FEMS Microbiology Ecology, 91, 2015, fiv031

doi: 10.1093/femsec/fiv031
Advance Access Publication Date: 20 March 2015
Research Article

RESEARCH ARTICLE

Phylogenetic structure of soil bacterial communities predicts ecosystem functioning

Eduardo Pérez-Valera, Marta Goberna and Miguel Verdú*

Centro de Investigaciones sobre Desertificación (CSIC-UVEG-GV), Carretera Moncada - Náquera, Km 4.5, E-46113 Moncada, Valencia, Spain

*Corresponding author: Centro de Investigaciones sobre Desertificación (CSIC-UVEG-GV), Carretera Moncada - Náquera, Km 4.5, E-46113 Moncada, Valencia, Spain. Tel: +34-96-342-41-62; Fax +34-96-342-41-60; E-mail: miguel.verdu@uv.es
One sentence summary: The relationship between phylogenetic diversity and ecosystem functioning is dependent on the taxonomic identity of the main coexisting bacterial lineages.
Editor: Angela Sessitsch

ABSTRACT

Quantifying diversity with phylogeny-informed metrics helps understand the effects of diversity on ecosystem functioning (EF). The sign of these effects remains controversial because phylogenetic diversity and taxonomic identity may interactively influence EF. Positive relationships, traditionally attributed to complementarity effects, seem unimportant in natural soil bacterial communities. Negative relationships could be attributed to fitness differences leading to the overrepresentation of few productive clades, a mechanism recently invoked to assemble soil bacteria communities. We tested in two ecosystems contrasting in terms of environmental heterogeneity whether two metrics of phylogenetic community structure, a simpler measure of phylogenetic diversity (NRI) and a more complex metric incorporating taxonomic identity (PCPS), correctly predict microbially mediated EF. We show that the relationship between phylogenetic diversity and EF depends on the taxonomic identity of the main coexisting lineages. Phylogenetic diversity was negatively related to EF in soils where a marked fertility gradient exists and a single and productive clade (*Proteobacteria*) outcompete other clades in the most fertile plots. However, phylogenetic diversity was unrelated to EF in soils where the fertility gradient is less marked and *Proteobacteria* coexist with other abundant lineages. Including the taxonomic identity of bacterial lineages in metrics of phylogenetic community structure allows the prediction of EF in both ecosystems.

Keywords: competitive exclusion; fitness differences; phylogenetic diversity; phylogenetic clustering; *Proteobacteria*; taxonomic identity

INTRODUCTION

The effect of biodiversity on ecosystem functioning has been widely studied, numerous pieces of evidence indicating a positive effect but some also reporting neutral or negative relationships (Zak *et al.* 2003; Hooper *et al.* 2005; Balvanera *et al.* 2006; Cardinale *et al.* 2012). Soil bacteria are primary actors in this relationship because of their exceptional diversity and key role on ecosystem functioning, through decomposing organic matter and controlling the planetary flows of energy and nutrients (Curtis, Sloan and Scannell 2002; Wardle *et al.* 2004; Van der Heijden, Bardgett and van Straalen 2008).

Species richness has been the measure of biodiversity traditionally used in studies relating biodiversity and ecosystem functioning (Cardinale *et al.* 2012). However, this approach disregards the fact that functional similarities among species are usually determined by their common evolutionary history, and therefore, phylogenetically related species tend to perform similar functions (Blomberg, Garland and Ives 2003; Martiny, Treseder and Pusch 2013; but see Revell, Harmon and Collar 2008

Received: 6 February 2015; **Accepted:** 14 March 2015

for other processes producing trait resemblance among close relatives). This is the reason why phylogenetically informed measures of diversity tend to be more informative than traditional richness measures (Lozupone and Knight 2007; Cadotte, Cardinale and Oakley 2008). Empirical evidence on the effect of phylogenetic diversity on ecosystem functioning is widespread across the tree of life (e.g. bacteria, Gravel *et al.* 2012; fungi, Maherali and Klironomos 2007; plants, Cadotte, Cardinale and Oakley 2008; Cadotte 2013; Navarro-Cano *et al.* 2014). Most of these studies have found a positive relationship between phylogenetic diversity and ecosystem functioning parameters, as expected when distantly related taxa perform complementary functions. However, neutral and negative relationships have also been described, particularly in bacteria, because phylogenetic diversity and taxonomic diversity may interactively influence ecosystem functioning (Severin, Östman and Lindström 2013; Venail and Vives 2013).

Phylogenetic diversity of bacterial communities in soils is low compared to those in other natural environments, contrasting with their extremely high species richness and diversity (Lozupone and Knight 2007). This paradoxical situation could be explained by adding a phylogenetic context to the modern coexistence theory (Chesson 2000; Mayfield and Levine 2010; HilleRis-Lambers *et al.* 2012; Godoy, Kraft and Levine 2014). The phylogenetic structure of soil bacterial communities is primarily driven by abiotic factors, such as acidity (Jones *et al.* 2009) and availability of organic resources (Goberna, García and Verdú 2014), that overrepresent certain clades. The composition of ecological communities is further determined by the balance between mechanisms shaping niche differences and fitness differences between lineages (Chesson 2000). Coexistence is maximized under large niche differences (i.e. absence of niche overlap), a situation where species do not compete for resources. This increases both diversity and productivity since the functional complementarity of coexisting organisms allows a more complete usage of resources. Complementarity effects have been shown to underlie the positive relationship between bacterial diversity and productivity in simple experimental communities, but it seems to be relatively unimportant in natural communities due to the high functional redundancy of bacteria (Griffiths *et al.* 2001; Bell *et al.* 2005; Venail and Vives 2013). In contrast, fitness differences between lineages tend to favor competitive exclusion because competitively superior lineages may consume too much of the resource on which other lineages depend (Chesson 2000). Mayfield and Levine (2010) noticed that fitness differences may produce outcompetition of entire clades when competitive superiority is a phylogenetically conserved trait. The immediate consequence of competitive exclusion of entire clades is the reduction of phylogenetic diversity in ecological communities, as occurs in soil bacterial communities worldwide (Goberna, García and Verdú 2014).

