Title

Soil microbial community dynamics and assembly under long-term land use change

Journal

FEMS Microbiology Ecology, 93(10)

ISSN

0168-6496

Authors

Goss-Souza, Dennis

Mendes, Lucas William

Borges, Clovis Daniel

et al.

Publication Date

2017-10-01

DOI

10.1093/femsec/fix109

Peer reviewed

FEMS Microbiology Ecology, 93, 2017, fix109

doi: 10.1093/femsec/fix109
Advance Access Publication Date: 1 September 2017
Research Article

RESEARCH ARTICLE

Soil microbial community dynamics and assembly under long-term land use change

Dennis Goss-Souza[1,2,3], Lucas William Mendes[1,†], Clovis Daniel Borges[1,2], Dilmar Baretta[4], Siu Mui Tsai[1,*] and Jorge L. M. Rodrigues[3,5]

[1]Center for Nuclear Energy in Agriculture, University of São Paulo, Piracicaba, SP, 13400–970, Brazil, [2]Applied Ecology Graduate Program, Luiz de Queiroz School of Agriculture, University of São Paulo, Piracicaba, SP, 13418–900, Brazil, [3]Department of Land, Air and Water Resources, University of California – Davis, Davis, CA 95616, USA, [4]Department of Soils and Sustainability, Santa Catarina State University, Chapecó, SC, 89815-630, Brazil and [5]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

*Corresponding author: Universidade de São Paulo Centro de Energia Nuclear na Agricultura Av. Centenario, 303, CP 96 Piracicaba, SP 13400-970 Brazil. Tel: +55-19-3429-4611; E-mail: tsai@cena.usp.br

One sentence summary: Deforestation affects the soil microbial communities causing loss of diversity in grassland areas, with consequences for taxa turnover and microbial functions.

Editor: Wietse de Boer

†Lucas William Mendes, http://orcid.org/0000-0003-0980-7006

ABSTRACT

We evaluated the bacterial and archaeal community dynamics and assembly in soils under forest, grassland and no-till cropping, using a high-throughput shotgun metagenomics approach. No significant alterations in alpha diversity were observed among different land uses, but beta diversity in grassland was lower than that observed in forest and no-till soils. Grassland communities showed assembly that predominantly followed the neutral model, i.e. high homogenizing selection with moderate dispersion, leading to biotic homogenization. Both no-till and forest soil communities were found to have assembly that predominantly followed a niche model, i.e. low rates of dispersal and weak homogenizing selection, resulting in maintenance of higher beta diversity relative to grasslands, indicating niche specialization or variable selection. Taken together, our results indicate that the patterns of assembly and their governing processes are dependent on the land use employed after deforestation, with consequences for taxa turnover and microbial functional potential.

Keywords: community ecology; deforestation; neutral and niche model; specialization; land use change; metagenome

INTRODUCTION

Change in land use due to agriculture has been identified as the most important driver of biodiversity loss (Chapin *et al.* 2000; Sala *et al.* 2000). This is particularly severe in tropical systems with rising societal demands for food, fiber and energy resources (Foley *et al.* 2005, 2011; Thomson *et al.* 2010), and an increasing number of studies have documented biodiversity losses for animals (Cardinale *et al.* 2012; McConkey and Farrill 2015) and plants (Feeley and Silman 2009; Pereira *et al.* 2010; Wardle *et al.* 2011). Recently, several studies have shown that microbial diversity losses also occur in tropical forests converted to grassland (Navarrete *et al.* 2011; Rodrigues *et al.* 2013; Mueller *et al.* 2014; Mendes *et al.* 2015b) or no-till systems (Figuerola *et al.* 2012; Lienhard *et al.* 2013; Carboretto *et al.* 2014), but the consequences of

Received: 5 May 2017; **Accepted:** 31 August 2017

land use change for the ecological processes ruling assembly remain unknown.

Ecologists have long established the mechanisms ruling community assembly, but quantifying the importance of the different processes for microorganisms remains a fundamental research challenge. Most ecological models based on selection were developed for interactions occurring among a few and/or discrete species (McArthur and Wison 1963; Leibold *et al.* 2004), but microbial communities are known to be more complex and species rich than those model/simpler communities (Tringe *et al.* 2005; Fierer and Lennon 2011). One of the challenges in working with communities containing a large number of species is that assembly configurations are nearly limitless. Thus, testing community ecology theories for microorganisms can only be achieved by developing a framework that takes into consideration the particulars of complex communities.

There are two contrasting models of the assembly of species into communities: (i) a niche-based (deterministic) model that considers abiotic and biotic factors as primary drivers of taxa differences, which means that a set of attributes may contribute to a variable selection effect on microbial communities (Carroll, Cardinale and Nisbet 2011); and (ii) a neutral (stochastic) model focusing on probabilistic and random events to explain the structuring of communities, which is based on homogenizing selection (Hubbell 2001; Alonso and McKane 2004; Dini-Andreote *et al.* 2015). Although only a few microbiological studies have examined the above models directly, a consensus has built up that considers microbial community assembly as the interplay between stochastic and deterministic processes (Ofiteru *et al.* 2010; Stegen *et al.* 2012; Ferrenberg *et al.* 2013; Mendes *et al.* 2014; Dini-Andreote *et al.* 2015). Indeed, experimental support has been provided for the importance of niche-based and stochastic factors on the assembly of soil microbial communities (Ofiteru *et al.* 2010; Barberán *et al.* 2012; Wang *et al.* 2013). Some results have shown that niche and neutral models of assembly can act in a combined way, ranging from competitive to stochastic exclusion. In this sense, the absence of immigration, leads to a complementarity of niches by competitive exclusion (i.e. taxa turnover). However, in some cases immigration may prevent the establishment of a community with high similarity by rescuing rare species from extinction (Gravel *et al.* 2006). Thus, the assembly results could be described as a set of complementary and redundant species, with environmental variables and immigration rates governing microbial distribution in ecosystems. Meanwhile, it remains of limited understanding how these processes take place in soils under different land uses and the consequences of forest-to-agriculture conversion for community assembly.

In this study, we examined the ecological processes governing microbial (Bacteria and Archaea) community assembly in agricultural soils after long-term conversion of the Brazilian Atlantic Rainforest, a biome with the highest number of endemic plant and animal species among all terrestrial biomes (Myers *et al.* 2000; Ribeiro *et al.* 2009). We tested the hypothesis that ecological processes differ across long-term land uses. In a corollary hypothesis, we tested whether variation in microbial functional potential is dependent on assembly. By combining high-throughput shotgun metagenomics and physicochemical soil properties, we aimed (i) to determine alterations in microbial taxonomical and functional potential profiles imposed by land uses; (ii) to assess the ecological processes affecting community assembly in forest, grassland and no-till soils; and (iii) to evaluate community assembly under neutral- and niche-based models.

In order to reach our goals and generate a comprehensive and integrative conclusion to our hypotheses, we combined several multivariate ordination and pairwise comparison methods. Thus, we described a three-dimensional (taxonomic, genetic and environmental) statistical inference about the contribution of ecological processes governing microbial assembly and functional potential, taking into account how environmental variables can selectively drive microbial communities in soils.

MATERIAL AND METHODS

Site description and soil collection

The sampling sites were located within the Brazilian Atlantic Rainforest Biome composed of remnants of the original forest cover, long-term grassland and no-till cropping sites, varying from 20 to 25 years of conversion (sampling site coordinates can be found in Supplementary Fig. S1a). The forest sites comprised a natural transition between mixed ombrophilous (wet) forest and semi-deciduous forest (details are given in the Supplementary data). Grassland sites were characterized as perennial with a predominance of *Axonopus affinis* (Poaceae). No-till cropping systems were characterized by successive cultivation of wheat in the winter, followed by maize and soybean in the summer. The site selection criteria were based on previous exploratory visits and the history of land use and management acquired through interviews with farmers and regional experts.

Samples were collected in July 2011 and January 2012, comprising winter and summer, respectively, in the municipalities of Chapecó (27°2′S; 52°41′W) and Xanxerê (26°50′S; 52°27′W), State of Santa Catarina, Brazil (Supplementary Fig. S1a). Non-deformed soil samples from the 0–10 cm profile were taken with sterile PVC tubes (5 cm diameter × 10 cm depth), yielding at least 500 g of soil each. A total of 36 soil samples were collected (3 sites × 3 replicates per site × 2 municipalities × 2 sampling periods). Each sample was composed of three subsamples in a geogrid system, equidistantly every 30 m from each other (Supplementary Fig. S1b). The litter layer was removed prior to sampling collection. Samples for microbial analysis were kept on dry ice and transported to the Cell and Molecular Biology Laboratory, Center for Nuclear Energy in Agriculture (Piracicaba, Brazil), while samples for physical and chemical analyses were maintained at 4°C and transported to the Soil Physicochemistry and Fertility Analysis Laboratory, Santa Catarina State University (Lages, Brazil).

