Microbial communities in Pampa soils : impact of landuse and climatic conditions

Lupatini, M.

Chapter 4

Shifts in soil bacterial community after eight years of land-use change

Authors

Afnan K. A. Suleiman*, Manoeli Lupatini*, Juliano T. Boldo, Marcos G. Pereira, Luiz F. W. Roesch

Published in: Systematic and Applied Microbiology (2013) 36:137-144

** This authors contributed equally to this study*

Abstract

The interaction between plants, soil and microorganisms is considered to be the major driver of ecosystem functions and any modification of plant cover and/or soil properties might affect the microbial structure, which, in turn, will influence ecological processes. Assuming that land-use changes are the major drivers of soil bacterial diversity and structure, it can be postulated that changes in plant cover causes significant shifts in soil bacterial community composition. To address this issue, this study used 16S rRNA pyrosequencing to detect differences in diversity, composition and/or relative abundance of bacterial taxa from an area covered by pristine forest, as well as eight-year-old grassland surrounded by the same forest. It was shown that a total of 69% of the operational taxonomic units (OTUs) were shared between sites. Overall, forest and grassland samples presented the same diversity and the clustering analysis did not show the occurrence of very distinctive bacterial communities between sites. However, 11 OTUs were detected in statistically significantly higher abundance in the forest samples but in lower abundance in the grassland samples, whereas 12 OTUs occurred in statistically significantly higher abundance in the grassland samples but in lower abundance in the forest samples. The results illustrated that the history of land use might influence present-day community structure.

Keywords

Pyrosequencing; Bacterial diversity; Grassland; Afforestation

4.1 Introduction

Soils are considered to be the most diverse microbial habitats on Earth. However, little is known about how environmental changes affect the microbiota and its functions (Fierer et al. 2007; Liebich et al. 2006). Land-use changes and agricultural management are major causes of biodiversity loss with negative consequences for the environment (Balvanera et al. 2006; Doran & Zeiss 2000; Navarrete et al. 2010). Changes in composition or species diversity of above-ground communities can affect the composition and function of below-ground communities and vice versa (van der Heijden et al. 2008). Particularly, changes in the above-ground vegetation affect the size, activity and composition of soil microbial communities (da C Jesus et al. 2009; Nüsslein & Tiedje 1999). Nevertheless, agricultural practices do not always deplete soil bacterial diversity, since shifts in microbial diversity and structure caused by different land-use changes may have a positive, negative or neutral impact (Singh et al. 2004). According to Jangid et al. (2011), microbial communities in relatively pristine deciduous forest and long-term mowed grassland soils were very similar, despite major differences in soil properties and vegetation.

Although changes in soil properties due to continuous cultivation appear to be a slow process, any land-use change can possibly cause a disturbance, which in turn might affect soil microbial communities. According to Allison and Martiny (2008), there are three potential impacts caused by disturbance. After disturbance, the microbial composition might be resistant and not change, might be altered and rapidly return to the original composition

(resilient) or might remain altered, which would imply a functional change. It was assumed in this study that recent changes in plant cover would not cause major changes in soil properties and that soil properties are the major drivers of soil bacterial diversity and structure. Therefore, the hypothesis was that the soil bacterial community from a pristine forest would not be different from the soil bacterial community from a cultivated grassland surrounded by the forest in the first years of cultivation.

Within this context, the aim of this work was to investigate bacterial communities from distinct land uses, and address the following question: what is the contribution of plant community composition on bacterial community patterns in the first years after land-use change? An area covered by pristine forest soil was analyzed in conjunction with eight-year-old grassland surrounded by this forest. The area had low human activity, no inputs of fertilizers (except for the manure added by animal activity) and a very low animal influence, which was ideal for testing the effect of plant cover removal on soil bacterial communities. The diversity and composition of bacterial taxa was analyzed by high throughput pyrosequencing of 16S rRNA gene amplified from DNA extracted directly from the soil samples.

4.2 Materials and methods

4.2.1 Site description and soil sampling

The sampling site was located in the Pampa biome, which has both subtropical and temperate climates with four well-characterized seasons. The climate is the most important factor determining the soil and vegetation. The soil in the major part of the region had an extremely sandy texture due to its sedimentary rook origin (Roesch et al. 2009; Tornquist et al. 2009). The dominant vegetation in this biome is grassland. However, there is a presence of forest formations limited mainly to gallery forests along rivers and forest

formations surrounded by natural grassland (Overbeck et al. 2007; Roesch et al. 2009). Not only the natural grassland is under land-use changes, but also the forests are converted for anthropogenic uses such as pasture or crop fields (Costella et al. 2013). The site analyzed here consisted of pristine forest (gallery forest) and eight-year-old grassland surrounded by the same forest (Table 4.1). The grassland is this study resulted from the deforestation of a small area (5,500 m^2) inside the forest for a non-commercial cattle settlement (Fig. S4.1). No fertilizers, except for the manure added by animal activity, were added to the soil and no exotic plants were introduced.

The soil samples were taken by drawing four randomly distributed 1 m^2 plots per land use (Baker et al. 2009) and a composite sample was collected by taking samples in every corner of the square. Equal amounts of sub-samples from cores were pooled and mixed to compose four samples from the native forest and four samples from the grassland. Bulk soil samples were collected by taking 5 cm diameter, 0 - 10 cm depth cores during the spring of 2010 and stored at -18°C until DNA extraction and chemical analyses were performed. For the soil chemical analyses, the four replicates from each land use were combined. The soil pH was determined in water (1:1 soil to water ratio) and the concentrations of Ca, Mg, Al, K, Na, P, total nitrogen, NH_4^+, NO_3^- + NO_2^- and total organic carbon (TOC) were quantified according to (Embrapa 1997). The dissolved organic carbon (DOC) was quantified according to Bartlett & Ross (1988). DNA was isolated from at least 1g of soil using the PowerSoil® DNA Isolation Kit (MO BIO Laboratories Inc., Carlsbad, CA, USA), according to the manufacturer's instructions. After DNA extraction, samples were purified with the DNeasy Tissue kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions, and the total DNA concentration was quantified using the NanoVue spectrophotometer (GE Healthcare, Harriet, USA).

4.2.2 16S rRNA gene amplification and pyrosequencing

The 16S rRNA gene fragments were sequenced using 454 GS FLX Titanium (Lib-L) chemistry for unidirectional sequencing of the amplicon libraries. Barcoded primers allow for combining amplicons of multiple samples into one amplicon library and, furthermore, enable the computational separation of the samples after the sequencing run. To do this, 8-base barcodes were added to the 5′-end of the reverse primers using the self-correcting barcode method of Hamady et al. (2008). The primers were attached to the GS FLX Titanium Adaptor A-Key (5′-CCATCTCATCCCTGCGTGTCTCCGACTCAG-3′) and Adaptor B-Key (5′-CCTATCCCCTGTGTGCCTTGGCAGTCTCAG-3′) sequences, modified for use with GS FLX Titanium emPCR Kits (Lib-L) and a 2-base linker sequence that was inserted between the 454 adapter and the 16S rRNA primers in order to reduce any effect the composite primer might have had on PCR efficiency. A total of eight independent PCR reactions were performed for each composite soil sample with the universal primers 27F and 338R for the amplification of the V1-V2 region of the 16S rRNA gene. PCR was performed with the GoTaq PCR core system (Promega, Madison, WI, USA). The mixtures contained 5 µL of 10x PCR buffer, 200 mM dNTPs, 100 mM of each primer, 2.5 U of *Taq* polymerase and approximately 100 ng of DNA template in a final volume of 50 µL. The PCR conditions were 94°C for 2 min, 30 cycles of 94°C for 45 s, 55°C for 45 s, and 72°C for a 1 min extension, followed by 72°C for 6 min. The PCR products were purified and combined in equimolar ratios with the quantitative DNA binding method (SequalPrep Kit, Invitrogen, Carlsbad, CA, USA) in order to create a DNA pool that was further used for pyrosequencing from the A-Key adaptor. All raw sequences were submitted to the NCBI Sequence Read Archive (SRA) under the accession number SRA013204.1.