Fitness differences may be produced by competitive asymmetries in which some lineages produce more per unit resource than others (Chesson 2000). This is the case of *Proteobacteria* and *Actinobacteria*, two bacterial lineages which are extremely competitive in terms of growth response when organic carbon substrates of varying recalcitrance are supplied to the soil, which is typically carbon limited (Goldfarb *et al.* 2011). This competitive superiority is phylogenetically conserved and therefore competitive exclusion leads to the overrepresentation of a few, very productive, lineages resulting in phylogenetic clustering both in experimental and natural soil communities (Goldfarb *et al.* 2011; Goberna *et al.* 2014). Under this scenario, highly productive communities dominated by competitive clades would feature low phylogenetic diversity levels, leading to an inverse relationship between phylodiversity and ecosystem functioning.

Here, we selected two ecosystems contrasting in terms of environmental heterogeneity, which is a main determinant of bacterial diversity (Ramette and Tiedje 2007). Differences between sites were particularly marked as regards the heterogeneity of resource availability, a factor that modifies the relationship between bacterial diversity and productivity (Jousset *et al.* 2011). In both ecosystems, we test whether (i) soil physical and chemical parameters determine the phylogenetic structure and (ii) the phylogenetic structure of bacterial communities predicts ecosystem functioning, measured through soil microbial productivity, metabolic efficiency and nutrient cycling increases, via overrepresentation of a particular productive clade.

MATERIALS AND METHODS

Study site

The study was carried out in two Mediterranean sites, differing in their climate, plant cover, lithology and soil type. We intentionally searched these contrasting ecosystems to test whether the phylogenetic structure of soil bacterial communities predicts ecosystem function under two extremes of environmental heterogeneity. Site 1 is characterized by the presence of a dense shrubland (100% plant cover) dominated by *Rosmarinus officinalis* L. and located in Teresa de Cofrentes (Valencia, Spain). Soils are Haplic Leptosols (Calcaric, Humic) (FAO–ISRIC–IUSS 2006) developed on limestones, mean annual rainfall is 446 mm and temperature 13.7°C. Topsoils (0–2 cm) were collected in ten 1 $\times$ 1 m plots located within a 150 m^2 area as described in Goberna *et al.* (2012). Site 2 is covered by a patchy shrub steppe dominated by *Ononis tridentata* L. and located in Algepsar dels Burutaus (Serra de Crevillent, Alacant, SE Spain). Soils are Leptic Regosols (Gypsiric, Calcaric) (FAO–ISRIC–IUSS 2006) developed on gypsum, mean annual rainfall is 220 mm and temperature 20°C. Topsoils (0–2 cm) were collected underneath 15 vegetation patches (defined as groups of plants growing underneath the canopy of an *O. tridentata* individual) and in the adjacent open spaces, all plots being located within a 1-ha area as described by Navarro-Cano *et al.* (2014). Sites 1 and 2, representing two extremes of environmental heterogeneity will be hereafter referred to as 'non-patchy' and 'patchy' ecosystems, respectively, based on the structure of their plant communities.

Plant community structure determined a low variance in the soil physical and chemical properties in the non-patchy ecosystem, which contrasted with the high variability of the same variables in the patchy ecosystem (Table 1). Further details on the soil physical and chemical environment in both sites can be found in previous studies (Goberna *et al.* 2012; Navarro-Cano *et al.* 2014). We characterized the soils of each plot with the scores of the first principal component (PC1-Soil) including the soil gravimetric humidity (GH), pH, electrical conductivity (EC), total organic C (TOC), pyrophosphate oxidizable C (PPi-OC) and total nitrogen (TN). PC1-Soil was then used as an abiotic predictor of phylogenetic structure of soil bacterial communities as described below. Both sites also exhibited large differences in the variability of several biochemical properties that are commonly used as proxies of ecosystem functioning, with the non-patchy ecosystem showing lower coefficients of variation compared to the patchy ecosystem (Table 1). Specifically, we used parameters that are indicators of general microbial activity and specific enzymatic activities involved in main steps of the nutrient cycles (Nannipieri, Grego and Ceccanti 1990). In particular, general

Table 1. Variability among sampling plots in physical, chemical and biochemical variables in the non-patchy and patchy ecosystems (data published by Goberna *et al.* 2012 and Navarro-Cano *et al.* 2014).

	Non-patchy ecosystem			Patchy ecosystem		
Variable	Mean	SD	CV	Mean	SD	CV
Gravimetric humidity (%)	23.34	5.27	22.57%	2.94	1.48	50.36%
Total organic C (g kg^{-1})	43.9	5.2	11.95%	59.7	39.7	66.53%
Pyrophosphate oxidizable carbon (g kg^{-1})	15.4	6.2	40.47%	1.87	1.57	83.65%
pH	8.05	0.17	2.14%	7.18	0.16	2.26%
Electrical conductivity (μS cm^{-1})	230	35.66	15.48%	2798	334	11.95%
Total N (%)	0.39	0.13	32.32%	0.39	0.29	73.78%
MBC (mg C kg^{-1})	469	194	41.32%	1411	1307	92.68%
MBC/TOC (%)	1.07	0.42	38.71%	1.89	1.11	58.95%
Basal respiration (mg C-CO_2 kg^{-1} d^{-1})	15.75	6.27	39.78%	76.26	71.88	94.25%
qCO2 (μg C-CO_2 mg^{-1} MBC h^{-1})	1.91	2.19	114.57%	2.18	1.08	49.47%
ATP (ng g^{-1})	2186	493	22.53%	424	310	73.11%
β-Glucosidase activity (μmol PNP g^{-1} h^{-1})	2.93	0.81	27.51%	5.88	6.07	103.28%
Phosphatase activity (μmol PNP g^{-1} h^{-1})	15.73	6.13	38.99%	16.11	14.78	91.73%
Urease activity (mg N-NH_4^+ g^{-1} h^{-1})	1.07	0.36	33.52%	2.05	1.49	72.59%

indicators of microbial activity included: (1) microbial biomass C (MBC) as a proxy of the microbial biomass; (2) ATP content, as an indicator of the total microbial activity; (3) basal respiration (BR), as an indicator of the activity of decomposers that mineralize organic C into CO_2; (4) microbial coefficient (MBC/TOC), which reflects the conversion efficiency of organic C into microbial C; and (5) metabolic quotient (qCO_2), which is the ratio between CO_2-C production and MBC and declines as the microbiota becomes efficient at conserving C. Specific indicators of microbial activity included (1) β-glucosidase (GA), (2) alkaline phosphatase (PA) and (3) urease activities (UA), which are hydrolytic enzymes that are respectively involved in C, P and N cycling. Further details on the soil biochemical properties in both sites can be found in previous studies (Goberna *et al.* 2012; Navarro-Cano *et al.* 2014).