Soil elemental analyses

A total of 35 soil parameters were analyzed. Soil chemical and physical properties were determined for each sample based on 500 g of soil, performed at the Laboratory of Soil Analysis at 'Luiz de Queiroz' College of Agriculture (ESALQ/USP, Piracicaba, Brazil). Soil pH was measured in a 1:2.5 soil/water suspension. Exchangeable Al^{3+}, Ca^{2+} and Mg^{2+} were extracted with 1 M KCl. Calcium and magnesium were determined by atomic absorption spectrometry and aluminum by acid–base titration. Phosphorus and potassium were extracted by ion-exchange resin. Potential acidity (H+Al) was estimated by an equation based on the pH determined in Shoemaker-McLean-Pratt buffer solution. Available micronutrients (Fe, Mn, Zn and Cu) were extracted by Mehlich-1 and determined by atomic absorption spectrometry. Boron was extracted with hot water and determined by spectrophotometry with azomethine-H at 420 nm. Some of the results allowed the calculation of other parameters such as exchangeable bases, the sum of Ca, Mg and K; cation exchange capacity, the sum of

Ca, Mg, K, Al and H; base saturation (V%), the percentage relation between exchangeable bases and cation exchange capacity; and Al saturation (m%), the percentage relation between exchangeable Al and cation exchange capacity. Total C, H, N and S were extracted and determined by the combustion catalytic oxidation method in a total organic carbon analyzer. Soil texture was determined using a Bouyoucos densimeter, after shaking the soil vigorously with 1 M NaOH as dispersant. Penetration resistance was measured by a penetrometer. Total porosity, macro- and microporosity were measured by standard methods. The gravimetric moisture was obtained as a percentage, through the difference between the weight of the sample at the moment of sampling and its dry weight after 48 h in an incubator at 105°C. Soil density was measured by Kopecky's ring method. Biopores were calculated by measuring the pores' continuity. Total porosity was calculated through the saturation method. Microporosity was obtained by the tension table method. Macroporosity was calculated by difference, deducting the microporosity from the total porosity. These parameters were analyzed at the Soil Physicochemistry and Fertility Analysis Laboratory, Santa Catarina State University, Lages, Brazil, following routine methodology (Gee and Bauder 1986; Tedesco *et al.* 1995; Claessen *et al.* 1997; Dhaliwal *et al.* 2011). Soil N-NH_4^+ and N-NO_3^- extraction and determination (Keeney and Nelson 1982) were performed at the Laboratory of Soil Fertility, University of São Paulo, Piracicaba, Brazil.

Soil DNA extraction and sequencing of shotgun metagenomic libraries

Total DNA was extracted from soil samples (0.25g) using the PowerLyzer PowerSoil DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA), following the manufacturer's instructions. Soil DNA was visualized in 1% Tris-saline buffer with sodium boric acid (Brody and Kern 2004) agarose gel electrophoresis and quantified with the Qubit fluorometer (Thermo Fischer Scientific, Waltham, MA, USA).

Shotgun metagenomic libraries were constructed with the Nextera DNA Sample Preparation Kit (Illumina Inc., San Diego, CA, USA), as specified by the manufacturer. Reactions were carried out in an Applied Biosystems GeneAmp PCR System 9700 (Thermo Fischer Scientific). Products were purified with the GFX PCR DNA and Gel Band Purification Kit (GE Healthcare, Little Chalfont, UK) and resuspended in 25 μl of water. The insertion of barcoding indexes was performed with the Nextera Index Kit (Illumina Inc.) with a combination of 4 × 6 indexes, totaling 24 samples per sequencing run. Products were then purified with Agencourt AMPure XP reagents (Beckman Coulter, Brea, CA, USA), on a magnetic rack, as per the manufacturer's instructions. DNA was dried at 37°C for 3 min in a Biological Oxygen Demand incubator (BOD MA415, Marconi, Piracicaba, Brazil). Samples were resuspended in 32.5 μl of RSB Buffer (supplied in the Nextera Index Kit), incubated for 2 min and transferred to a new 96-well plate. DNA samples were quantified with the Qubit fluorometer (Thermo Fischer Scientific) and equimolar concentrations (80 ng) of shotgun metagenomic libraries pooled for sequencing. Prior to sequencing, the pooled samples were quantified with the KAPA SYBR FAST Universal qPCR Kit (Kappa Biosystems, Woburn, MA, USA) according to the manufacturer's protocol. The PCR reaction was performed in the StepOne Plus thermocycler (Thermo Fischer Scientific) and consisted of a predenaturation at 95°C for 5 min, followed by 35 cycles of denaturation at 95°C for 30 s and annealing at 60°C for 45 s.

The final step of sequencing was performed with the MiSeq Reagent Kit V2 (Illumina Inc.). Ten microliters of NaOH (0.2 M) were added to 10 μl of pooled shotgun metagenomic libraries (2 nM). The solution was mixed briefly in a vortex and incubated at room temperature for 5 min followed by addition of 980 μl of HT1 buffer. A total of 600 μl of the mixture (17 pM of DNA) was loaded in the sequencing cartridge and inserted into the MiSeq Desktop Sequencer (Illumina Inc.) at the Center for Nuclear Energy in Agriculture, University of Sao Paulo, Brazil. The sequencing consisted of a paired-end reads (2 × 250 bp) 500 cycles/39 h run.

Annotation of metagenomic data and analysis

Raw sequences were sorted based on assigned barcodes, followed by filtering to discard sequences with low-quality bases (quality score lower than 20) under default parameters in Miseq software (Illumina Inc.). The remaining paired-end sequences were merged using FLASH version 1.2.5 (Magoč and Salzberg 2011) and additional quality trimming was conducted using the Phred algorithm with SeqyClean script (http://github.com/ibest/seqyclean). Pair-ended DNA sequences were annotated with Metagenomics Rapid Annotation (MG-RAST) pipeline version 3.3.9 (Meyer *et al.* 2008). Taxonomic and functional potential profiles were generated using the normalized abundance of sequences matches to the M5nr (Wilke *et al.* 2012) and SEED (Overbeek *et al.* 2013) databases, respectively. A table of frequency of hits to each individual taxon (taxonomy) or subsystem (function) for each metagenome was generated and normalized by dividing by the total number of hits to remove bias indifference in read lengths and sequencing efforts. To identify hits, BlastX searches with a minimum alignment length of 50 bp and an *E*-value cut-off of 1×10^{-5} was used. The shotgun metagenome data are available at MG-RAST under the project ID 5426.

Statistical analysis

All statistical analyses were performed comparing the land uses, two seasonal effects and the geography (the two municipalities). Data were presented together when the effect of season and/or geography was negligible. The normalized matrices from taxa (M5nr database) and functional subsystems (SEED database) generated from MG-RAST were used for downstream analyses. Shannon's alpha-diversity and Whittaker's global beta diversity were calculated based on the taxonomic relative abundance matrix at the genus level and the functional potential relative abundance matrix at subsystem Level 2, using PAST software, v.3.0 (Hammer, Harper and Ryan 2001). In order to compare the structure of the microbial communities among samples, we conducted nonmetric multidimensional scaling plots based on Bray–Curtis dissimilarity matrix using the above described taxonomic and functional potential abundance matrices. Resulting coordinates were used to calculate the mean distance between neighbor samples (MDN) through nearest neighbor algorithm and the mean density of coordinates (MDC) through convex hull area. Additionally, we used permutational multivariate analysis of variance (PERMANOVA; Anderson 2001) to test whether sample categories harbored significantly different community structures. Nonmetric multidimensional scaling rotations of PCA were performed with Canoco software v.5.0 (Lepš and Šmilauer 2005), while MDN, MDC and PERMANOVA were performed using PAST software.

To determine the differences in abundance of bacterial and archaeal groups among soil samples, the Statistical Analysis of Metagenomic Profiles (STAMP) software v.3.0 (Parks and Beiko 2010) was used. The *q*-values were calculated using two-sided Welch's *t*-test (Welch 1934), while confidence intervals were calculated using the Newcombe–Wilson method (Newcombe 1998) with corrections performed using the Benjamini–Hochberg false discovery rate (Benjamini and Hochberg 1995).

In order to identify the main environmental drivers of microbial taxonomic assembly and functional potential profile, we performed distance-based redundancy analysis (Ramette and Tiedje 2007; Blanchet, Legendre and Borcard 2008) of Bray–Curtis dissimilarity matrices, with stepwise forward selection. Only non-colinear (inflation factor < 20) and highly significant (P-value ≤ 0.05) environmental factors were considered in the model. In addition, a variation partitioning of redundancy analysis, generated by principal coordinates analysis of neighbor matrices with stepwise forward selection was performed to evaluate the possible single and joint effects of environmental variables, space and time on the variation of taxonomic and functional potential profiles of microbial communities in each land use. Latitude and longitude were used as constraining spatial coordinates. Both analyses were performed using Canoco software, v.5.0.

In addition, network analyses were performed to determine correlations within taxonomic groups and functional potential, their interactions and their connectivity with environmental factors. For this, all-possible rank pairwise Spearman's correlation with Benjamini–Hochberg false discovery rate correction was calculated with the Vegan package from R software, v.3.1.2 (R Development Core Team 2007). Correlation patterns across land uses were visualized with Gephi software v.0.8.2 (Bastian, Heymann and Jacomy 2009), using only high Spearman's pairwise correlations, with cut-off at $-0.7 \leq R \geq 0.7$, with P-value ≤ 0.01 for taxonomy, function and environmental variables.

Analysis of the microbial community assembly

To test whether neutral or niche-based mechanisms best explain the microbial community assembly, we calculated rank abundance distributions and immigration rates, based on absolute abundances for taxonomic genus level matrices, for each sample, from each land use system. We fitted our data to five different rank abundance models, two of them fitted for neutral- (Hubbell 2001; Volkov *et al.* 2003) and three for niche-based assembly (Ofiteru *et al.* 2010; Pholchan *et al.* 2013). The neutral zero-sum model was calculated with TeTame software v.1.9 (Jabot, Etiene and Chave 2008). The neutral broken stick model (null model) and three niche-based models (log-normal, pre-emption and Zipf) were calculated using the function 'radfit' from the Vegan package in R software, version 3.1.2. Models were compared based on the Akaike information criterion—the lower the value, the better the microbial community fitted to a specific model. The dispersal rate was calculated with TeTame software, through Etienne's formula (Etienne and Alonso 2005).