4.2.3 Processing of pyrosequencing data and statistical analyses

The raw sequences obtained were processed using the Quantitative Insights Into Microbial Ecology (QIIME) (Caporaso et al. 2010) with the default parameters. Briefly, bacterial sequences were first quality trimmed by removing short sequences (<200 bp), sequences that presented low average quality scores (<25), sequences that did not present a perfect match to the sequence barcode and primer, sequences that presented more than two undetermined bases (Hamady et al. 2008). Additionally, to identify potential chimeric sequences, the dataset was subject to the ChimeraSlayer implemented in mothur (Schloss et al. 2009). After removing low-quality sequences, the multiplexed reads were assigned to the corresponding soil samples based on their barcodes. Bacterial sequences were grouped into operational taxonomic units (OTUs) using a 97% identity threshold and the most abundant sequence from each OTU was selected as a representative sequence for that OTU. Afterwards, the sequences were taxonomically classified using the RDP naïve Bayesian rRNA Classifier (Wang et al. 2007), which assigns complete taxonomic information from domain to species to each sequence in the database with 80% taxonomy confidence and an e-value of 0.001. Good's coverage was calculated for each taxonomic level (Phylum, Class, Order, Family and Genus) (Good 1953). The representative set of sequences was also used for aligning the sequences against a reference database and to build a phylogenetic tree necessary for downstream measurements. These taxonomic assignments were used to build an OTU table, which was a matrix of OTU abundance for each sample with meaningful taxonomic identifiers for each OTU. The total number of sequences obtained from the native forest and the grassland are shown in Table 4.2.

4.2.4 Measurement of differences between the bacterial communities

To explore the similarities and differences between the two sites tested, jackknifed principal coordinates analysis (PCoA) and a hierarchical clustering analysis were performed in order to find clusters of similar groups of samples. PCoA is an ordination method based on multivariate statistical analysis that maps the samples in different dimensions and reflects the similarity of the biological communities. A matrix using the UniFrac metrics (weighted and unweighted) for each pair of sites was calculated. The distances were turned into points in space with the number of dimensions one less than the number of samples. The first three principal dimensions, which usually contain most of the variation found in the samples, were used to plot a three-dimensional graph that illustrated the distribution of soils according to their similarity. To test whether the results were robust for sample size, a sequence-jackknifing technique was used in which the PCoA clusters were regenerated using a subset of 600 sequences randomly selected from each soil for 100 replicate trials, and this was used to create a graph made up of the mean values obtained. Ellipses were drawn around the mean values representing the interquartile ranges (measurement of statistical dispersion obtained by sequencing jackknifing). If the ellipses are small, the same result would likely be achieved with a different set of sequences from the same site, but if the ellipses are large a different result might be expected. Furthermore, a hierarchical cluster tree was constructed on the basis of the distance matrix calculated by the unweighted UniFrac algorithm. To assess the uncertainty in hierarchical cluster analysis 1000 bootstrap re-samplings were computed. The jackknifed PCoA and the hierarchical cluster analysis were performed using QIIME (Caporaso et al. 2010).

To compare the diversity between bacterial communities from the soil

samples, the diversity of each sample was estimated using the Shannon-Weaver index (Shannon 2001) and Faith's index of phylogenetic diversity (Faith 1992). For these measurements, the diversity metrics were calculated for a randomly selected subset of 12,393 sequences per soil, as alpha diversity indexes are correlated with the number of sequences collected (Lemos et al. 2011). To find which OTUs were abundantly different between the two sites a chi-square test (based on 50,000 Monte Carlo iterations) was calculated in order to obtain a *p*-value for the null hypothesis that there was no difference between all possible pairwise combinations of soil samples from the native forest and the grassland. The *p*-values (≤0.01) were ordered and processed in order to find a false discovery rate (FDR) less than or equal to 1%. The test was performed using the OTU table summarized at the genus level with the sub-sampled number of sequences (12,393 sequences) for each sample in PANGEA (Giongo et al. 2010).

4.3 Results

4.3.1 Vegetation and soil chemical analysis

The most common native tree species that were found in the forest belonged to the families of *Boraginaceae, Euphorbiaceae, Fabaceae, Lauraceae, Malvaceae, Meliaceae, Myrtaceae,* and *Rutaceae*. The most dominant grass species found in the grassland belonged to the *Poaceae* family. The number of plant families indicated greater plant diversity in the forest and a dominance of a single family in the grassland. The location, altitude and soil chemical analyses are presented in Table 4.1. The pH and sodium content did not differ between soils from the native forest and those from the grassland. All the other variables measured presented higher contents in the native forest than in the grassland, except for K that was greater in the grassland than in the forest. Some nutrients, such as P and NO_3^- + NO_2^- were found to be at least 1.5-fold

higher in the native forest than in the grassland. The total organic carbon was 2.2-fold higher in the native forest than in the grassland.

4.3.2 Assessment of taxon distribution and bacterial diversity

After filtering the reads by base quality and removing reads smaller than 200 bases, a total of 170,046 sequences were obtained from the eight soil samples collected in the native forest and the grassland from the Pampa biome. From all samples, 127,238 (74.83%) were classified below the domain level. The number of high quality sequences per sample varied from 12,393 to 37,225 and the average number of sequences per sample was 21,256 (Table 4.2). The classified sequences were affiliated to 20 bacterial phyla but only eight phyla were found at a relative abundance of greater than 1% (Fig. 4.1 A), and ten phyla were found at a relative abundance smaller than 1% (Fig. 4.1 B). The major phylogenetic groups did not differ between forest and grassland soil samples.

Table 4.1 Location, altitude and soil chemical analyses of native forest and grassland soils from Brazilian Pampa biome

	Native forest	Grassland
Coordinates	30° 24′ 09.3″ S 53° 52′ 59.1″ W	30° 24′ 08.9″ S 50° 53′ 05.9″ W
Altitude (m)	616	616
pH	5.8	5.6
Ca + Mg (cmolc kg^{-1})	39.0	23.2
Al (cmolc kg^{-1})	0.50	0.13
Na (cmolc kg^{-1})	0.014	0.014
K (cmolc kg^{-1})	0.6	1.0
P (cmolc kg^{-1})	39	12
Total N (%)	0.76	0.40
$NH4^+$ (mg kg^{-1})	180	120
$NO3^-$ + $NO2^-$ (mg kg^{-1})	102	30.8
Total organic carbon (%)	7.3	3.3
Dissolved organic carbon (g kg^{-1})	0.24	0.21

To identify shifts in bacterial diversity between the forest and the grassland, two diversity indices, the Shannon-Weaver index and the phylogenetic diversity (PD) index, were calculated. For the calculations, a random subset of sequences (12,393 per sample) was sampled in order to correct for the differences between samples related to the sampling coverage (Table 4.2). Overall, forest samples and grassland samples presented the same diversity. The Shannon index ranged from 10.01 to 10.58 for samples from the forest and from 10.18 to 10.53 for samples from the grassland.

The average Shannon diversity index revealed no significant

differences between site according to the Tukey's test at a 5% probability error. The PD index ranged from 145.37 to 172.23 for samples from the forest, and from 141.84 to 165.02 for samples from the grassland. Although the average PD index was larger in the samples from the forest, the Tukey's range test at a 5% probability error revealed no significant differences between sites.