Soil DNA extraction and pyrosequencing

Soil DNA from the non-patchy ecosystem was extracted within 24 h after sampling from ca. 0.25 g soil with the PowerSoil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA). Extracted DNA was checked for quality by electrophoresis in 1% agarose gels run in 0.5 × Tris–acetate–EDTA buffer. Amplifications of the 16S rRNA gene were carried out using the universal bacterial primers 8F (5′-AGAGTTTGATCCTGGCTCAG-3′; Turner *et al.* 1999) and 534R (5′-ATTACCGCGGCTGCTGGC-3′; Muyzer, de Waal and Uitterlinden 1993). Each sample contained a synthetized forward primer, including a 454 sequencing adaptor (5′-CCATCTCATCCCTGCGTGTCTCCGACTCAG-3′) and a unique 8-nucleotide barcode in its 5′ end randomly selected from those published by Hamady *et al.* (2008). The reverse primer incorporated a 454 sequencing adaptor in its 5′ end (5′-CCTATCCCCTGTGTGCCTTGGCAGTCTCAG-3′).

PCR reactions were performed in a Flexcycler (Analytik Jena, Jena, Germany) in 50 μl volumes. Each reaction contained a final concentration of 1 × Platinum PCR SuperMix High Fidelity (Invitrogen, Carlsbad, USA), 0.3 μM of each primer and 0.4 mg mL^{-1} bovine serum albumin. A volume of 1.5 μl DNA was directly applied to the reaction mix. Thermal cycling consisted of 5 min at 94°C, 20 cycles including 45 s at 94°C, 45 s at 54°C and 90 s at 72°C and terminated with 10 min at 72°C. Purification of PCR products (100 μl) was carried out with the NucleoSpin Extract II Kit (Macherey-Nagel, Düren, Germany). Afterwards, they were eluted in 50 μl DNAase free 1 × TE (Tris-EDTA) buffer and checked for quality and size in 2% agarose gels run in 1 × TAE buffer (80 V, 45 min). Non-template controls followed the same procedure. Purified tagged amplicons were quantified in duplicate using the Quant-iT PicoGreen dsDNA Kit (Invitrogen, Carlsbad, USA) and pooled in equimolar amounts. Pyrosequencing was performed by GATC Biotech (Konstanz, Germany) with the Roche 454 GS-FLX system using titanium chemistry.

Similar procedures were used for DNA extraction, PCR amplification and pyrosequencing of soil samples in 30 plots from the patchy ecosystem. Details are given in Goberna *et al.* (2014).

Sequence analysis and phylogeny reconstruction

For the non-patchy ecosystem, 10 604 sequences were obtained. Short sequences (<200 bp) were removed, along with those with ambiguous base calls or with homopolymers exceeding 6 bp. Primers and barcodes were trimmed. After denoising, chimeric sequences and singletons were excluded from the analysis. Operational taxonomic units (OTUs) were defined at an identity level of 97% and taxonomically classified using BLASTn against a curated GreenGenes database (DeSantis *et al.* 2006). This initial sequence processing was performed by MR DNA (Shallowater, TX, USA). A final 2289 OTUs were aligned with PyNAST (Caporaso *et al.* 2010a) by using QIIME (Caporaso *et al.* 2010b). Then, we constructed a community matrix showing the abundance of the total 2289 OTUs in each of the 10 plots. As proposed by Kembel *et al.* (2012), the relative abundance of each OTU was corrected by the estimated number of 16S rRNA gene copies. Bacterial phylogeny was reconstructed using RAxML 7.3.0 (Stamatakis 2006). We built five independent maximum-likelihood phylogenetic trees with the GTRGAMMA substitution model. Previously, hypervariable regions were removed using the Lane mask (Lane 1991). To avoid high phylogenetic uncertainty resulting from the usage of short sequences, tree topology was constrained to match the basal relationships of the megatree built from the Silva database (Release 108, Quast *et al.* 2013). All phylogenetic trees were selected among the best of 1000 iterations and rooted using *Archaeoglobus profundus*.

Sequences were deposited in the European Nucleotide Archive (http://www.ebi.ac.uk/ena/data/view/PRJEB6166).

In the patchy ecosystem, we worked with 24 162 sequences after removal of low-quality sequences and artifacts. After excluding singletons, these were collapsed into a final 3290 OTUs. Sequence processing and phylogeny reconstruction were similar to those described above and details are given in Goberna *et al.* (2014).

Phylogenetic community structure

We described the phylogenetic structure of bacterial communities by using two phylogeny-weighted metrics. First, we calculated the abundance-weighted net relatedness index (NRI), one of the most commonly used metrics in community phylogenetics, with the picante package for R (Kembel *et al.* 2010). This computes NRI $= -(\mathrm{MPD}_{obs} - \mathrm{MPD}_{rand})/\mathrm{sd}\dot{}\mathrm{MPD}_{rand}$, where MPD_{obs} is the average of all pairwise phylogenetic distances between the taxa in a local community, MPD_{rand} is the average of MPD calculated in n randomly constructed communities after shuffling all taxa in the regional pool and $\mathrm{sd}\dot{}\mathrm{MPD}_{rand}$ is the standard deviation of MPD_{rand} (Webb *et al.* 2002). This allows examining whether co-occurring taxa are more (positive NRI) or less (negative NRI) closely related than expected by chance. Thus, positive NRI values are related to phylogenetic clustering while negative values indicate phylogenetic overdispersion.