RESULTS

Soil attributes

We performed Tukey's honest significant difference test for 35 soil physicochemical and three geographical parameters measured for forest, grassland and no-till samples taken during winter and summer (Supplementary Table S1). The total N, total C, soil organic matter, potential acidity (H+Al), aluminum (Al^{3+}), Al^{3+} saturation, total porosity, macroporosity and bioporosity decreased after long-term conversion for both grassland and no-till soils, while pH, PO_4^{2-}, Ca^{2+}, Mg^{2+}, base saturation, soil density, microporosity and penetration resistance increased ($P < 0.001$). Concentrations of K^+ and $N\text{-}NH_4^+$ only increased in the non-till system. When comparisons were performed between sampling periods, $N\text{-}NH_4^+$, $N\text{-}NO_3^-$, total S concentration and the C:N ratio were higher for all land uses during the summer ($P < 0.001$).

Taxonomic and functional potential diversities across land uses, geographical locations and sampling times

A total of 27 million of sequences were obtained for 36 soil samples using a shotgun metagenomic approach. A list the average and minimum number of sequences generated per sample, as well as the number of hits to each database for each sample (see Supplementary Table S2). The taxonomic and functional potential diversities across land uses were measured with both Shannon's alpha diversity and Whittaker's global beta diversity indices for two municipalities in winter and summer. There were no significant changes in alpha diversity with forest-to-agriculture conversion ($P > 0.05$). Although not significant, alpha diversities were higher when sampling was performed in the summer (Fig. 1a and 1b). Whittaker's global beta diversity values were lower for grasslands in both municipalities, with no changes between the forest and no-till system (Fig. 1a and b). There was a shift in Whittaker's global beta diversity between locations and sampling times, but when an analysis for explanatory variables with stepwise forward selection was performed, both were found not to be significant.

Community structure and composition across land uses

Based on the result that grassland showed a lower Whittaker's global beta diversity index in comparison with forest and no-till for both locations, we asked whether community structure displayed distinct patterns among different land uses. A non-metric multidimensional scaling plot showed that grassland samples grouped together and were separated from forest and no-till samples (PERMANOVA, $P < 0.05$; Fig. 2). The nearest neighbors analysis of forest microbial communities showed significant over-dispersion on taxa assembly (MDN = 0.048; MDC = 250; $P < 0.001^{***}$), while no-till (MDN = 0.038; MDC = 408; $P = 0.14$) and grassland (MDN = 0.027; MDC = 943; $P = 0.06$) communities clustered according to land use (Fig. 2). When there is a combination of both high distances across samples and low density of grouped samples with significant P-value, it suggests that samples are over-dispersed; the opposite means that samples are grouped and consequently more homogeneous (Mendes *et al.* 2014).

Sequence reads comprised 32 phyla with 28 belonging to Bacteria and four to Archaea based on a non-redundant assignment of the M5nr database. Overall, the most abundant phylum was Proteobacteria with an average of 46.6% of the sequences, followed by the phylum Actinobacteria representing 20.9% of the total. The Acidobacteria and Firmicutes were represented by a total of 9.5 and 4.1%, respectively. Sequences matching the phyla Verrucomicrobia, Planctomycetes, Chloroflexi, Cyanobacteria and Bacteroidetes ranged from 2 to 3.3%. The same patterns were observed when comparing summer and winter

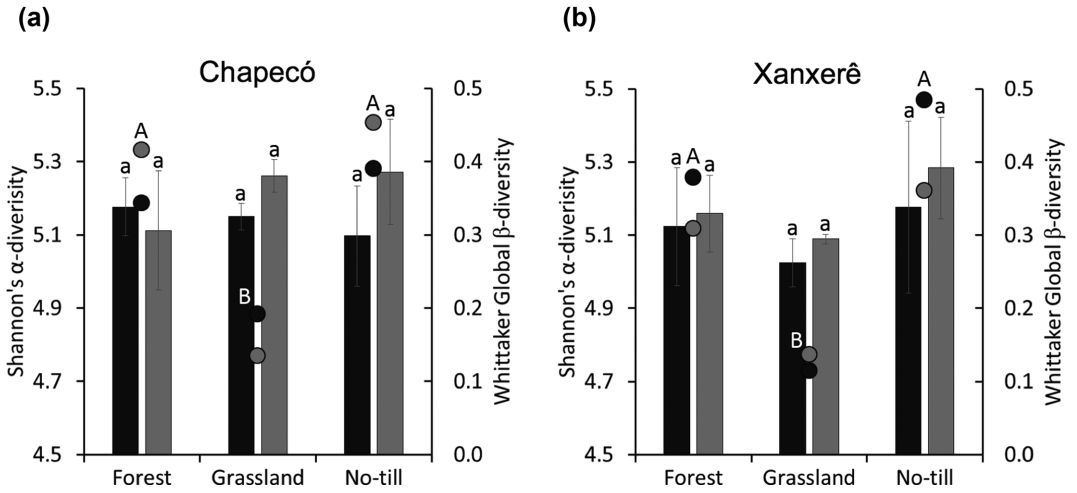

Figure 1. Soil microbial diversity at taxonomic genus level based on shotgun metagenomics of three land uses: forest, grassland and no-till at two municipalities: Chapecó (**a**) and Xanxerê (**b**). Bars represent Shannon's alpha-diversity and dispersion points represent Whittaker's Global beta-diversity for two sampling seasons: winter (black) and summer (gray). Error bars show standard deviation of triplicate samples. Different letters refer to significant differences of alpha (lower case letters) and beta (upper case letters) diversity between treatments based on Tukey's test ($P < 0.05$)

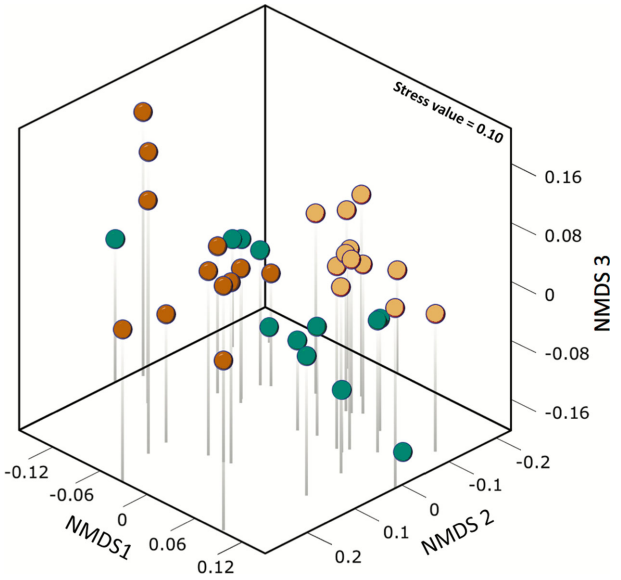

Figure 2. Non-metric multidimensional scaling plot generated from a Bray–Curtis dissimilarity matrix of metagenomic profiles based on taxonomic genus level. Colored circles as follows: forest, green; grassland, light-brown; no-till, dark-brown.

sampling periods (Fig. 3a). When the relative abundance of bacterial and archaeal groups was compared among samples, long-term conversion of forest to grassland and forest to no-till showed significant changes for 11 and 8 phyla, respectively (Fig. 3b and c, *q*-values are shown in the panels). The relative abundances of the Planctomycetes, Verrucomicrobia, Fibrobacteres, Synergistetes and Euryarchaeota increased with conversion to grassland (Fig. 3b). The largest proportional decrease in response to conversion was by the phylum Proteobacteria, followed by Gemmatimonadetes, Nitrospirae and Thaumarchaeota. The long-term no-till conversion resulted in an increase (combined total of 3%) in relative abundance of several phyla, namely Bacteroidetes, Chloroflexi, Gemmatimonadetes, Nitrospirae, Planctomycetes, Dictyoglomi and Poribacteria, with a corresponding decrease (2.7%) for the Acidobacteria (Fig. 3c).

Soil functional potential shifts after long-term land use change

In order to verify whether the taxonomic differences observed with forest conversion were reflected at functional potential level, the shotgun metagenomic libraries were classified according to the SEED subsystems database. Out of 48 possible categories, soil metagenomes were assigned to 28 different functional subsystems. Thirteen categories were significantly altered by forest conversion with no seasonal effect (Fig. 4, corrected *q*-values are shown in the lower panels). Long-term grasslands resulted in changes in relative abundance for a total of 12 categories with increases in potential functions associated with carbohydrates, respiration, virulence, sulfur metabolism, regulation and cell signaling, motility and chemotaxis, and iron acquisition and metabolism, while amino acids and derivatives, fatty acids, lipids and isoprenoids, cofactors and vitamins, metabolism of aromatic compounds, and photosynthesis decreased (Fig. 4b). When forests were converted to long-term no-till systems, only four categories were significantly changed, with potential functions associated to amino acids and nucleosides and nucleotides increasing and those related to photosynthesis and virulence decreasing (Fig. 4c).