Table 4.2 Total number of sequences, Good's coverage and diversity index

	Forest				Grassland			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Total nº. of sequence	16,337	16,994	37,225	17,240	12,393	14,328	25,797	29,732
	Sequence coverage (%)							
Phylum	99.98	99.99	99.99	99.98	100	100	100	100
Order	99.96	99.96	99.99	99.97	99.97	99.99	99.99	99.98
Class	99.93	99.95	99.98	99.95	99.97	99.97	99.98	99.97
Family	99.85	99.91	99.96	99.91	99.90	99.87	99.96	99.95
Genus	99.67	99.76	99.85	99.77	99.69	99.69	99.86	99.87
3% dissimilarity cutoff	85.36	83.05	90.95	85.55	84.77	82.39	89.75	90.50
	Diversity index							
[a] **Phylogenetic diversity**	154.5	172.3	145.4	148.0	141.8	165.0	143.8	144.2
Shannon	10.3	10.6	10.0	10.6	10.2	10.5	10.4	10.2

[a] All samples were sub-sampled to 12,393 sequences prior to diversity index calculations. The average phylogenetic diversity for the forest samples was 155.0 and for the grassland samples it was 148.7. The average Shannon diversity index for the forest samples was 10.4 and for the grassland samples 10.3. The means did not differ statistically between the forest samples and the grassland samples by the Tukey test at a 5% probability error.

4.3.3 Similarity between communities based on membership and structure

The weighted and unweighted PCoA analyzes (Fig. 4.2A and B) did not show the occurrence of very distinct groups of soil bacterial communities. In addition, the analysis of microbial communities using hierarchical cluster analysis showed that the bacterial communities from the same site (forest or grassland) were more similar to each other than bacterial communities in different sites, as observed by the two highly supported clusters made up of samples from the forest soil and the grassland soil (Fig. 4.3).

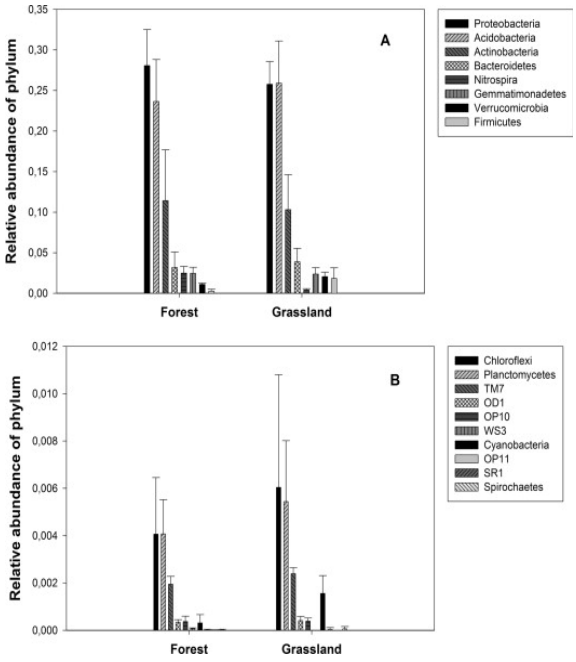

Fig. 4.1 Relative abundance of phyla for each soil library. (A) represents the relative abundance greater than 1%; (B) represents the relative abundance smaller than 1%. Bars represent the standard error ($n = 4$).

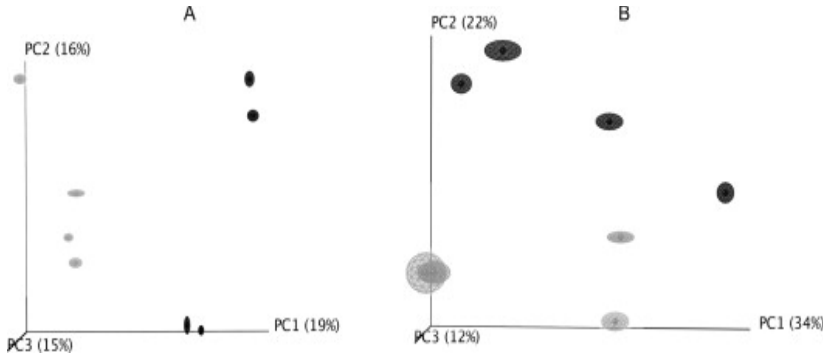

Fig. 4.2 Jackknifed PCoA plots with (A) unweighted UniFrac distance metric, which accounts for presence/absence of taxa and (B) weighted UniFrac distance metric, which accounts for changes in the relative abundance of taxa. The clusters were generated using a subset of 600 sequences from each site for 100 replicate trials. The positions of the points are the average for the jackknife replicates and ellipses were drawn around the mean values to represent the IQRs. Dark grey - grassland and light grey - natural forest.

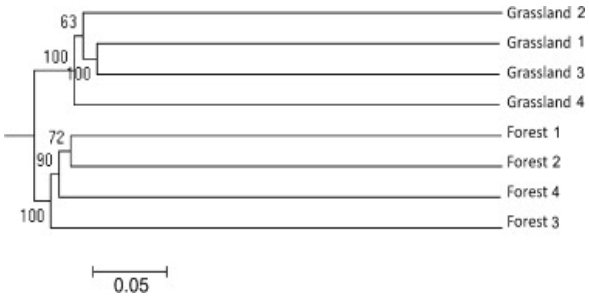

Fig. 4.3 Hierarchical cluster constructed on the basis of the distance matrix calculated by the unweighted UniFrac algorithm. Numbers at branch points indicate the percentage of 1000 bootstrap re-samplings.

4.3.4 Co-occurrence of OTUs among soil samples

An important component of this analysis was to identify those bacteria that were responsible for the differences observed between forest and grassland soil samples. To determine the OTUs that were statistically different between sites, an exact chi-square test was performed. On the basis of the test, only eleven OTUs were found to be in higher abundance in the forest samples but in lower abundance in the grassland samples (Table 4.3). The unclassified Bacteria also presented different abundances between the sites. A total of seven unclassified OTUs presented higher abundances in the native forest, while two unclassified OTUs showed higher abundance in the grassland. Among those sequences that could not be classified to known taxa, the OTUs 5084, 6116 and 4424 presented the greatest difference between forest and grassland.

Table 4.3 List of the closest bacterial relatives of operational taxonomic units (OTUs) whose abundances differed statistically ($p \leq 0.01$; FDR ≤ 0.01) between forest and grassland soils

	% of all forest sequences	% of all grassland sequences	Fold difference
[a]Classifiable OTUs with greater abundance in the native forest			
Acidobacteria;Acidobacteria;Gp6;Gp6	15.26	11.26	1.4
Acidobacteria;Acidobacteria;Gp22;Gp22	0.61	0.28	2.2
Actinobacteria;Actinobacteria;Acidimicrobiales	0.59	0.21	2.8
Actinobacteria;Actinobacteria;Actinomycetales; Micromonosporaceae	5.11	2.88	1.8