Second, we used the phylogenetic fuzzy-weighting method proposed by Pillar and Duarte (2010). Compared to NRI, which is blind to the taxonomic identity of coexisting lineages (i.e. similar NRIs can be obtained for communities dominated by closely related *Actinobacteria* or for communities dominated by closely related *Proteobacteria*), the fuzzy-weighting method identifies the representativeness of different lineages across the sites (see Duarte, Prieto and Pillar 2012 for a detailed explanation). Briefly, this method calculates a matrix (matrix P), that describes the species phylogenetic composition of each plot taking into account the phylogenetic neighborhood of each OTU. To obtain matrix P, we transformed the pairwise phylogenetic distance matrix on similarities between species. Then, we used similarities to weight the species composition matrix by a fuzzy set algorithm (Pillar and Duarte 2010). In matrix P, each OTU has a value per plot that increases as the phylogenetic distance between neighboring OTUs decreases. Matrix P was calculated using the SYNCSA package implemented in R (Debastiani and Pillar 2012). Principal components analysis (PCA) with Euclidean distance was run to reduce the dimensionality of the matrix P. The loadings of each OTU indicate the relative contribution of that OTU to differentiate plots along the first principal component axis (plot scores). Consequently, each plot score captures the whole variation of species abundances weighted by phylogenetic relatedness. To identify which phyla were responsible for the phylogenetic community structure, we ran a linear model with the plot scores along the first principal component axis (PCPS1 hereafter) as the dependent variable and the relative abundance of the most abundant phyla as independent variables.

Statistical analyses

To check whether spatial autocorrelation in the bacterial community composition across plots should be taken into account in subsequent analyses, we correlated OTU composition and geographic distance matrices through Mantel tests in the ADE4 package for R (Mantel 1967; Dray and Dufour 2007). We tested whether physical and chemical soil parameters determine the phylogenetic structure of bacterial communities by performing Bayesian generalized linear models (GLMs) with NRI and PCPS1 used individually as the dependent variables and the PC1-Soil as the independent variable. The NRI values per plot were very similar across the five phylogenetic trees in both sites ($r > 0.77$; $P < 0.005$ for all the correlations). Similarly, PCPS1 values per plot were very similar for all the trees ($r > 0.98$; $P < 0.005$ for all the correlations). Although these correlations indicate that phylogenetic uncertainty was small, we accommodated such small uncertainty by running five GLMs for each site, each one using the phylogenetic information calculated from an independent tree and integrated over the posterior samples by drawing 1000 random samples across models. The models were run with the help of MCMC techniques as implemented in the MCMCglmm package for R (Hadfield 2010). We used the default priors and ran 13 000 MCMC iterations with a burn-in period of 3000 iterations. Convergence of the chain was tested by means of an autocorrelation statistic. The statistical significance of the factors in the model was estimated by calculating the 95% credible interval of their posterior distribution.

Bayesian GLMs were also used to test which metrics of phylogenetic community structure predicted the ecosystem functioning more accurately. We ran five GLMs per site, using each ecosystem functioning parameter individually as the dependent variable and both NRI and PCPS1 as independent variables in the same model. The relative abundance of the most abundant clades was also used as a predictive parameter of ecosystem functioning. Clade relative abundances were estimated as the sum of the relative abundances of all OTUs that belonged to that particular clade, which were corrected based on their estimated 16S rRNA gene copy numbers (see details above). All analyses were performed using the software R 3.1.1 (R Core Team 2014).

RESULTS

Soil bacterial communities had 602 ± 13 and 430 ± 24 OTUs per plot (mean ± SE) in the non-patchy and patchy ecosystems, respectively. *Proteobacteria* was the most dominant phylum in both ecosystems followed by *Actinobacteria* (Fig. 1). There was not spatial autocorrelation across plots in the bacterial community composition (non-patchy ecosystem, $r = -0.205$, $P = 0.924$; patchy ecosystem, $r = -0.054$, $P = 0.691$; Mantel tests) nor in the phylogenetic structure measured as NRI (non-patchy ecosystem, $r = 0.04$, $P = 0.349$; patchy ecosystem, $r = -0.049$, $P = 0.74$) or PCPS1 (non-patchy ecosystem, $r = 0.29$, $P = 0.06$; patchy ecosystem, $r = -0.044$, $P = 0.78$).

The phylogenetic structure of the bacterial communities in both ecosystems was clustered, as indicated by NRI significantly higher than zero [NRI post-mean estimate (95% credible interval) = 1.70 (0.68, 2.93) for the non-patchy and NRI = 2.49 (1.37, 3.56) for the patchy ecosystems]. Phylogenetically clustered plots (i.e. high NRI values) were those with higher abundances of *Proteobacteria* and/or *Actinobacteria* (see positive estimates in the NRI models in Table 2). The contribution of both phyla was significantly positive but differed in their relative importance, with *Proteobacteria* and *Actinobacteria* equally contributing in the non-patchy ecosystem but *Actinobacteria* contribution becoming non-significant in the patchy ecosystem.

The metrics of the community structure that accounts for the variability in the taxonomic identity and the phylogenetic relatedness (PCPS1) explained 50 and 71% of the total variance of the phylogenetic structure in the non-patchy and