Environmental factors associated with taxonomic and functional potential patterns

We performed a Euclidean distance-based redundancy analysis for 39 explanatory variables (35 soil physicochemical parameters, three geographical coordinates and time of sampling) to identify factors modulating soil microbial community structure at taxonomic and functional potential levels. When all 39 factors were computed in our model, the total variations in taxonomic and functional potential community structures were 40.2 and 4.8%, respectively. However, when a stepwise forward selection was performed, a subset of predictors ($P_{adjusted} \leq 0.05$) with the highest scores contributed to 36.8% of the explanation in taxonomic structure (Fig. 5a, note that 31.7% is represented in the first two axes of the distance-based redundancy analysis, while 5.1% is explained in the remaining axes). The P content was the main factor associated with differences in community

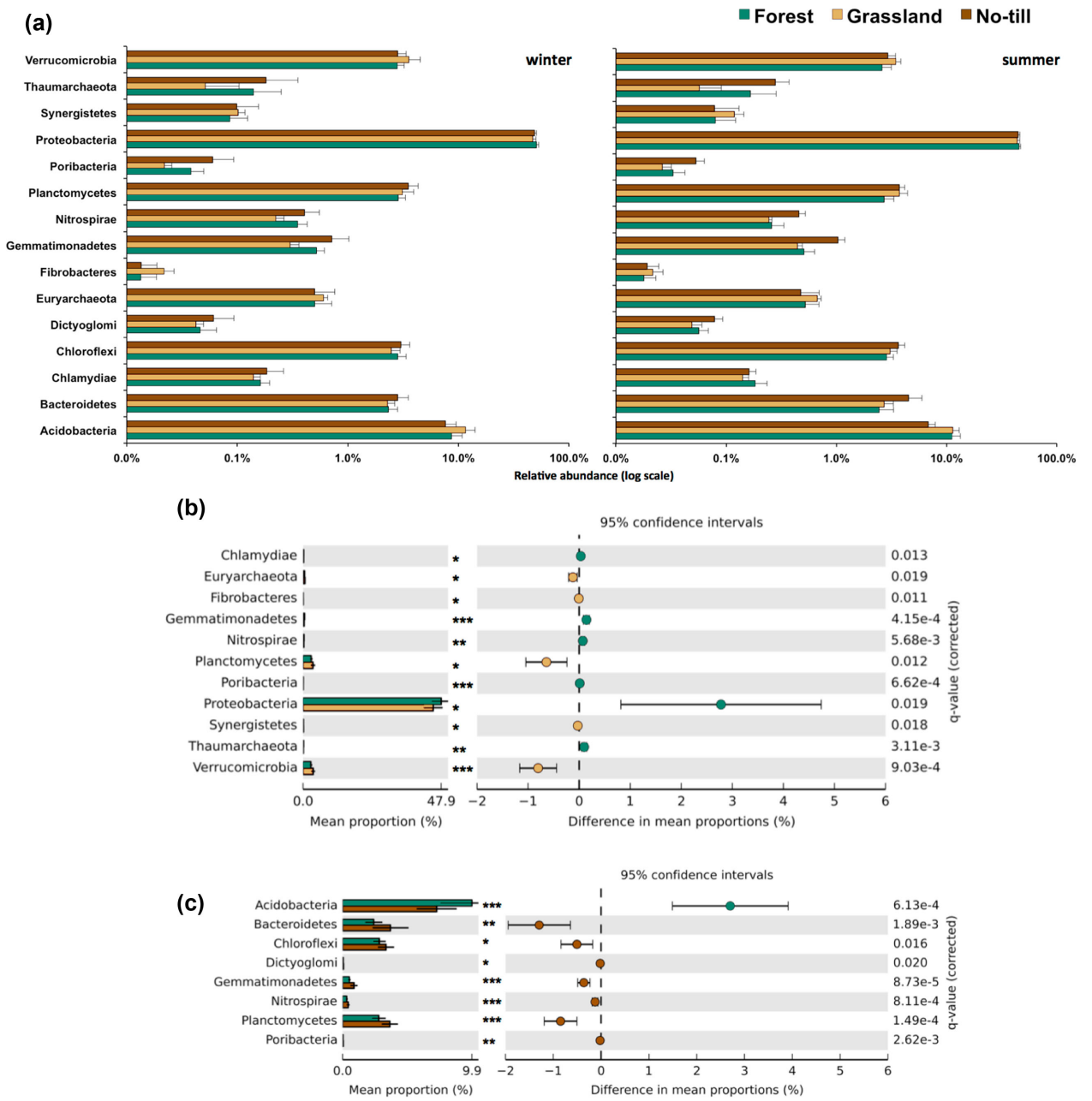

Figure 3. Relative abundance of soil bacterial and archaeal communities at phylum level based on shotgun metagenomics. Only significantly altered taxa are shown. Forest, green bars and circles; grassland, light-brown bars and circles; no-till, dark-brown bars and circles. Error bars show standard deviation. (**a**) Percentages of sequence reads are shown for two sampling seasons. (**b**) Differences in mean proportions for comparison between forest and grassland. (**c**) Differences in mean proportions are shown for comparison between forest and no-till. Welch's *t*-test with corrected *q*-values calculated using the Benjamini–Hochberg false discovery rate was performed for the significance levels: $^{*}q \leq 0.05$, $^{**}q \leq 0.01$, $^{***}q \leq 0.001$.

composition, explaining 20.9% of the total variation (pseudo-$F = 9.0$; $P_{adjusted} = 0.01$), followed by Ca^{2+}/Mg^{2+} ratio and N-NO_3^- content explaining 8.1% (pseudo-$F = 3.8$; $P_{adjusted} = 0.025$) and 7.8% (pseudo-$F = 3.9$; $P_{adjusted} = 0.031$), respectively. When the same analysis was performed for functional potential structure, the forward-selected variables ($P_{adjusted} \leq 0.05$) explained 28.7% of the total variation (Fig. 5b). Phosphorus was again the major factor contributing to 14.3% (pseudo-$F = 5.7$; $P_{adjusted} = 0.014$) of the explained variation, followed by penetration resistance (8.3%; pseudo-$F = 3.5$; $P_{adjusted} = 0.014$), C/N ratio (6.6%; pseudo-$F = 3.0$; $P_{adjusted} = 0.03$) and longitude (5.8%; pseudo-$F = 2.8$, $P_{adjusted} = 0.01$) (Fig. 5b, only 28.7% is represented in the first two axes).

We calculated pairwise Spearman's correlation coefficients to establish co-occurrence patterns between all taxonomic, functional potential and environmental data within each land use (Supplementary Tables S3 and S4). Grassland samples had the highest number of significant correlations (637), followed by no-till (184) and forest (65). All land uses showed a higher number of total positive correlations in comparison with the negative ones. Correlations were consistently higher for grassland at taxa–taxa, function–function, taxa–function and

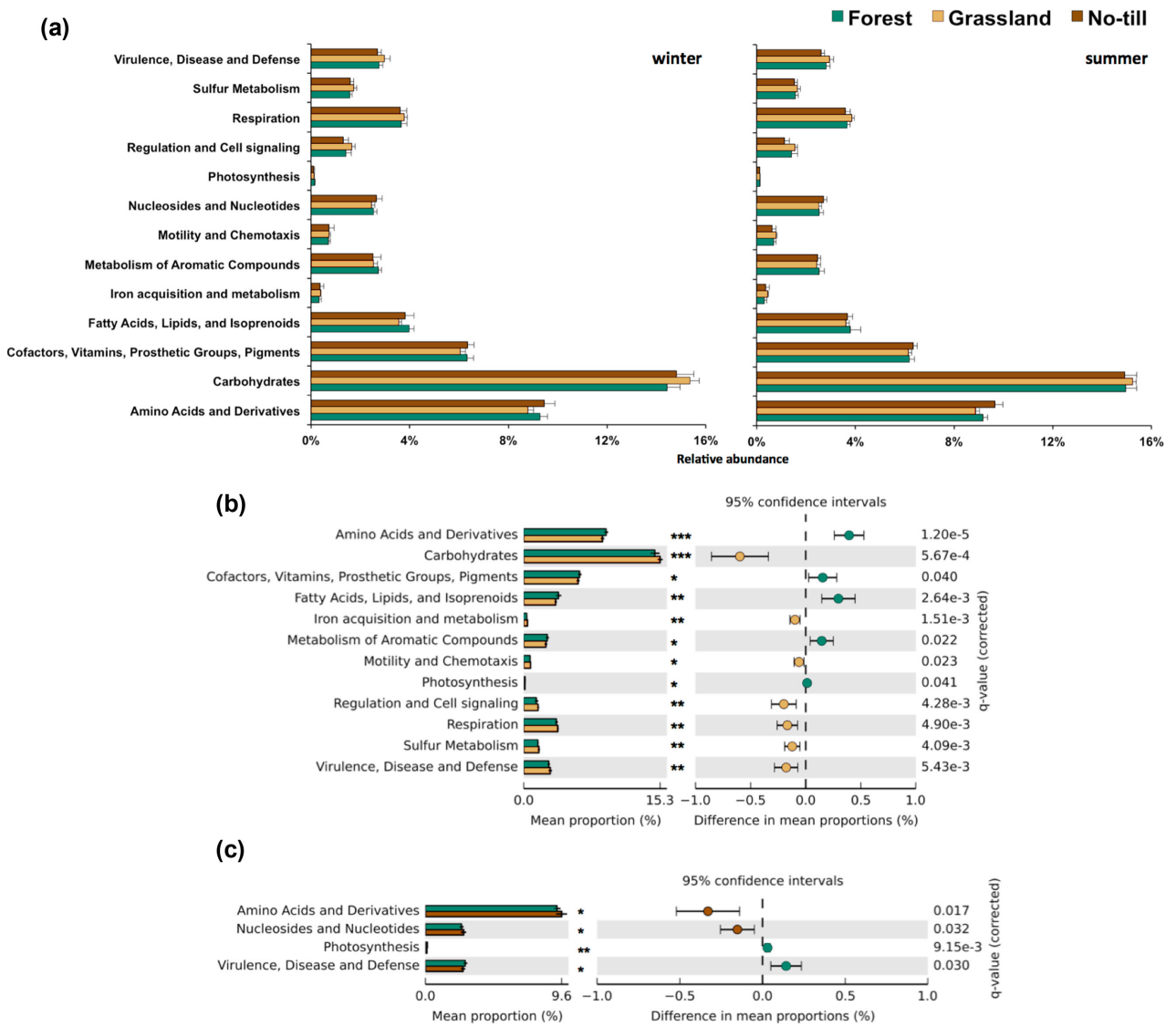

Figure 4. Relative abundance of soil potential functional categories at subsystems level 1 based on shotgun metagenomics. Only significantly altered categories are shown. Forest, green bars and circles; grassland, light-brown bars and circles; no-till, dark-brown bars and circles. Error bars show standard deviation. (**a**) Percentages of sequence reads are shown for two sampling seasons. (**b**) Differences in mean proportions for comparison between forest and grassland. (**c**) Differences in mean proportions for comparison between forest and no-till. Welch's *t*-test with corrected *q*-values calculated using the Benjamini–Hochberg false discovery rate was performed for the significance levels: $^{*}q \leq 0.05$, $^{**}q \leq 0.01$, $^{***}q \leq 0.001$.