	% of all forest sequences	% of all grassland sequences	Fold difference
Actinobacteria;Actinobacteria;Actinomycetales; Mycobacteriaceae;Mycobacterium	0.71	0.41	1.8
Actinobacteria;Actinobacteria;Actinomycetales; Propionibacteriaceae;Microlunatus	1.24	0.82	1.5
Nitrospira;Nitrospira;Nitrospirales; Nitrospiraceae;Nitrospira	3.91	0.57	6.9
Proteobacteria	5.02	3.57	1.4
Proteobacteria;Alphaproteobacteria	5.44	3.42	1.6
Proteobacteria;Alphaproteobacteria;Rhizobiales	14.57	10.70	1.4
Proteobacteria;Deltaproteobacteria	4.36	3.37	1.3
Classifiable OTUs with greater abundance in the grassland			
Acidobacteria;Acidobacteria;Gp1;Gp1	1.73	5.83	3.4
Acidobacteria;Acidobacteria;Gp4;Gp4	8.48	10.36	1.2
Actinobacteria;Actinobacteria;Actinomycetales	4.56	5.91	1.3
Bacteroidetes;Sphingobacteria; Sphingobacteriales;Chitinophagaceae; Terrimonas	1.14	1.87	1.6
Chloroflexi;Anaerolineae;Anaerolineales; Anaerolineaceae	0.51	0.78	1.5
Firmicutes;Bacilli;Bacillales;Bacillaceae; Bacillus	0.38	2.74	7.2
Planctomycetes;Planctomycetacia;	0.26	0.42	1.6

	% of all forest sequences	% of all grassland sequences	Fold difference
Planctomycetales;Planctomycetaceae			
Proteobacteria;Alphaproteobacteria; Rhodospirillales	0.15	0.45	3.0
Proteobacteria;Alphaproteobacteria; Sphingomonadales;Sphingomonadaceae	0.41	0.69	1.7
Proteobacteria;Betaproteobacteria	2.63	3.76	1.4
Verrucomicrobia;Spartobacteria; Spartobacteria;genera;incertaesedis	0.58	1.04	1.8
Verrucomicrobia;Subdivision3; Subdivision3 genera incertae sedis	0.62	1.36	2.2
Unclassified			
OTUs with greater abundance in the native forest			
Bacteria 1137	0.081	0.006	1.3
Bacteria 3541	0.020	0.002	1.0
Bacteria 4424	2.058	0.056	3.6
Bacteria 5084	0.141	0.002	7.0
Bacteria 5735	0.061	0.004	1.5
Bacteria 5785	0.121	0.010	1.2
Bacteria 6116	0.101	0.002	5.0
OTUs with greater abundance in the grassland			
Bacteria 2294	0.726	0.107	1.5

	% of all forest sequences	% of all grassland sequences	Fold difference
Bacteria 2768	0.020	0.008	4.0

[a] Each OTU was classified at the highest taxonomic level with 80% taxonomy confidence and an *e*-value of 0.001. The unclassified Bacteria correspond to an OTU that did not match any of the sequences in the database according to the criteria mentioned above.

The analysis of the OTUs that were partitioned between samples showed that most of the taxa were shared between forest and grassland (69%). However, 16.4% of the taxa were found only in the forest soil and 14.6% only in the grassland. The exclusive OTUs from each site are shown in Table S4.1. The abundance of a genus was analyzed statistically to provide support for the analysis of shared OTUs using a t-test on QIIME. The OTUs found exclusively in the forest samples belonged to the genera *Sphingobium*, *Methylotenera* and *Pedobacter*, and to the phylum *WS3*. The OTUs found exclusively in the grassland samples belonged to the genera *Dechloromonas*, *Zoogloea* and *Geobacter*.

4.4 Discussion

As microorganisms play key roles in nutrient cycling and other important functions in soils, the shifts in microbial communities caused by land-use changes might directly affect the functioning of ecosystems, such as biogeochemical cycles (Berthrong et al. 2009). In this study, differences in diversity, composition and/or relative abundance of bacterial taxa were tested from bulk soil samples collected in two sites: pristine forest and eight-years-old

grassland resulted by the deforestation of a small area inside the forest. The area chosen for sampling was ideal for testing the effect of removing plant cover on soil bacterial communities, since it presented low human activity, no inputs of fertilizers and a very low animal influence. As the samples were taken in one single period of time (during the spring) it is important to mention that the results obtained represent a "*snapshot*" of the microbial community status, and temporal variations in plant growth and seasonal fluctuations are not considered.

We found that the replacement of forest for grassland reduce most of the soil chemical properties measured. In agreement with our results, many researchers reported that the conversion of natural forest to other forms of land use (*e.g.*, crop fields or pasture) lead to a reduction in soil chemical properties such as organic content, N and Ca, suggesting that continuous use of soil for anthropogenic purposes might be responsible for deterioration in soil quality (Braimoh & Vlek 2004). Due the removal of forest, soil attributes change as a consequence of the loss of input from forest litter, increase rates of organic matter decomposition and nutrient depletion caused by prolonged nutrient mining without sufficient replenishment of nutrients and of nutrient losses by soil erosion (Tan et al. 2005; Bringhurst & Jordan 2015). Yet, the magnitude of these changes vary with land cover, land management intensity and time (Houghton et al., 1999; Kizilkaya & Dengiz 2010). Geissen et al. (2009) studied the effect of land-use changes after 15 years and did not detect chemical soil degradation but did detect severe compaction of soils under permanent pasture. In certain ecosystems, the effect of agricultural practices on soil properties (*e.g.*, forestland replaced by cropland and pasture) was clearly detected only 50 years after the land-use change (Kizilkaya & Dengiz 2010).

Our approach was based on pyrosequencing of the 16S rRNA genes amplified from microbial DNA extracted directly from four soil samples from each site. This approach is considered to present high levels of robustness and

resolution (Liu et al. 2007; Lozupone & Knight 2007). However, it should be mentioned that an incomplete dataset was examined and other factors, not assessed in this study, such as biases at the steps of DNA extraction, PCR amplification, primer choice and sequencing, might present some degree of interference in the results obtained. Nevertheless, as the sequencing errors were removed, these biases were unlikely to have missed many taxa that could have resulted in the main findings being incorrect. The selection of primers is still under debate among researchers, since no primer is truly "universal" and the choice of a reference database and taxonomy can have a dramatic impact on the resulting classification accuracy (Soergel et al. 2012). The most widely used PCR primer sets span hypervariable regions V1-V3 but *in silico* tests have revealed that primers designed for amplification of this region underestimate the richness because they neglect candidate divisions (Winsley et al. 2012). On the other hand, *in silico* predictions may not reflect the real performance of the primers. The V3-V4 region, for example, has shown high-simulated accuracy and good classification consistency but the set of primers designed to amplify this region has been proven to produce biases caused by amplification of artifacts arising from the combination of these two specific V3-forward and V4-reverse primers (Claesson et al. 2010). Although we are aware that our primer choice excluded some phyla during amplification, mainly *Verrucomicrobia* (Bergmann et al. 2011), and that the barcoded primers used might be a source of bias (Berry et al. 2011), we opted to amplify the DNA with the 27F and 338R primers and performed the sequencing through the reverse end. This was because this set of primers generate low rates of PCR artifacts and the sequences produced provided relatively good cluster recovery even for short (≤250 bases) pyrosequencing reads (Liu et al. 2007). Thus, PCR primers rarely amplify all bacterial members of a community and any PCR-based approach is likely to miss some bacterial groups or at least underestimate the abundance of some bacterial taxa. Although microbial surveys are always limited by these

practical problems, it is possible to obtain robust comparisons across samples when the data analysis is conducted in a consistent manner (Bent & Forney 2008; Hamady et al. 2008).

