BD
FC
Po


D.
0.3
0.0
-0.3
CAP2
-0.3
0.0
0.3
CAP1
TAW
CIC
Mg
SOC
TA
APR
FC
Po