taxa–environment combinations in comparison with no-till and forest, with the latter showing the lowest numbers of correlations (positive and negative) for all combinations. In order to visualize the connectivity patterns among the three dataset, we generated networks for strong (Spearman's $\rho \geq 0.7$) and highly significant ($P \leq 0.01$) pairwise correlations (Fig. 6). The forest network had a larger diameter (7) and higher modularity (0.65) and average path length (2.65) in comparison with converted systems, but lower number of nodes (55), edges (65), average clustering (0.10) and average degree (2.40). The complexity of the network in grassland was higher than that measured for the no-till system, evidenced by a higher number of nodes, edges, positive and negative edges, average degree and average weighted degree (Supplementary Table S5). The grassland network structure was found to be non-modular (modularity = 0.11).

Four phyla, Firmicutes, Verrucomicrobia, Elusimicrobia and Thaumarchaeota, had a high number of positive correlations among themselves as well as with other factors in the forest network. The Proteobacteria, although not directly connected with the above phyla, presented a high number of correlations with functional potential categories and environmental variables. After long-term land use change, co-occurring phyla were altered. The taxa with more connections in the no-till network were Deinococcus–Thermus, Euryarchaeota, Nitrospirae, Planctomycetes and Proteobacteria, but none of them was predominant. The grassland network had a high number of correlations for Deinococcus–Thermus, Cyanobacteria, Aquificae, Tenericutes and Deferribacteres, which were strongly connected to one another. Functional potential categories with more correlations in the forest were metabolism of aromatic compounds, cell division and cell cycle, DNA metabolism, and

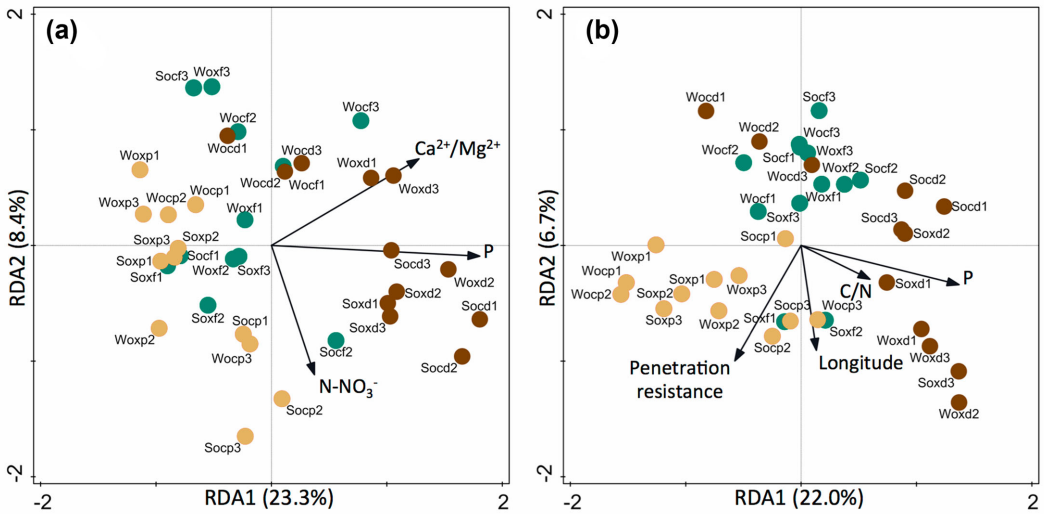

Figure 5. Distance-based redundancy analysis (RDA) of soil microbial communities at (**a**) genus level and (**b**) functional categories at level 2. Plots were generated using Euclidean distance matrices with 1000 Monte-Carlo permutations. Vectors represent environmental variables, which were forward selected ($P_{adjusted} \leq 0.05$). Samples are colored as follows: forest, green circles; grassland, light-brown circles; no-till, dark-brown circles.

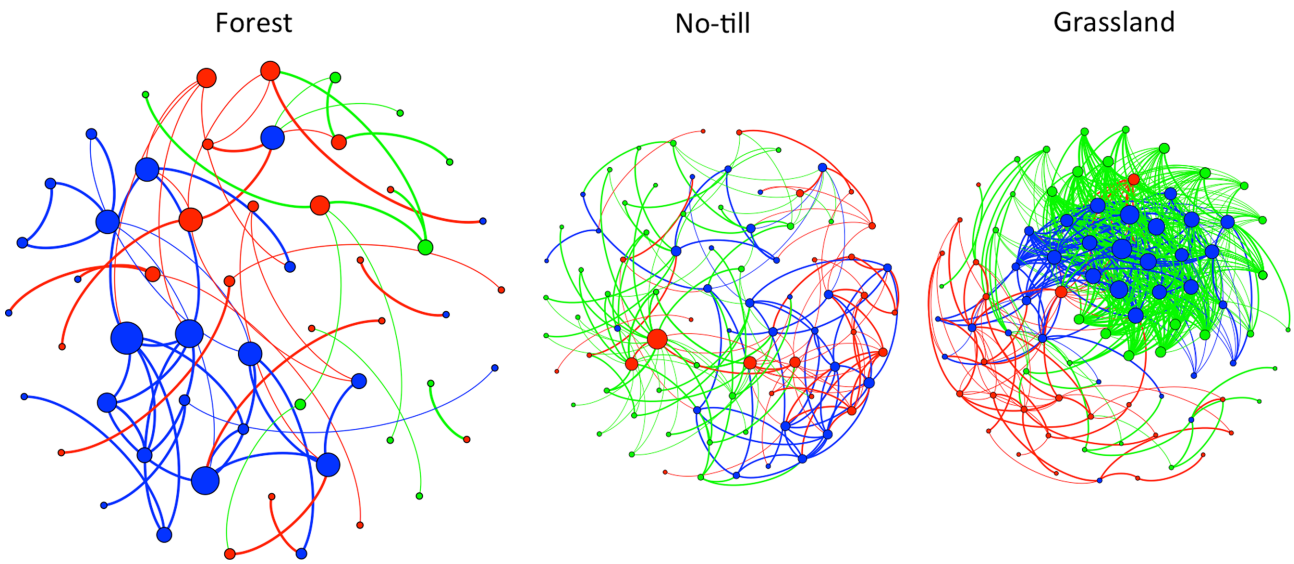

Figure 6. Networks based on Spearman's pairwise correlation matrices at phylum level (blue nodes), potential functional groups level 1 (red nodes) and environmental variables (green nodes). Bold edges represent strong and positive correlations ($\rho \geq 0.7$), while regular edges represent strong and negative correlations ($\rho \leq -0.7$). Only significant edges are shown ($P \leq 0.01$). Size of each node is proportional to its number of edges.

sulfur metabolism. Long-term forest conversion showed a high number of positive connections for potential functions related to metabolism of aromatic compounds and RNA metabolism in grassland and sulfur metabolism and metabolism of aromatic compounds in the no-till system.