To detect relevant bacterial patterns within our samples, the datasets of 16 rRNA sequences were analyzed using phylogenetic- and taxon-based approaches. The methods based on phylogeny are useful for exploring similarities and differences based on a phylogenetic tree, while OTU-based approaches need a rigid OTU definition based on a cutoff distance. Since there are no accepted dissimilarity cutoffs for the different microbial taxonomic levels, the 3% dissimilarity clustering threshold proposed by Kunin et al. (2010) was used. According to the authors, diversity estimates are grossly overestimated when clustering thresholds are higher than 97% identity. Therefore, phylogenetic parental sequences can be grouped differently than those based on OTU identification. In this regard, two different metrics were applied to calculate bacterial diversity among samples: the Shannon diversity index (H′) and the phylogenetic diversity index (PD). Shannon index is an OTU-based analysis and measures the average degree of uncertainty in predicting to what species an individual chosen at random from a collection of S species and N individuals will belong. The value increases as the number of species increases and as the distribution of individuals among the species becomes even (Ludwig 1988). The phylogenetic diversity is defined and calculated as the sum of the branch-lengths of the minimal subtree connecting the taxa in the subset (Faith 1992). This evaluation is based on a single phylogenetic tree and is sensitive to the quality of the branch length and topology. Another problem associated with measurements of microbial diversity using diversity indexes is related to uneven sequence sampling. Diversity index values increase with sample size and make normalization of the number of sequences in all samples crucial. Within this work the calculations of both diversity indexes mentioned above were performed with sub-samples of 12,393

sequences. This reduced the bias associated with the sample size and allowed for a better comparison between the samples.

Generally, in agricultural systems with low vegetative diversity and high xenobiotic inputs, overall species diversity may be reduced to a bottleneck, from which species diversification is possible, albeit from a limited number of phyla (Roesch et al. 2009). Agricultural practices such as tillage, application of pesticides and nutrients, machinery traffic used for the soil and crops management, modify the physical and chemical properties of soil and, consequently, alter soil microbial diversity and ecological functions (Bissett et al. 2011; Lozupone & Knight 2007). In the same line, Hossain and Sugiyama (2011) suggested that soils exposed to frequent human disturbances might show modification of microbial structure or reduction of the microbial diversity. Contrary, we found that the removal of forest did not reduce the bacterial diversity. In general, our findings are in agreement with other studies, in which deforestation or low plant diversity due the result of forest conversion for pasture did not necessarily lead to a reduction of the bacterial diversity (da C Jesus et al. 2009; Ding et al. 2013). Deforestation might not always be reflected in loss of microbial community diversity likely may be explained due to the large presence of inactive bacterial cells in soils detected by DNA-based methods (Lennon & Jones 2011). Recently, Fierer and Lennon (2011) revised the concepts on generation and maintenance of diversity in microbial communities. Dormancy "refers to an organism's ability to enter a reversible state of low metabolic activity when faced with unfavorable environmental conditions" (Lennon & Jones 2011) and it may work as a microbial seed bank that helps to maintain the high levels of microbial biodiversity observed in nearly all ecosystems (Jones & Lennon 2010). As our approach was not sensitive to microbial activity, the metabolic status of our samples was unclear. However, following the concept of a seed bank proposed by Lennon and Jones (2011), the similarity in diversity levels of our soil samples may reflect a

reservoir of biodiversity that could potentially be resuscitated in the future under different environmental conditions.

The interaction between plants, soil and microorganisms is the driver of ecosystem functions and a disturbance might affect the microbial structure, which, in turn, will influence the soil processes and ecological functions (Singh et al. 2004). In our experiment, the disturbance was constant but relatively recent (eight years), and before removing the plant cover the same soil microbial diversity and structure was expected to be found. According to Miki et al. (2010), a change in the composition of a plant community leads to a change in the litter quality that in turn alters the nutrient cycling process and resultant soil conditions. Due to differences in vegetation composition, a clear discrimination between the microbial diversity and community structure from the forest and the grassland soils would be expected (Mitchell et al. 2010). However, a large overlap (69% of shared OTUs) was found between both microbial communities and no clear discrimination between them. Analysis of shared OTUs would be reasonable only when the sequencing coverage was enough to detect most of the OTUs present (90% or more), since the power for detecting overlapping species from multiple environments is strongly related to the sequencing intensity (Lemos et al. 2011). To circumvent the problem associated with detection of overlapping taxa, we first calculated how well each sample was representative for the bacterial community. The data summarized at the genus level provided reasonable coverage (greater than 99%), therefore the analysis of bacterial genera that were either unique or shared by specific soil samples was sensitive enough to detect the changes in the number of sequences, as well as the presence/absence of taxonomic units.

According to Martiny et al. (2006), the present-day community structure may have been driven by historical events (*e.g.*, prevalence of any type of vegetation, weather conditions) that might influence this particular community structure. Although the soil bacterial community did not suffer

great alteration after removing the natural forest, we were able to detect shifts related to specific bacterial groups. A total of 11 OTUs were found in statistically significant higher abundance in the forest samples but in lower abundance in the grassland samples. The *Nitrospira* genus, for example, was found in greater numbers in the forest than in the grassland. On the other hand, 12 bacterial taxa were found in higher abundance in the grassland samples but in lower abundance in the forest samples. This observation might be indicative that soil bacterial communities under the influence of environmental change will gradually be replaced by another community composed of different species that survive better with the new conditions. The bacterial community structure might change with time and, without any significant changes in soil properties, the plant cover will be the major driver of bacterial diversity and structure, as proposed by Mitchell et al. (2010). In addition, the new bacterial community may be functionally equivalent to the original one even if it has a different structure (Allison & Martiny 2008). The results obtained by other studies indicate that a greater degree of disturbance would be necessary to cause major shifts in microbial diversity and structure for the soil tested in this work. The results suggest the prevalence of a resilient microbial community less influenced by plant cover in which the history of land use might influence present-day community structure.

4.5 Supplementary material

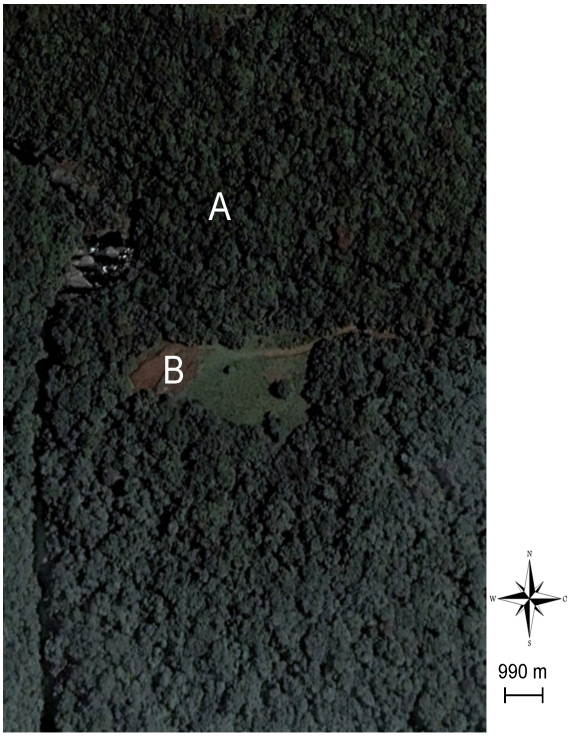

Fig. S4.1 (A) Pristine forest and (B) eight-years-old grassland resulted by the deforestation of a small area (5,500 m^2) inside the forest.

Table S4.1 List of the closest bacterial relative of Operational Taxonomic Unities (OTUs) uniquely found in native forest or grassland.