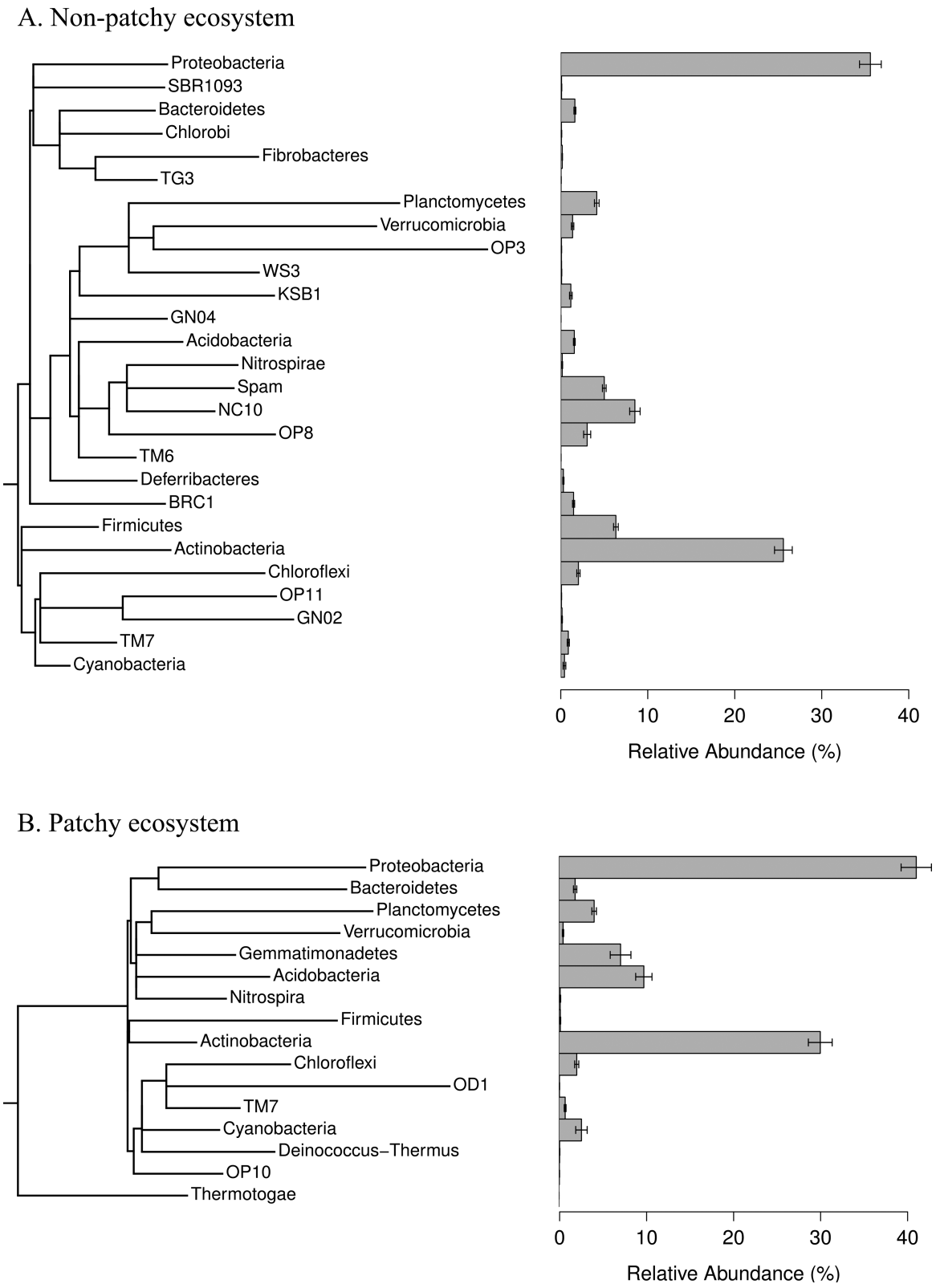

Figure 1. Phylogenetic tree of major basal groups in the non-patchy and patchy ecosystems showing the average relative abundances of each phylum across the study plots. Error bars indicate the standard error of the mean.

Table 2. Linear model explaining the contribution (% variance) of the abundance of the dominant phyla (% of OTUs belonging to *Proteobacteria* and *Actinobacteria*) on the mean NRI and on the mean plot scores along the first principal component axis of the phylogenetic community structure (PCPS1) across the five phylogenetic trees.

	NRI			PCPS1		
	Estimate ± SE	t	% variance	Estimate ± SE	t	%variance
Non-patchy ecosystem						
Intercept	−11.9 ± 3.02	−3.96**		136.7 ± 36.26	3.77**	
% *Proteobacteria*	0.22 ± 0.05	4.49**	33.4	0.61 ± 3.64	3.64**	52.5
% *Actinobacteria*	0.21 ± 0.06	3.45*	41.9	−8.43 ± 0.741	−11.25***	45.2
Patchy ecosystem						
Intercept	−3.54 ± 1.35	−2.61*		−197.4 ± 7.22	−27.37***	
% *Proteobacteria*	0.09 ± 0.02	3.28**	35.5	5.31 ± 0.14	35.64***	97.2
% *Actinobacteria*	0.07 ± 0.03	2.11*	0.09	−0.67 ± 018	−3.56**	0.009

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

patchy ecosystems, respectively. The contribution of the most abundant phyla to PCPS1 differed between ecosystems, with similar contributions of *Proteobacteria* and *Actinobacteria* in the non-patchy ecosystem but with an overwhelming contribution of *Proteobacteria* in the patchy ecosystem (Table 2). Interestingly, the phylogenetic position of both phyla in distant clades (see trees in Fig. 1) was accounted for by PCPS1 and clearly segregated the plots with preponderance of *Proteobacteria* in the right extreme from those with preponderance of *Actinobacteria* in the left extreme (see positive estimates for *Proteobacteria* and negative for *Actinobacteria* in the PCPS1 linear models in Table 2).

The first axis of the PCA grouping soil physical and chemical variables (PC1-Soil) accounted for 68 and 84% of the variance in non-patchy and patchy ecosystems, respectively. The loading factors showed that PC1-Soil represented a fertility gradient of increasing oxidizable carbon and humidity contents in both ecosystems (non-patchy ecosystem: TN 0.27, EC 0.36, pH 0.40, TOC 0.43, GH 0.47, PPi-OC 0.48; patchy ecosystem: pH –0.32, EC 0.40, GH 0.40, PPi-OC 0.43, TN 0.44, TOC 0.44). While the magnitude of such a gradient was slight in the non-patchy ecosystem (e.g. TOC ranged from 3.3 to 5% across plots), it was extremely accentuated in the patchy ecosystem (e.g. TOC ranged from 1.8 to 12.5% across plots; Table 1). This fertility gradient could not predict the NRI in the non-patchy ecosystem where both *Actinobacteria* and *Proteobacteria* had relevant contributions [NRI versus PC1-Soil = 0.15 (–0.19, 0.56)]. However, once the identities of both phyla and the variability in phylogenetic relatedness across plots were accounted for, the fertility gradient significantly explained the phylogenetic structure of the community [PCPS1 versus PC1-Soil = 11.66 (2.68, 23.19)]. In the patchy ecosystem where the taxonomic relevance of a single phylum (*Proteobacteria*) was disproportionate, the fertility gradient significantly explained both NRI [NRI versus PC1-Soil = 0.59 (0.36, 0.84)] and PCPS1 [PCPS1 versus PC1-Soil = 17.61 (23.36, 11.33)].