Microbial community assembly across land uses

The data of all samples were fitted to theoretical species abundance distribution models to test whether neutral or niche-based mechanisms could best explain the composition and structure of microbial communities in the three datasets. The comparison of different rank abundance distribution models based on the corrected Akaike information criterion indicated that most of the forest and no-till samples fitted a lognormal distribution, while grassland samples fitted the zero-sum multinomial model (Supplementary Table S6). When the rate of dispersal was calculated for each community, forest and no-till samples had average values of 0.08 and 0.07, respectively, suggesting their microbial populations had a very low tendency to migrate, while grassland samples showed consistently higher dispersal rates (average = 0.38) (Supplementary Table S6). In addition, the weight of selection was calculated using a variation partitioning of redundancy analysis that decomposed environment, geography + time, and their interaction. The variation explained by selection in grassland was 83.8% ($P \leq 0.001$), followed by forest and no-till with 40.6% ($P \leq 0.05$) and 20.9%, respectively (Fig. 7a), implying that a strong selection is an important driver

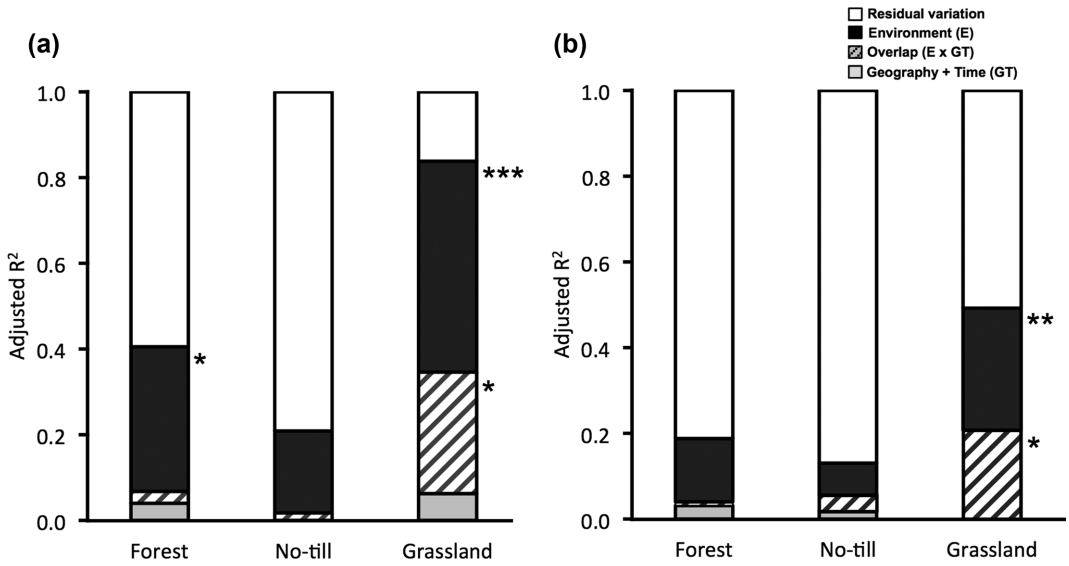

Figure 7. Variation partitioning of redundancy analysis with forward selection. Data show the adjusted coefficient of multiple determination (R^2) from simple effects of environmental variables (black), geography + time (gray), and their interactions (dashed gray). Latitude and longitude were used as constraining spatial coordinates. (**a**) Taxonomy at genus level based on M5nr database. (**b**) Function based on the SEED subsystems level 2 database. $*P_{adjusted} \leq 0.05$, $**P_{adjusted} \leq 0.01$, $***P_{adjusted} \leq 0.001$.

of community assembly in grassland, which leads to a community homogenization.

DISCUSSION

A few studies have simultaneously tested the importance of a range of abiotic factors, including climate and soil properties, and biotic factors, such as vegetation composition, across a wide range of spatial scales (Ramette and Tiedje 2007; Baker *et al.* 2009; Figuerola *et al.* 2012). According to de Vries *et al.* (2012), this represents a major gap in knowledge given the potential for both abiotic and biotic factors to explain variation in soil microbial communities. This gap of understanding of the factors that explain variation in microbial communities at larger spatial scales is surprising given their functional importance in regulating ecosystem processes, such as carbon and nitrogen cycling (Olivier *et al.* 2011) and the resistance of nutrient cycles to climate change-related disturbances (Thomson *et al.* 2010; de Vries *et al.* 2012).

Integrating dimensions to understand microbial patterns after long-term land use conversion

Many studies have described the consequences of land use change for microbial dynamics (Drenovsky *et al.* 2010; Fierer *et al.* 2012; Kuramae *et al.* 2012; Lauber *et al.* 2013), but the processes governing the alterations are often neglected (Knelman and Nemergut 2014). Our results showed that soil parameters are altered according to land use. It is well known that management practices affect soil quality and nutrient dynamics (Alam *et al.* 2013), and the microbial communities respond to these changes (Mendes *et al.* 2015a). Our data showed that phosphorus (P) and penetration resistance most explained the community variation. The addition of P fertilizers is a common practice to overcome P limitation in grassland systems (Bennett and Adams 2001), which may affect the activities of nitrifying and denitrifying microorganisms. Furthermore, several studies identified increased microbial respiration after P addition suggesting increased microbial activity (Mehnaz and Dijkstra 2016). Regarding penetration resistance, previous studies have shown that the conversion of forest to pasture areas causes an increase in pH and density, as well a decrease in soil porosity (Piccolo, Neill and Cerri 1994; Moraes *et al.* 2002), and such physical alteration in soil may have an influence on other chemical parameters. Furthermore, forest-to-grassland conversion has been also shown to alter microbial abundance as a result of changes in C and N concentrations (Nüsslein and Tiedje 1999; Cenciani *et al.* 2009; Rodrigues *et al.* 2013).

Several studies have shown that transformations in the environment due to land use change have a direct effect on the soil microbial community, by a response of specific microbial groups, which alters the structure, composition and diversity (Navarrete *et al.* 2011; Kuramae *et al.* 2012; Mendes *et al.* 2015a). Overall, we found that community alpha diversity did not vary among land uses in space and time; however, the beta diversity decreased in grassland. Our results agree with previous reports that tropical forest-to-agriculture conversion maintains (Lee-Cruz *et al.* 2013) or increases (Jesus *et al.* 2009) alpha diversity, but leads to significant decreases in beta diversity in grassland systems (Rodrigues *et al.* 2013; Mendes *et al.* 2015b). Our results also showed that grassland samples were grouped apart from forest and no-till, revealing a distinct microbial community structure. The analysis of the community composition showed that the forest soil presented higher abundance of Acidobacteria, Proteobacteria and Chlamydiae when compared with the agricultural sites. The phyla Acidobacteria and Proteobacteria have been described as a dominant group in forest soils, with significant decrease with land use change (Jesus *et al.* 2009; Miyashita *et al.* 2013; Navarrete *et al.* 2013). Members of Proteobacteria are important in global carbon, nitrogen and sulfur cycling, functions that are abundant in forest soils (Mendes *et al.* 2015b). Chlamydiae has been associated with degradation if plant biomass (Kanokratana *et al.* 2010), which explain their high abundance in forest soils. On the other hand, several groups showed an increased abundance after the conversion of the forest to agriculture, such as Verrucomicrobia, Planctomycetes, Bacteroidetes, Chloroflexi and Nitrospirae. Chloroflexi are aerobic thermophiles, which grow in high temperatures, a common characteristic of pasture soils (Rodrigues *et al.* 2013). The Bacteroidetes phylum includes

plant-growth-promoting bacteria (Soltani *et al.* 2010) and cellulose decomposing bacteria (Verkhovtseva, Kubarev and Mineev 2007), which might be related to cultivation in agricultural areas. The high abundance of Nitrospirae may be related to the inorganic fertilization in the agricultural areas, since members of this group are related to nitrogen transformations, such as nitrite oxidation (Attard *et al.* 2010). Also, the functional potential profile was distinct among the land use sites revealing a differential response of specific microbial groups to soil chemical alterations. Mendes *et al.* (2015b) showed that, after conversion of native forest to agricultural areas, the functional potential profiles of the communities are changed, where the altered sites present high abundance of potential functions related to soil stress as a response of deforestation and soil management. In addition, the co-occurrence network analysis reinforced the effect of land use change on the microbial community dynamics, where grassland presented a more complex network compared with forest and no-till samples, revealing the distinct response of this community to soil management.

Ecological processes governing microbial assembly and functional potential

The aim of this study was to assess the effects of land use change on the structure of the soil microbial community, examining the ecological processes that contribute to community assembly in a long-term forest-to-agriculture conversion. In order to test the effects of land use change on community assembly we analyzed our data based on ecological process. Our results strongly support the hypothesis that grassland and no-till microbial communities were assembled differently. Our results suggest that forest and no-till microbial communities have assemblages based on niche-based (deterministic) processes and, thus, are more prone to be influenced by selective pressures imposed by the surrounding environment, indicating variable selection and niche specialization. Differently, grassland communities have a neutral (stochastic) assembly, a pattern that emphasizes random changes in species relative abundance (Dini-Andreote *et al.* 2015). While diversification and dispersal apparently acted on the composition of microbial communities, selection and drift clearly altered the relative abundance of their members. This is consistent with previous observations that the above processes shifted their strength at temporal and spatial scales (Nemergut *et al.* 2013; Stegen *et al.* 2013). Furthermore, we determined that the long-term forest to grassland conversion reduced diversification. This was true not only for the beta diversity, a measure of shared taxa between samples in space and/or time, but also for structure and distance to the nearest neighbor, both measures of community similarity. Together, these results indicate that grassland communities have a more homogeneous composition. When the rate of dispersal was calculated for each community, forest and no-till samples had average values of 0.08 and 0.07, respectively, suggesting their microbial populations had a very low tendency to migrate, corroborating the results of previous microbial assembly studies (Gravel *et al.* 2006; Dini-Andreote *et al.* 2015). Conversely, grassland samples showed consistently higher dispersal rates (average = 0.38). A combination of decreased diversification and increased dispersal, as observed in grassland, often leads to biotic homogenization (Loreau, Mouquet and Holt 2003; Olden 2006; Rodrigues *et al.* 2013). Also, our results showed that strong selection is an important driver of community assembly in grassland, which leads to a community homogenization. Owing to the difficulty of directly measuring the contribution of drift for non-discrete populations in large communities, the process was estimated as a component of the partitioning of redundancy analysis residue—non-explained variation + stochastic forces (Vellend 2010; Stegen *et al.* 2013). When selection is weak (low R^2) and residual variation is high, such as observed in our study for no-till and forest, drift is likely to play an important role in community assembly (Nemergut *et al.* 2013; Knelman and Nemergut 2014).