Native Forest	**Grassland**
Actinobacteria;Actinobacteria; Actinomycetales;Segniliparaceae; Segniliparus	*Firmicutes;Clostridia*
Bacteroidetes;Bacteroidia;Bacteroidales	*Proteobacteria;Gammaproteobacteria; Xanthomonadales*
Bacteroidetes;Bacteroidia;Bacteroidales; Porphyromonadaceae	*Firmicutes;Bacilli;Bacillales; Bacillaceae;Lysinibacillus*
Proteobacteria;Gammaproteobacteria; Xanthomonadales;Xanthomonadaceae; Thermomonas	*Proteobacteria;Deltaproteobacteria; Bdellovibrionales;Bacteriovoracaceae; Bacteriovorax*
Proteobacteria;Betaproteobacteria; Burkholderiales;Comamonadaceae; Hylemonella	*Actinobacteria;Actinobacteria; Actinomycetales; Promicromonosporaceae;Xylanimonas*
SR1;SR1_genera_incertae_sedis	*Proteobacteria;Betaproteobacteria; Burkholderiales;Alcaligenaceae; Azohydromonas*
Actinobacteria;Actinobacteria; Actinomycetales;Nocardiaceae;Nocardia	*Chloroflexi;Anaerolineae;Anaerolineale; Anaerolineaceae;Anaerolinea*
Planctomycetes;Planctomycetacia; Planctomycetales;Planctomycetaceae; Pirellula	*Proteobacteria;Alphaproteobacteria; Rhodospirillales;Rhodospirillaceae; Azospirillum*
Proteobacteria;Gammaproteobacteria; Xanthomonadales;Xanthomonadaceae; Stenotrophomonas	*Actinobacteria;Actinobacteria; Actinomycetales;Streptosporangiaceae; Sphaerisporangium*
Proteobacteria;Gammaproteobacteria; Xanthomonadales;Sinobacteraceae;Nevskia	*Proteobacteria;Alphaproteobacteria; Rhizobiales;Methylobacteriaceae; Methylobacterium*
Actinobacteria;Actinobacteria; Actinomycetales;Bogoriellaceae;Georgenia	*Proteobacteria;Alphaproteobacteria; Rhizobiales;Bradyrhizobiaceae; Balneimonas*
Firmicutes;Clostridia;Clostridiales; Ruminococcaceae	*Proteobacteria;Gammaproteobacteria; Enterobacteriales;Enterobacteriaceae; Pantoea*
Proteobacteria;Gammaproteobacteria; Pseudomonadales;Pseudomonadaceae; Cellvibrio	*Proteobacteria;Alphaproteobacteria; Rhodospirillales;Rhodospirillaceae; Magnetospirillum*
Actinobacteria;Actinobacteria; Actinomycetales;Cellulomonadaceae	*Proteobacteria;Deltaproteobacteria; Desulfuromonadales*
Bacteroidetes;Flavobacteria;Flavobacteriales; Cryomorphaceae	*Proteobacteria;Betaproteobacteria; Burkholderiales;Comamonadacea*
Proteobacteria;Deltaproteobacteria;	*Proteobacteria;Deltaproteobacteria;*

Native Forest	**Grassland**
Myxococcales;Phaselicystidaceae; Phaselicystis	*Desulfuromonadales;Geobacteraceae*
Proteobacteria;Betaproteobacteria; Methylophilales;Methylophilaceae; Methylophilus	*Proteobacteria;Deltaproteobacteria; Desulfobacterales;Desulfobulbaceae; Desulfobulbus*
Actinobacteria;Actinobacteria; Actinomycetales;Promicromonosporaceae; Promicromonospora	*Actinobacteria;Actinobacteria; Actinomycetales;Actinosynnemataceae; Lechevalieria*
Actinobacteria;Actinobacteria; Actinomycetales;Nocardiaceae;Williamsia	*Actinobacteria;Actinobacteria; Actinomycetales;Geodermatophilaceae; Modestobacter*
Proteobacteria;Deltaproteobacteria; Myxococcales;Nannocystaceae;Nannocystis	*Actinobacteria;Actinobacteria; Actinomycetales;Propionibacteriaceae; Friedmanniella*
Bacteroidetes;Sphingobacteria; Sphingobacteriales;Cytophagaceae; Dyadobacter	*Proteobacteria;Alphaproteobacteria; Rhizobiales;Methylobacteriaceae; Microvirga*
Firmicutes;Bacilli	*Proteobacteria;Alphaproteobacteria; Sphingomonadales; Sphingomonadaceae;Sphingosinicella*
Proteobacteria;Alphaproteobacteria; Rhizobiales;Hyphomicrobiaceae;Devosia	*Proteobacteria;Betaproteobacteria; Burkholderiales;Burkholderiaceae; Chitinimonas*
Proteobacteria;Gammaproteobacteria; Xanthomonadales;Xanthomonadaceae; Dokdonella	*Proteobacteria;Betaproteobacteria; Burkholderiales;Comamonadaceae; Roseateles*
Proteobacteria;Gammaproteobacteria; Xanthomonadales;Xanthomonadaceae; Dokdonell;Verrucomicrobia;Opitutae	*Proteobacteria;Deltaproteobacteria; Myxococcales;Myxococcaceae*
Actinobacteria;Actinobacteria; Actinomycetales;Micromonosporaceae; Rugosimonospora	*Actinobacteria;Actinobacteria; Actinomycetales;Geodermatophilaceae; Blastococcus*
Actinobacteria;Actinobacteria; Actinomycetales;Nocardiaceae; Smaragdicoccus	*Proteobacteria;Betaproteobacteria; Rhodocyclales;Rhodocyclaceae;Zoogloea*
Actinobacteria;Actinobacteria; Actinomycetales;Cryptosporangiaceae; Cryptosporangium	*Acidobacteria;Acidobacteria_Gp19; Gp19*
Actinobacteria;Actinobacteria; Actinomycetales;Streptomycetaceae; Streptacidiphilus	*Planctomycetes;Planctomycetacia; Planctomycetales;Planctomycetaceae; Gemmata*
Bacteroidetes;Sphingobacteria; Sphingobacteriales;Cytophagaceae;Emticicia	*Proteobacteria;Alphaproteobacteria; Rhodospirillales;Rhodospirillaceae; Defluviicoccus*
Chloroflexi;Anaerolineae;Anaerolineales;	*Firmicutes;Bacilli;Bacillales;*

Native Forest	Grassland
Anaerolineaceae;Longilinea	*Paenibacillaceae;Brevibacillus*
Proteobacteria;Betaproteobacteria; Burkholderiales;Comamonadaceae; Comamonas	*Firmicutes;Clostridia;Clostridiales; Peptococcaceae;Desulfosporosinus*
Bacteria;Proteobacteria;Alphaproteobacteria; Caulobacterales;Caulobacteraceae; Brevundimonas	*Proteobacteria;Alphaproteobacteria; Rhizobiales;Rhizobiaceae;Kaistia*
Proteobacteria;Alphaproteobacteria; Rhodobacterales;Rhodobacteraceae; Paracoccus	*Proteobacteria;Alphaproteobacteria; Rhodobacterales;Rhodobacteraceae; Pannonibacter*
Proteobacteria;Gammaproteobacteria; Xanthomonadales;Xanthomonadaceae; Pseudoxanthomonas	*Proteobacteria;Gammaproteobacteria; Enterobacteriales;Enterobacteriaceae; Serratia*
Proteobacteria;Deltaproteobacteria; Myxococcales;Cystobacteraceae;Hyalangium	*Verrucomicrobia;Verrucomicrobiae; Verrucomicrobiales; Verrucomicrobiaceae;Luteolibacter*
Proteobacteria;Alphaproteobacteria; Rhizobiales;Phyllobacteriaceae;Aminobacter	*Firmicutes;Clostridia;Clostridiales; Clostridiaceae;Clostridium*
Proteobacteria;Alphaproteobacteria; Rhodospirillales;Acetobacteraceae; Roseomonas	*Cyanobacteria;Cyanobacteria*
Chloroflexi;Chloroflexi	*Chloroflexi;Thermomicrobia; Sphaerobacterales;Sphaerobacteraceae; Sphaerobacter*
Bacteroidetes;Sphingobacteria;Sphingobacte riales;Sphingobacteriaceae;Pedobacter	*Actinobacteria;Actinobacteria; Actinomycetales;Nocardioidaceae; Actinopolymorpha*
WS3;WS3_genera_incertae_sedis	*Actinobacteria;Actinobacteria; Actinomycetales;Nocardiaceae; Rhodococcus*
Bacteroidetes;Sphingobacteria;Sphingobacte riales;Cytophagaceae;Cytophaga	*Actinobacteria;Actinobacteria; Actinomycetales;Micromonosporaceae; Planosporangium*
Proteobacteria;Alphaproteobacteria; Sphingomonadales;Sphingomonadaceae; Sphingobium	*Actinobacteria;Actinobacteria; Actinomycetales;Nakamurellaceae; Nakamurella*
Bacteroidetes;Sphingobacteria; Sphingobacteriales;Chitinophagaceae; Filimonas	*Proteobacteria;Betaproteobacteria; Rhodocyclales;Rhodocyclaceae; Dechloromonas*
Acidobacteria;Acidobacteria_Gp9;Gp9	*Proteobacteria;Deltaproteobacteria; Desulfuromonadales;Geobacteraceae; Geobacter*
Proteobacteria;Deltaproteobacteria; Myxococcales;Haliangiaceae;Haliangium	*Verrucomicrobia;Verrucomicrobiae; Verrucomicrobiales; Verrucomicrobiaceae;Haloferula*