NRI did not predict any of the ecosystem functioning variables related to soil microbial productivity, metabolic efficiency and biogeochemical cycling in the non-patchy ecosystem while PCPS1 significantly explained most of the general indicators of microbial activity. Specifically, PCPS1 was negatively associated to MBC and MBC/TOC and positively to BR and qCO_2 (Fig. 2 upper panel, Model 1). Plots with high abundances of *Actinobacteria* were those with high MBC and high efficiency in converting organic C into microbial C (MBC/TOC) and conserving C (as indicated by the negative relationship with qCO_2) (Fig. 2 upper panel, Model 2). Plots with abundant *Proteobacteria* were those with high activity of decomposers that mineralize organic C into CO_2 (as indicated by BR) and PA (Fig. 2 upper panel Model 2). In the patchy ecosystem, most of the ecosystem functioning parameters, including indicators of both general microbial activity and specific enzymatic processes, were predicted by both NRI and PCPS1 (Fig. 2 bottom panel, Model 1). All these relationships were positive and were also explained by the relative abundance of *Proteobacteria* (Fig. 2 bottom panel, Model 2).

DISCUSSION

Our results show that the relationship between phylogenetic diversity and ecosystem functioning is dependent on the taxonomic identity of the main coexisting bacterial lineages. We show that the soil environment structures a phylogenetically clustered community and discuss the mechanisms underlying the relationship between such phylogenetic community structure and ecosystem functioning. To understand this relationship, we invoke the need to include the species identity in phylogenetic diversity metrics to account for variation in phylogenetically weighted abundances across communities.

The soil abiotic variables determined a fertility gradient that explained the phylogenetic structure of soil bacterial communities in both ecosystems. This correlates well with previous observations showing that the amount of oxidizable substances is a good predictor of the phylogenetic community structure of soil bacteria worldwide (Goberna, García and Verdú 2014). Our ability to explain the bacterial phylogenetic community structure through abiotic factors depended on the level of environmental heterogeneity. At high environmental heterogeneity (patchy ecosystem), the abiotic environment explained the community structure regardless the inclusion (PCPS1) or not (NRI) of lineage identity, while at low environmental heterogeneity (non-patchy ecosystem), only the most complex measure of community structure including lineage identity was predicted by the abiotic environment. Future studies in other ecosystems are needed to refine the relationship between environmental heterogeneity and the power of phylogenetic metrics to detect community structure in soil bacterial communities. Another important picture emerging from the present study is that environmentally mediated changes in the composition of bacterial communities left a phylogenetic signature in the community structure with profound implications in ecosystem functions. Detecting which lineage has been overrepresented under particular environmental parameters is key to understand the meaning of the phylogenetic clustering in the communities.