Community dynamics after long-term land use conversion

Together, the above results suggest that land use is a more important driver for microbial assembly than the conversion itself. While the no-till system in our study is a successive rotation of soybean and/or maize in the summer, followed by wheat in the winter, the grassland is dominated by the perennial species *Axonopous affinis*. In grassland soil, the grass root system increases the soil connectivity by removing potential barriers to migration (e.g. aggregate-to-aggregate continuity) and leads to higher microbial dispersal in space and time (Murphy and Foster 2014). As a consequence, the overall microbial diversity is impacted. However, the no-till cropping with rotation increases carbon storage by maintaining residues from previous cultivations and supplying microbial communities with more heterogeneous sources, indicating variable selection. Classical ecological theory states that species coexistence increases as a result of niche partitioning due to resource availability (Carroll, Cardinale and Nisbet 2011) and empirical support has been provided for plant and animal communities (Finke and Snyder 2008; Mori *et al.* 2013). We found that resource heterogeneity in no-till increased niche specialization and reduced the importance of the process of homogenizing selection and dispersal on microbial assembly.

Niche and neutral processes have been studied in terrestrial ecosystems (Stegen *et al.* 2012; Ferrenberg *et al.* 2013; Dini-Andreote *et al.* 2015), but the relationship between assembly and functions has yet to be investigated (Knelman and Nemergut 2014). Given that our results point towards distinctive assemblies according to land use, we argue as follows: (i) if community assembly occurs mostly by neutral processes, a few environmental conditions should exert a stronger effect on microbial functional potential, leading to homogenizing selection; and (ii) if community assembly is mostly determined by niche-based processes, the heterogeneity of the environment should modulate functional potential, which means variable selection and niche complementarity (Blüthgen and Klein 2011). While the grassland communities were found to be neutral, forest and no-till communities had a predominantly niche-based assembly. We attribute this observation in no-till to functional niche partitioning, in which species can coexist by differential resource utilization. When resources are more heterogeneous, lower rates of dispersal and variable selection lead to a decrease in competition, and thus functional potentials are maintained through niche specialization (Blüthgen and Klein 2011). A second implication of niche partitioning in no-till is the buffering effect against species losses (Bihn, Gebauer and Brandl 2010; Mouillot *et al.* 2013), as observed by similar values of global beta diversity when compared with those found in forest. In grassland, root exudates are mostly composed of easily degradable and homogenous C sources that might intensify the process of specific selection, thus leading to an increase in functional potential categories related to carbohydrates. As a response to the pressure of a few and strong environmental variables, microbial communities increase functional potential categories

associated with stress responses (regulation and cell signaling, respiration, and virulence, disease and defense) and motility and chemotaxis (Fig. 4b), the latter associated with the moderate tendency to migration.

Forest conversion to agricultural systems has been generally regarded as detrimental to biodiversity. Most studies describe community patterns without evaluating processes governing their assembly or consequences for function. We have shown that soil microbial community assembly is dependent on land use and influenced by the net biodiversity outcome. Furthermore, the relative weight of divergent ecological processes, depending on land use, determined a niche- or neutral-based model of assembly with implications for function. Because these models are not mutually exclusive, future investigations stand to gain valuable information from assessing their combined effect in different ecosystems and across multiple scales of space and time. Understanding microbial community assembly and characterizing fundamental processes that guide dynamic responses in community organization have the potential to provide important insights into ecosystem modeling of biodiversity and microbial conservation.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSEC* online.

AUTHOR CONTRIBUTIONS

All authors discussed the results and commented on the manuscript. DG-S, SMT and DB designed the project. DG-S, CDB and DB collected the soil samples. DG-S, CDB and LWM performed the molecular biology analyses. DG-S, JLMR and LWM analyzed the metadata. DG-S, LWM, JLMR and SMT wrote the manuscript.

ACKNOWLEDGEMENT

We thank Adam Burns (University of Oregon), Lucas Silva (University of California—Davis), and Stephanie Eichorst (University of Vienna) for thoughtful feedback and discussions on neutral and niche models.NDING

This study was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo (No. 2008/58114-3 and No. 2014/50320-4) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (SISBIOTA-CNPq No. 563251/2010-7). DG-S received a scholarship from Coordenacão de Aperfeiçoamento de Pessoal de Nível Superior (DS-CAPES PPGIEA/USP, PDSE-CAPES No. 99999.003968/2014-06) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (PRONEX-CNPq No. 140317/2014-7). SMT thanks CNPq (CNPq-PQ 311008/2016-0).

Conflict of interest. None declared.

REFERENCES

Alam MK, Salahin N, Rashid MH *et al.* Effects of different tillage practices and cropping patterns on soil physical properties and crop productivity. *J Trop Resour Sustain Sci* 2013;**1**:51–61.

Alonso D, McKane AJ. Sampling Hubbell's neutral theory of biodiversity. *Ecol Lett* 2004;**7**:901–10.

Anderson M. A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 2001;**26**:32–46.

Attard E, Poly F, Commeaux C *et al.* Shifts between *Nitrospira*- and *Nitrobacter*-like nitrite oxidizers underlie the response of soil potential nitrite oxidation to changes in tillage practices. *Environ Microbiol* 2010;**12**:315–26.

Baker KL, Langenheder S, Nicol GW *et al.* Environmental and spatial characterisation of bacterial community composition in soil to inform sampling strategies. *Soil Biol Biochem* 2009;**41**:2292–8.

Barberán A, Bates ST, Casamayor EO *et al.* Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J* 2012;**6**:343–51.

Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. *Proceedings of the Third International ICWSM Conference*. AAAI Publications, 2009.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc* 1995;**57**:289–300.

Bennett LT, Adams MA. Response of a perennial grassland to nitrogen and phosphorus additions in sub-tropical, semi-arid Australia. *J Arid Environ* 2001;**48**:289–308.

Bihn JH, Gebauer G, Brandl R. Loss of functional diversity of ant assemblages in secondary tropical forests. *Ecology* 2010;**91**:782–92.

Blanchet FG, Legendre P, Borcard D. Forward selection of explanatory variables. *Ecology* 2008;**89**:2623–32.

Blüthgen N, Klein AM. Functional complementarity and specialisation: The role of biodiversity in plant-pollinator interactions. *Basic Appl Ecol* 2011;**12**:282–91.

Brody JR, Kern SE. Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis. *Biotechniques* 2004;**36**:214–6.

Carbonetto B, Rascovan N, Álvarez R *et al.* Structure, composition and metagenomic profile of soil microbiomes associated to agricultural land use and tillage systems in Argentine Pampas. *PLoS One* 2014;**9**:e99949.

Cardinale BJ, Duffy JE, Gonzalez A *et al.* Biodiversity loss and its impact on humanity. *Nature* 2012;**486**:59–67.

Carroll IT, Cardinale BJ, Nisbet RM. Niche and fitness differences relate the maintenance of diversity to ecosystem function. *Ecology* 2011;**92**:1157–65.

Cenciani K, Lambais MR, Cerri CC *et al.* Bacteria diversity and microbial biomass in forest, pasture and fallow soils in the southwestern Amazon basin. *Rev Bras Cienc Solo* 2009;**33**:907–16.

Chapin FS, Zavaleta ES, Eviner VT *et al.* Consequences of changing biodiversity. *Nature* 2000;**405**:234–42.

Claessen MEC, Barreto WO, Paula JL *et al.* *Manual de métodos de análise de solo, 2nd edn*. Rio de Janeiro: Embrapa-CNPS, 1997.

Dhaliwal GS, Gupta N, Kukal SS *et al.* Standardization of automated Vario EL III CHNS analyzer for total carbon and nitrogen determination in soils. *Commun Soil Sci Plant Anal* 2011;**42**:971–9.

De Vries FT, Liiri ME, Bjørnlund L *et al.* Land use alters the resistance and resilience of soil food webs to drought. *Nat Clim Chang* 2012;**2**:276–80.

Dini-Andreote F, Stegen JC, van Elsas JD *et al.* Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc Natl Acad Sci U S A* 2015;**112**:1326–32.

Drenovsky RE, Steenwerth KL, Jackson LE *et al*. Land use and climatic factors structure regional patterns in soil microbial communities. *Glob Ecol Biogeogr* 2010;**19**:27–39.

Etienne RS, Alonso D. A dispersal-limited sampling theory for species and alleles. *Ecol Lett* 2005;**8**:1147–56.

Feeley KJ, Silman MR. Extinction risks of Amazonian plant species. *Proc Natl Acad Sci U S A* 2009;**106**:12382–7.

Ferrenberg S, O'Neill SP, Knelman JE *et al*. Changes in assembly processes in soil bacterial communities following a wildfire disturbance. *ISME J* 2013;**7**:1102–11.

Fierer N, Lauber CL, Ramirez KS *et al*. Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J* 2012;**6**:1007–17.

Fierer N, Lennon JT. The generation and maintenance of diversity in microbial communities. *Am J Bot* 2011;**98**:439–48.

Figuerola ELM, Guerrero LD, Rosa SM *et al*. Bacterial indicator of agricultural management for soil under no-till crop production. *PLoS One* 2012;**7**:e51075.

Finke DL, Snyder WE. Niche partitioning increases resource exploitation by diverse communities. *Science* 2008;**321**:1488–90.

Foley JA, DeFries R, Asner GP *et al*. Global consequences of land use. *Science* 2005;**309**:570–4.

Foley JA, Ramankutty N, Brauman KA *et al*. Solutions for a cultivated planet. *Nature* 2011;**478**:337–42.

Gee GW, Bauder JW. Particle-size analysis. In: Klute, A (ed). *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods*. Madison: American Society of Agronomy, 1986, 383–411.

Gravel D, Canham CD, Beaudet M *et al*. Reconciling niche and neutrality: the continuum hypothesis. *Ecol Lett* 2006;**9**:399–409.