Native Forest	**Grassland**
Proteobacteria;Betaproteobacteria; Burkholderiales;Oxalobacteraceae; Janthinobacterium	*Proteobacteria;Alphaproteobacteria; Rhodospirillales;Rhodospirillaceae; Skermanella*
Proteobacteria;Alphaproteobacteria; Rhizobiales;Xanthobacteraceae; Azorhizobium	*Proteobacteria;Alphaproteobacteria; Sphingomonadales;Erythrobacteraceae*
Bacteroidetes;Sphingobacteria; Sphingobacteriales;Flammeovirgaceae	*Firmicutes;Bacilli;Bacillales; Bacillales_incertae_sedis;Solibacillus*
Actinobacteria;Actinobacteria; Actinomycetales;Micromonosporaceae; Catellatospora	
Acidobacteria;Holophagae;Holophagales; Holophagaceae	
Bacteroidetes;Flavobacteria;Flavobacteriales; Flavobacteriaceae;Chryseobacterium	
Proteobacteria;Deltaproteobacteria; Myxococcales;Cystobacteraceae;Stigmatella	
Proteobacteria;Alphaproteobacteria; Rhodospirillales;Rhodospirillaceae; Telmatospirillum	
Proteobacteria;Betaproteobacteria; Methylophilales;Methylophilaceae; Methylotenera	

4.6 References

Allison, S. D. and Martiny, J. B. H. (2008). *Resistance, resilience, and redundancy in microbial communities*, Proceedings of the National Academy of Sciences 105:11512-11519.

Baker, K. L.; Langenheder, S.; Nicol, G. W.; Ricketts, D.; Killham, K.; Campbell, C. D. and Prosser, J. I. (2009). *Environmental and spatial characterisation of bacterial community composition in soil to inform sampling strategies*, Soil Biology and Biochemistry 41:2292-2298.

Balvanera, P.; Pfisterer, A. B.; Buchmann, N.; He, J.-S.; Nakashizuka, T.; Raffaelli, D. and Schmid, B. (2006). *Quantifying the evidence for biodiversity effects on ecosystem functioning and services*, Ecology Letters 9:1146-1156.

Bartlett, R. J. and Ross, D. S. (1988). *Colorimetric determination of oxidizable carbon in acid soil solutions*, Soil Science Society of America Journal 52:1191.

Bent, S. J. and Forney, L. J. (2008). *The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity*, The ISME Journal 2:689-695.

Bergmann, G. T.; Bates, S. T.; Eilers, K. G.; Lauber, C. L.; Caporaso, J. G.; Walters, W. A.; Knight, R. and Fierer, N. (2011). *The under-recognized dominance of Verrucomicrobia in soil bacterial communities*, Soil Biology and Biochemistry 43:1450-1455.

Berry, D.; Ben Mahfoudh, K.; Wagner, M. and Loy, A. (2011). *Barcoded primers used in multiplex amplicon pyrosequencing bias amplification*, Applied and Environmental Microbiology 77:7846-7849.

Berthrong, S. T.; Schadt, C. W.; Pineiro, G. and Jackson, R. B. (2009). *Afforestation alters the composition of functional genes in soil and biogeochemical processes in South American grasslands*, Applied and Environmental Microbiology 75:6240-6248.

Bissett, A.; Richardson, A. E.; Baker, G. and Thrall, P. H. (2011). *Long-term land use effects on soil microbial community structure and function*, Applied Soil Ecology 51:66-78

Bossio, D.; Girvan, M.; Verchot, L.; Bullimore, J.; Borelli, T.; Albrecht, A.; Scow, K.; Ball, A.; Pretty, J. and Osborn, A. (2005). *Soil microbial community response to land use change in an agricultural landscape of Western Kenya*, Microbial Ecology 49:50-62.

Braimoh, A. K. and Vlek, P. L. G. (2004). *The impact of land-cover change on soil properties in northern Ghana*, Land Degradation & Development 15:65-74.

Bringhurst, K. and Jordan, P. (2015). *The impact on nutrient cycles from tropical forest to pasture conversion in Costa Rica*, Sustain. Water Resource Management 1:3-13.

da C Jesus, E.; Marsh, T. L.; Tiedje, J. M. and de S Moreira, F. M. (2009). *Changes in land use alter the structure of bacterial communities in Western Amazon soils*, The ISME Journal 3:1004-1011.

Caporaso, J. G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F. D.; Costello, E. K.; Fierer, N.; Peña, A. G.; Goodrich, J. K.; Gordon, J. I. and et al. (2010). *QIIME allows analysis of high-throughput community sequencing data*, Nature Methods 7:335-336.

Claesson, M. J.; Wang, Q.; O'Sullivan, O.; Greene-Diniz, R.; Cole, J. R.; Ross, R. P. and O'Toole, P. W. (2010). *Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions*, Nucleic Acids Research 38:e200.

Costella, E.; Garcia, B. A.; Costa, L. S. d.; Corneleo, N. d. S.; Schünemann, A. L. and Stefenon, V. M. (2013). *Anthropogenic use of gallery forests in the Brazilian Pampa*,

Actascibiolsci 35:211-217.

da C Jesus, E.; Marsh, T. L.; Tiedje, J. M. and de S Moreira, F. M. (2009). *Changes in land use alter the structure of bacterial communities in Western Amazon soils*, The ISME Journal 3:1004-1011.

Ding, G.-C.; Piceno, Y. M.; Heuer, H.; Weinert, N.; Dohrmann, A. B.; Carrillo, A.; Andersen, G. L.; Castellanos, T.; Tebbe, C. C. and Smalla, K. (2013). *Changes of soil bacterial diversity as a consequence of agricultural land use in a semi-arid ecosystem*, PLoS ONE 8:e59497.

Doran, J. W. and Zeiss, M. R. (2000). *Soil health and sustainability: managing the biotic component of soil quality*, Applied Soil Ecology 15:3-11.

Embrapa (1997). *Manual de métodos de análise de solo*, Rio de Janeiro: Embrapa Solos.

Faith, D. P. (1992). *Systematics and conservation: on predicting the feature diversity of subsets of taxa*, Cladistics 8:361-373.

Fierer, N.; Breitbart, M.; Nulton, J.; Salamon, P.; Lozupone, C.; Jones, R.; Robeson, M.; Edwards, R. A.; Felts, B.; Rayhawk, S. and et al. (2007). *Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil*, Applied and Environmental Microbiology 73:7059-7066.