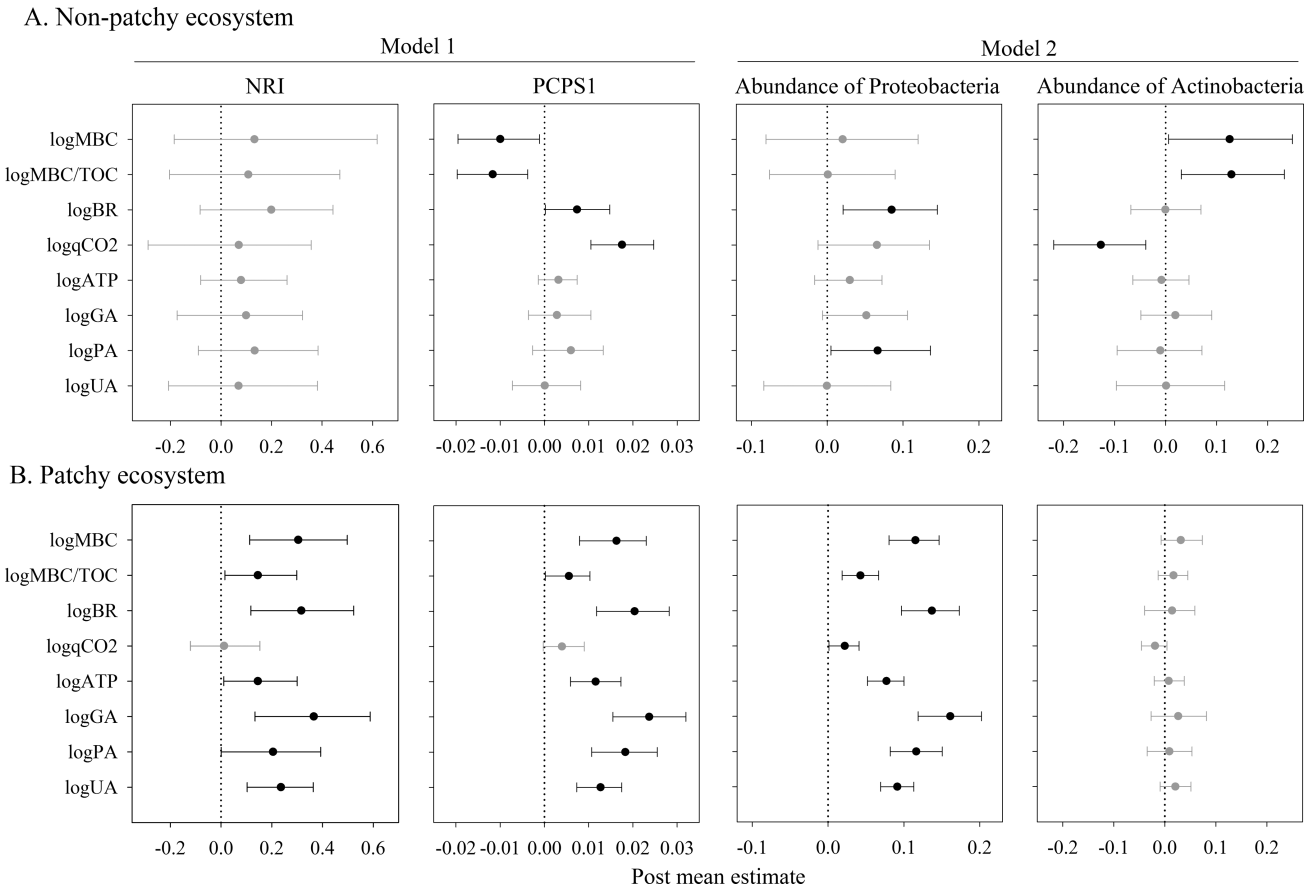

Figure 2. Bayesian post-mean estimates and their expected 95% credible intervals for the effect of NRI and PCPS1 (Model 1) and the relative abundance of *Proteobacteria* and *Actinobacteria* (Model 2) on the eight soil microbial indicators from (**A**) non-patchy and (**B**) patchy ecosystems. All variables were log-transformed to improve normality. Effects with intervals not including zero are significant (black-colored intervals), whereas those including zero are not significant (gray-colored intervals).

Ecosystem functioning was also better predicted by the metrics accounting for the identity of the lineages. In our non-patchy ecosystem, both *Actinobacteria* and *Proteobacteria* were key components structuring productive bacterial communities (Goldfarb *et al.* 2011). As both phyla are distantly related, their co-existence in more fertile plots was not translated into increased phylogenetic clustering as shown by the lack of correlation between the fertility gradient and NRI. Similarly, NRI could not predict any ecosystem function in this non-patchy ecosystem. However, the phylogenetic structure metrics accounting for the identity of both phyla predicted most of the general indicators of ecosystem functioning. On the other side, in the patchy ecosystem we found that communities phylogenetically clustered because of the overrepresentation of a particular clade (*Proteobacteria*) were the most productive. In this case, the coexistence of closely related *Proteobacteria* in fertile plots was translated into increased phylogenetic clustering, and therefore NRI could also predict high ecosystem functioning at low phylogenetic diversities.

Our results contrast with the common findings in 'macro'organisms that indicate that phylogenetic diversity is positively related to ecosystem functioning (Cadotte, Cardinale and Oakley 2008; Flynn *et al.* 2011; Cadotte 2013). They agree, however, with other lines of evidence showing variable responses of ecosystem functioning parameters to bacterial phylogenetic diversity. In simple experimental communities, positive and neutral responses of community productivity to increasing levels of phylogenetic diversity have been described (Gravel *et al.* 2012; Venail and Vives 2013). In some instances, positive responses could be experimentally attributed to complementarity effects based on the overyielding of the mixtures compared to their constituent species (Venail and Vives 2013), but this pattern is not consistent in the literature (Gravel *et al.* 2012). In more complex microcosms, bacterial productivity showed mostly negative, but also neutral and positive responses, to phylogenetic diversity (Severin, Östman and Lindström 2013). These authors suggest that negative responses are mediated by the overrepresentation of productive ß-*Proteobacteria* with the ability to consume an aromatic carbon compound. Similarly, our results in natural soil communities indicate that fast growing, competitively superior clades in the presence of soil organic carbon outcompete other clades, thus reducing phylogenetic diversity but rising indicators of ecosystem functioning. These results are consistent with fitness differences as the predominant mechanism causing high productivity at low phylodiversity through competitive exclusion (Mayfield and Levine 2010; Carroll, Cardinale and Nisbet 2011; HilleRisLambers *et al.* 2012).

In short, microbially mediated ecosystem functions can be predicted by the phylogenetic structure of soil bacterial communities because this metrics contains information on both the outcome of the ecological processes determining species coexistence and the functionality of these coexisting lineages. We suggest that outcompetition of big clades by very competitive and productive lineages explains both the phylogenetic diversity patterns of bacterial communities and the relationship between

diversity and ecosystem functioning. Capturing the ecological and evolutionary idiosyncrasies of the soil bacterial communities is crucial to understand the relationship between diversity and ecosystem functioning. The improvement in our prediction ability of the ecosystem functions performed by soil bacteria is of paramount importance given the relevance of these processes (i.e. biogeochemical cycling of nutrients, decomposition of organic matter, etc.) at the planetary level.

FUNDING

This work was supported by the Spanish Ministry of Science and Innovation (R+D Project CGL2011-29585-C02-01), the EU (FP7-PEOPLE-2009-RG-248155) and the BBVA foundation (project Mintegra; I Convocatoria de ayudas de la fundación BBVA a proyectos de investigación). EPV acknowledges support by the FPI programme and MG by the Ramón y Cajal Programme (Spanish Ministry of Economy and Competitiveness).

Conflict of interest. None declared.

REFERENCES

Balvanera P, Pfisterer AB, Buchmann N, *et al.* Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol Lett* 2006;**9**:1146–56.

Bell T, Newman JA, Silverman BW, *et al.* The contribution of species richness and composition to bacterial services. *Nature* 2005;**436**:1157–60.

Blomberg SP, Garland TJR, Ives AR. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evol Int J Org Evol* 2003;**57**:717–45.

Cadotte MW. Experimental evidence that evolutionarily diverse assemblages result in higher productivity. *P Natl Acad Sci USA* 2013;**110**:8996–9000.

Cadotte MW, Cardinale BJ, Oakley TH. Evolutionary history and the effect of biodiversity on plant productivity. *P Natl Acad Sci USA* 2008;**105**:17012–7.

Caporaso JG, Bittinger K, Bushman FD, *et al.* PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 2010a;**26**:266–7.

Caporaso JG, Kuczynski J, Stombaugh J, *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010b;**7**:335–6.

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.

Chesson P. Mechanisms of maintenance of species diversity. *Annu Rev Ecol Syst* 2000;**31**:343–66.