Hammer Ø, Harper DAT, Ryan PD. PAST: Paleontological statistics software package for education and data analysis. *Paleontol Electron* 2001;**4**:1–9.

Hubbell SP. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton: Princeton University Press, 2001.

Jabot F, Etienne RS, Chave J. Reconciling neutral community models and environmental filtering: theory and an empirical test. *Oikos* 2008;**117**:1308–20.

Jesus EC, Marsh TL, Tiedje JM *et al*. Changes in land use alter the structure of bacterial communities in Western Amazon soils. *ISME J* 2009;**3**:1004–11.

Kanokratana P, Uengwetwanit T, Rattanachomsri U *et al*. Insights into the phylogeny and metabolic potential of a primary tropical peat swamp forest microbial community by metagenomic analysis. *Microb Ecol* 2010;**61**:518–28.

Keeney DR, Nelson DW Page AL, Miller RH, Keeney DR (eds). *Methods of Soil Analysis, Part 2, Chemical and Microbiological Properties*. Madison: American Society of Agronomy, 1982, 643–98.

Knelman JE, Nemergut DR. Changes in community assembly may shift the relationship between biodiversity and ecosystem function. *Front Microbiol* 2014;**5**:1–4.

Kuramae EE, Yergeau E, Wong LC *et al*. Soil characteristics more strongly influence soil bacterial communities than land-use type. *FEMS Microbiol Ecol* 2012;**79**:12–24.

Lauber CL, Ramirez KS, Aanderud Z *et al*. Temporal variability in soil microbial communities across land-use types. *ISME J* 2013;**7**:1641–50.

Lee-Cruz L, Edwards DP, Tripathi BM *et al*. Impact of logging and forest conversion to oil palm plantations on soil bacterial communities in Borneo. *Appl Environ Microbiol* 2013;**79**:7290–7.

Leibold MA, Holyoak M, Mouquet N *et al*. The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* 2004;**7**:601–13.

Lepš J, Šmilauer P. Multivariate analysis of ecological data using CANOCO. *Bull Ecol Soc Am* 2005;**86**:6–6.

Lienhard P, Tivet F, Chabanne A *et al*. No-till and cover crops shift soil microbial abundance and diversity in Laos tropical grasslands. *Agron Sustain Dev* 2013;**33**:375–84.

Loreau M, Mouquet N, Holt RD. Meta-ecosystems: a theoretical framework for a spatial ecosystem ecology. *Ecol Lett* 2003;**6**:673–9.

MacArthur RH, Wilson EO. An equilibrium theory of insular zoogeogrphy. *Evolution* 1963;**17**:373–87.

Magoč T, Salzberg SL. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011;**27**:2957–63.

McConkey KR, Farrill GO. Cryptic function loss in animal populations. *Trends Ecol Evol* 2015;**30**:182–9.

Mehnaz KR, Dijkstra FA. Denitrification and associated N_2O emissions are limited by phosphorus availability in a grassland soil. *Geoderma* 2016;**284**:34–41.

Mendes LW, Brossi MJL, Kuramae EE *et al*. Land-use system shapes soil bacterial communities in Southeastern Amazon region. *Appl Soil Ecol* 2015a;**95**:151–60.

Mendes LW, Kuramae EE, Navarrete AA *et al*. Taxonomical and functional microbial community selection in soybean rhizosphere. *ISME J* 2014;**8**:1577–87.

Mendes LW, Tsai SM, Navarrete AA *et al*. Soil-borne microbiome: linking diversity to function. *Microb Ecol* 2015b;**70**:255–65.

Meyer F, Paarmann D, D'Souza M *et al*. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 2008;**9**:1–8.

Miyashita NT, Iwanaga H, Charles S *et al*. Soil bacterial community structure in five tropical forests in Malaysia and one temperate forest in Japan revealed by pyrosequencing analyses of 16S rRNA gene sequence variation. *Genes Genet Syst* 2013;**88**:93–103.

Moraes JFL, Neill C, Volkoff B *et al*. Soil carbon and nitrogen stocks following forest conversion to pasture in the Western Brazilian Amazon Basin. *Acta Scientiarum* 2002;**24**:1369–76.

Mori AS, Shiono T, Koide D *et al*. Community assembly processes shape an altitudinal gradient of forest biodiversity. *Glob Ecol Biogeogr* 2013;**22**:878–88.

Mouillot D, Bellwood DR, Baraloto C *et al*. Rare species support vulnerable functions in high-diversity ecosystems. *PLoS Biol* 2013;**11**:e1001569.

Mueller RC, Paula FS, Mirza BS *et al*. Links between plant and fungal communities across a deforestation chronosequence in the Amazon rainforest. *ISME J* 2014;**8**:1548–50.

Murphy CA, Foster BL. Soil properties and spatial processes influence bacterial metacommunities within a grassland restoration experiment. *Restor Ecol* 2014;**22**:685–91.

Myers N, Mittermeier RA, Mittermeier CG *et al*. Biodiversity hotspots for conservation priorities. *Nature* 2000;**403**:853–558.

Navarrete AA, Kuramae EE, de Hollander M *et al*. Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. *FEMS Microbiol Ecol* 2013;**83**:607–21.

Navarrete AA, Taketani RG, Mendes LW *et al*. Land-use systems affect archaeal community structure and functional diversity in Western Amazon soils. *Rev Brasil Cien Solo* 2011;**35**:1527–40.

Nemergut DR, Schmidt SK, Fukami T *et al*. Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* 2013;**77**:342–56.

Newcombe RG. Improved confidence intervals for the difference between binomial proportions based on paired data. *Stat Med* 1998;**17**:2635–50.

Nüsslein K, Tiedje JM. Soil bacterial community shift correlated with change from forest to pasture vegetation in a tropical soil. *Appl Environ Microbiol* 1999;**65**:3622–6.

Ofiteru ID, Lunn M, Curtis TP *et al*. Combined niche and neutral effects in a microbial wastewater treatment community. *Proc Natl Acad Sci U S A* 2010;**107**:15345–50.

Olden JD. Biotic homogenization: a new research agenda for conservation biogeography. *J Biogeogr* 2006;**33**:2027–39.

Ollivier J, Töwe S, Bannert A *et al*. Nitrogen turnover in soil and global change. *FEMS Microbiol Ecol* 2011;**78**:3–16.

Overbeek R, Olson R, Pusch GD *et al*. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 2013;**42**:1–9.

Parks DH, Beiko RG. Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 2010;**26**:715–21.

Pereira HM, Leadley PW, Proença V *et al*. Scenarios for global biodiversity in the 21st century. *Science* 2010;**330**:1496–501.

Pholchan MK, Baptista JC, Davenport RJ *et al*. Microbial community assembly, theory and rare functions. *Front Microbiol* 2013;**4**:1–9.

Piccolo MC, Neill C, Cerri CC. Natural abundance of ^{15}N in soils along forest-to-pasture chronosequences in the western Brazilian Amazon Basin. *Oecologia* 1994;**99**:112–7.

Ramette A, Tiedje JM. Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. *Proc Natl Acad Sci U S A* 2007;**104**:2761–6.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation For Statistical Computing, 2007. http://www.R-project.org.

Ribeiro MC, Metzger JP, Martensen AC *et al*. The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biol Conserv* 2009;**142**:1141–53.

Rodrigues JLM, Pellizari VH, Mueller R *et al*. Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. *Proc Natl Acad Sci U S A* 2013;**110**;988–93.

Sala OE, Chapin FS, Armesto JJ *et al*. Global biodiversity scenarios for the year 2100. *Science* 2000;**287**:1770–4.

Soltani AA, Khavazi K, Asadi-Rahmani H *et al*. Plant growth promoting characteristics in some *Flavobacterium* spp. isolated from soils of Iran. *J Agric Sci* 2010;**2**:106–15.

Stegen JC, Lin X, Fredrickson JK *et al*. Quantifying community assembly processes and identifying features that impose them. *ISME J* 2013;**7**:2069–79.

Stegen JC, Lin X, Konopka AE *et al*. Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* 2012;**6**:1653–64.

Tedesco MJ, Gianello C, Bissani CA *et al*. *Analysis of Soil, Plants and Other Materials*. Porto Alegre: Universidade Federal do Rio Grande do Sul, 1995.

Thomson AM, Calvin KV, Chini LP *et al*. Climate mitigation and the future of tropical landscapes. *Proc Natl Acad Sci U S A* 2010;**107**;19633–8.

Tringe SG, von Mering C, Kobayashi A *et al*. Comparative metagenomics of microbial communities. *Science* 2005;**308**:554–7.

Vellend M. Conceptual synthesis in community ecology. *Q Rev Biol* 2010;**85**:183–206.

Verkhovtseva N, Kubarev E, Mineev V. Agrochemical agents in maintaining the structure of the soil microbial community. *Russ Agric Sci* 2007;**33**:100–2.

Volkov I, Banavar JR, Hubbell SP *et al*. Neutral theory and relative species abundance in ecology. *Nature* 2003;**424**:1035–7.

Wang J, Shen J, Wu Y *et al*. Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J* 2013;**7**:1310–21.

Wardle DA, Bardgett RD, Callaway RM *et al*. Terrestrial ecosystem responses to species gains and losses. *Science* 2011;**332**:1273–8.

Welch BL. The generalization of Student's problem when several different population variances are involved. *Biometrika* 1934;**34**:28–35.

Wilke A, Harrison T, Wilkening J *et al*. The M5nr: a novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. *BMC Bioinformatics* 2012;**13**:1–5.