Fierer, N. and Lennon, J. T. (2011). *The generation and maintenance of diversity in microbial communities*, American Journal of Botany 98:439-448.

Geissen, V.; Sánchez-Hernández, R.; Kampichler, C.; Ramos-Reyes, R.; Sepulveda-Lozada, A.; Ochoa-Goana, S.; de Jong, B.; Huerta-Lwanga, E. and Hernandez-Daumas, S. (2009). *Effects of land-use change on some properties of tropical soils - an example from Southeast Mexico*, Geoderma 151:87-97.

Giongo, A.; Crabb, D. B.; Davis-Richardson, A. G.; Chauliac, D.; Mobberley, J. M.; Gano, K. A.; Mukherjee, N.; Casella, G.; Roesch, L. F.; Walts, B. and et al. (2010). *PANGEA: pipeline for analysis of next generation amplicons*, The ISME Journal 4:852-861.

Good, I. J. (1953). *The population frequencies of species and the estimation of population parameters*, Biometrika 40:237-264.

Hamady, M.; Lozupone, C. and Knight, R. (2009). *Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data*, The ISME Journal 4:17-27.

Hamady, M.; Walker, J. J.; Harris, J. K.; Gold, N. J. and Knight, R. (2008). *Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex*, Nature Methods 5:235-237.

van der Heijden, M. G. A.; Bardgett, R. D. and van Straalen, N. M. (2008). *The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems*, Ecology Letters 11:296-310.

Hossain, Z. and Sugiyama, S.-i. (2011). *Geographical structure of soil microbial communities in northern Japan: effects of distance, land use type and soil properties*, European Journal of Soil Biology 47:88-94.

Jangid, K.; Williams, M. A.; Franzluebbers, A. J.; Schmidt, T. M.; Coleman, D. C. and Whitman, W. B. (2011). *Land-use history has a stronger impact on soil microbial community composition than aboveground vegetation and soil properties*, Soil Biology and Biochemistry 43:2184-2193.

Jones, S. E. and Lennon, J. T. (2010). *Dormancy contributes to the maintenance of microbial diversity*, Proceedings of the National

Academy of Sciences 107:5881-5886.

Kizilkaya, R and Dengiz, O. (2010). *Variation of land use and land cover effects on some soil physico-chemical characteristics and soil enzyme activity*, Zemdirbyste-Agriculture 97:15-24.

Kunin, V.; Engelbrektson, A.; Ochman, H. and Hugenholtz, P. (2010). *Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates*, Environmental Microbiology 12:118-123.

Kuramae, E.; Gamper, H.; van Veen, J. and Kowalchuk, G. (2011). *Soil and plant factors driving the community of soil-borne microorganisms across chronosequences of secondary succession of chalk grasslands with a neutral pH*, FEMS Microbiology Ecology 77:285-294.

Lemos, L. N.; Fulthorpe, R. R.; Triplett, E. W. and Roesch, L. F. (2011). *Rethinking microbial diversity analysis in the high throughput sequencing era*, Journal of Microbiological Methods 86:42-51.

Lennon, J. T. and Jones, S. E. (2011). *Microbial seed banks: the ecological and evolutionary implications of dormancy*, Nature Reviews Microbiology 9:119-130.

Liebich, J.; Schadt, C. W.; Chong, S. C.; He, Z.; Rhee, S.-K. and Zhou, J. (2006). *Improvement of oligonucleotide probe design criteria for functional gene microarrays in environmental applications*, Applied and Environmental Microbiology 72:1688-1691.

Liu, Z.; Lozupone, C.; Hamady, M.; Bushman, F. D. and Knight, R. (2007). *Short pyrosequencing reads suffice for accurate microbial community analysis*, Nucleic Acids Research 35:e120.

Lozupone, C. A. and Knight, R. (2007). *Global patterns in bacterial diversity*, Proceedings of the National Academy of Sciences 104:11436-11440.

Ludwig, J. A. (1988). *Statistical ecology: a primer in methods and computing,* New York: John Wiley & Sons.

Martiny, J. B. H.; Bohannan, B. J.; Brown, J. H.; Colwell, R. K.; Fuhrman, J. A.; Green, J. L.; Horner-Devine, M. C.; Kane, M.; Krumins, J. A.; Kuske, C. R. and et al. (2006). *Microbial biogeography: putting microorganisms on the map*, Nature Reviews Microbiology 4:102-112.

Miki, T.; Ushio, M.; Fukui, S. and Kondoh, M. (2010). *Functional diversity of microbial decomposers facilitates plant coexistence in a plant-microbe-soil feedback model*, Proceedings of the National Academy of Sciences 107:14251-14256.

Mitchell, R. J.; Hester, A. J.; Campbell, C. D.; Chapman, S. J.; Cameron, C. M.; Hewison, R. L. and Potts, J. M. (2010). *Is vegetation composition or soil chemistry the best predictor of the soil microbial community?*, Plant and Soil 333:417-430.

Nüsslein, K. and Tiedje, J. M. (1999). *Soil bacterial community shift correlated with change from forest to pasture vegetation in a tropical soil*, Applied and Environmental Microbiology 65:3622-3626.

Navarrete, A. A.; Cannavan, F. S.; Taketani, R. G. and Tsai, S. M. (2010). *A molecular survey of the diversity of microbial communities in different Amazonian agricultural model systems*, Diversity 2:787-809.

Overbeck, G. E.; Müller, S. C.; Fidelis, A.; Pfadenhauer, Jö.; Pillar, V. D.; Blanco, C. C.; Boldrini, I. I.; Both, R. and Forneck, E. D. (2007). *Brazil's neglected biome: The South Brazilian Campos*, Perspectives in Plant Ecology, Evolution and Systematics 9:101-116.

Roesch, L. F. W.; Vieira, F. C. B.; Pereira, V. A.; Schünemann, A. L.; Teixeira, I. F.; Senna, A. J. T. and Stefenon, V. M. (2009).

The Brazilian Pampa: a fragile biome, Diversity 1:182-198.

Schloss, P. D.; Westcott, S. L.; Ryabin, T.; Hall, J. R.; Hartmann, M.; Hollister, E. B.; Lesniewski, R. A.; Oakley, B. B.; Parks, D. H.; Robinson, C. J. and et al. (2009). *Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities*, Applied and Environmental Microbiology 75:7537-7541.

Shannon, C. E. (2001). *A mathematical theory of communication*, Bell System Technical Journal 5:3-55.

Singh, B. K.; Millard, P.; Whiteley, A. S. and Murrell, J. (2004). *Unravelling rhizosphere-microbial interactions: opportunities and limitations*, Trends in Microbiology 12:386-393.

Soergel, D. A. W.; Dey, N.; Knight, R. and Brenner, S. E. (2012). *Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences*, The ISME Journal 6:1440-1444.

Tan, Z. X.; Lal, R. and Wiebe, K. D. (2005). *Global soil nutrient depletion and yield reduction*, Journal of Sustainable Agriculture 26:123-146.

Tornquist, C. G.; Giasson, E.; Mielniczuk, J.; Cerri, C. E. P. and Bernoux, M. (2009). *Soil organic carbon stocks of Rio Grande do Sul, Brazil*, Soil Science Society of America Journal 73:975-982.

Wang, Q.; Garrity, G. M.; Tiedje, J. M. and Cole, J. R. (2007). *Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy*, Applied and Environmental Microbiology 73:5261-5267.

Winsley, T.; van Dorst, J. M.; Brown, M. V. and Ferrari, B. C. (2012). *Capturing greater 16s rRNA gene sequence diversity within the domain bacteria*, Applied and Environmental Microbiology 78:5938-5941.