Curtis TP, Sloan WT, Scannell JW. Estimating prokaryotic diversity and its limits. *P Natl Acad Sci USA* 2002;**99**:10494–9.

Debastiani VJ, Pillar VD. SYNCSA—R tool for analysis of metacommunities based on functional traits and phylogeny of the community components. *Bioinformatics* 2012;**28**:2067–8.

DeSantis TZ, Hugenholtz P, Larsen N, *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microb* 2006;**72**:5069–72.

Dray S, Dufour AB. The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw* 2007;**22**:1–20.

Duarte LDS, Prieto PVMB, Pillar VD. Assessing spatial and environmental drivers of phylogenetic structure in Brazilian *Araucaria* forests. *Ecography* 2012;**35**:952–60.

FAO–ISRIC–IUSS. *World reference base for soil resources*. Rome: FAO, 2006.

Flynn DFB, Mirotchnick N, Jain M, *et al.* Functional and phylogenetic diversity as predictors of biodiversity–ecosystem-function relationships. *Ecology* 2011;**92**:1573–81.

Goberna M, García C, Insam H, *et al.* Burning fire-prone mediterranean shrublands: immediate changes in soil microbial community structure and ecosystem functions. *Microb Ecol* 2012;**64**:242–55.

Goberna M, García C, Verdú M. A role for biotic filtering in driving phylogenetic clustering in soil bacterial communities. *Glob Ecol Biogeogr* 2014;**23**:1346–55.

Goberna M, Navarro-Cano JA, Valiente-Banuet A, *et al.* Abiotic stress tolerance and competition-related traits underlie phylogenetic clustering in soil bacterial communities. *Ecol Lett* 2014;**17**:1191–201.

Godoy O, Kraft NJB, Levine JM. Phylogenetic relatedness and the determinants of competitive outcomes. *Ecol Lett* 2014;**17**:836–44.

Goldfarb KC, Karaoz U, Hanson CA, *et al.* Differential growth responses of soil bacterial taxa to carbon substrates of varying chemical recalcitrance. *Front Microbiol* 2011;**2**:94.

Gravel D, Bell T, Barbera C, *et al.* Phylogenetic constraints on ecosystem functioning. *Nat Commun* 2012;**3**:1117.

Griffiths BS, Ritz K, Wheatley R, *et al.* An examination of the biodiversity–ecosystem function relationship in arable soil microbial communities. *Soil Biol Biochem* 2001;**33**:1713–22.

Hadfield JD. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J Stat Softw* 2010;**33**:1–22.

Hamady M, Walker JJ, Harris JK, *et al.* Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods* 2008;**5**:235–7.

HilleRisLambers J, Adler PB, Harpole WS, *et al.* Rethinking community assembly through the lens of coexistence theory. *Annu Rev Ecol Evol S* 2012;**43**:227–48.

Hooper DU, Chapin FS, Ewel JJ, *et al.* Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecol Monogr* 2005;**75**:3–35.

Jones RT, Robeson MS, Lauber CL, *et al.* A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME J* 2009;**3**:442–53.

Jousset A, Schmid B, Scheu S, *et al.* Genotypic richness and dissimilarity opposingly affect ecosystem functioning. *Ecol Lett* 2011;**14**:537–45.

Kembel SW, Cowan PD, Helmus MR, *et al.* Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 2010;**26**:1463–4.

Kembel SW, Wu M, Eisen JA, *et al.* Incorporating 16s gene copy number information improves estimates of microbial diversity and abundance. *PLoS Comput Biol* 2012;**8**:e1002743.

Lane DJ. 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds). *Nucleic Acid Techniques in Bacterial Systematics*. New York, NY, USA: John Wiley & Sons, 1991, 115–75.

Lozupone CA, Knight R. Global patterns in bacterial diversity. *P Natl Acad Sci USA* 2007;**104**:11436–40.

Maherali H, Klironomos JN. Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science* 2007;**316**:1746–8.

Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer Res* 1967;**27**:209–20.

Martiny AC, Treseder K, Pusch G. Phylogenetic conservatism of functional traits in microorganisms. *ISME J* 2013;**7**:830–8.

Mayfield MM, Levine JM. Opposing effects of competitive exclusion on the phylogenetic structure of communities: phylogeny and coexistence. *Ecol Lett* 2010;**13**:1085–93.

Muyzer G, de Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microb* 1993;**59**:695–700.

Nannipieri P, Grego S, Ceccanti B. Ecological significance of the biological activity in soil. In: Bollag J, Stotzky G (eds). *Soil Biochemistry*. New York, NY, USA: Marcel Dekker, 1990, 293–355.

Navarro-Cano JA, Goberna M, Valiente-Banuet A, *et al.* Plant phylodiversity enhances soil microbial productivity in facilitation-driven communities. *Oecologia* 2014;**174**:909–20.

Pillar VD, Duarte LDS. A framework for metacommunity analysis of phylogenetic structure. *Ecol Lett* 2010;**13**:587–96.

Quast C, Pruesse E, Yilmaz P, *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**:D590–6.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2014.

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

Revell LJ, Harmon LJ, Collar DC. Phylogenetic signal, evolutionary process, and rate. *Syst Biol* 2008;**57**:591–601

Severin I, Östman Ö, Lindström ES. Variable effects of dispersal on productivity of bacterial communities due to changes in functional trait composition. *PLoS One* 2013;**8**:e80825.

Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 2006;**22**:2688–90.

Turner S, Pryer KM, Miao VPW, *et al.* Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J Eukaryot Microbiol* 1999;**46**:327–38.

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

Venail PA, Vives MJ. Phylogenetic distance and species richness interactively affect the productivity of bacterial communities. *Ecology* 2013;**94**:2529–36.

Wardle DA, Bardgett RD, Klironomos JN, *et al.* Ecological linkages between aboveground and belowground biota. *Science* 2004;**304**:1629–33.

Webb CO, Ackerly DD, McPeek MA, *et al.* Phylogenies and community ecology. *Annu Rev Ecol Syst* 2002;**33**:475–505.

Zak DR, Holmes WE, White DC, *et al.* Plant diversity, soil microbial communities, and ecosystem function: are there any links? *Ecology* 2003;**84**:2042–50.