Received Date : 14-Jul-2015
Revised Date : 21-Sep-2015
Accepted Date : 23-Sep-2015
Article type : Letters

Relating belowground microbial composition to the taxonomic, phylogenetic, and functional trait distributions of trees in a tropical forest

Albert Barberán[1*], Krista L. McGuire[2,3*], Jeffrey A. Wolf[3,4], F. Andrew Jones[5,6], S. Joseph Wright[6], Benjamin L. Turner[6], Adam Essene[7], Stephen P. Hubbell[4,6], Brant C. Faircloth[8], Noah Fierer [1,9 †]

[1]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

[2]Department of Biology, Barnard College, Columbia University, NY, USA

[3]Department of Ecology, Evolution and Environmental Biology, Columbia University, NY, USA

[4]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA

[5]Department of Botany and Plant Pathology, Oregon State University, OR, USA

[6]Smithsonian Tropical Research Institute, Apartado 0843–03092, Balboa, Ancon, Republic of Panama

[7]Department of Biological Sciences, Fordham University, NY, USA

[8]Department of Biological Sciences and Museum of Natural Science, Louisiana State University, LA, USA

[9]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, USA

**co-first authors*

†corresponding author

Statement of Authorship: A.B., K.L.M., and N.F. developed and framed research questions. J.A.W., F.A.J., S.J.W. B.L.T, K.L.M., N.F., S.P.H., and B.C.F. provided data used in this study and logistical support. A.E. helped with sample processing. Data analyses were led by A.B. with help from K.L.M., N.F. and J.A.W. The paper was primarily written by N.F., K.L.M., and A.B. with help from all co-authors.

Running title: Relating soil microbes to tree distributions

Keywords: soil, bacteria, fungi, microbial ecology, trees, roots, tropical forest, Barro Colorado Island, phylogeny, functional traits.

Type of article: Letters

148 words in the abstract, 4,752 words in the main text, 49 references, and 3 figures.

Corresponding author:
Noah Fierer
University of Colorado
CIRES, UCB 216
Boulder, CO 80309-0216
phone: 303-492-5615
email: Noah.Fierer@colorado.edu

ABSTRACT

The complexities of the relationships between plant and soil microbial communities remain unresolved. We determined the associations between plant aboveground and belowground (root) distributions and the communities of soil fungi and bacteria found across a diverse tropical forest plot. Soil microbial community composition was correlated with the taxonomic and phylogenetic structure of the aboveground plant assemblages even after controlling for differences in soil characteristics, but these relationships were stronger for fungi than for bacteria. In contrast to

expectations, the species composition of roots in our soil core samples was a poor predictor of microbial community composition perhaps due to the patchy, ephemeral, and highly overlapping nature of fine root distributions. Our ability to predict soil microbial composition was not improved by incorporating information on plant functional traits suggesting that the most commonly measured plant traits are not particularly useful for predicting the plot-level variability in belowground microbial communities.

INTRODUCTION

Soil microbes engage in complex feedbacks with plants (Wardle *et al.* 2004; Bever *et al.* 2013), although the specific mechanisms that link aboveground and belowground communities often remain unresolved. Predicted associations between plant and microbial distributions arise from known plant effects on soil conditions, including the amounts and types of organic carbon (C) inputs, soil pH, or soil nutrient availability through litter and root exudates (van der Heijden *et al.* 2008; Prescott & Grayston 2013). These soil modifications can have direct and indirect effects on local microbial communities by favoring the growth of symbiotic bacteria and fungi (including nitrogen-fixing bacteria and mycorrhizal fungi) or microbial pathogens (Wardle *et al.* 2004). However, plant and microbial assembly can also occur independently in response to similar abiotic properties of soil (Hines *et al.* 2006), or can occur at different spatial and temporal scales (Bardgett *et al.* 2005), making it difficult to identify generalizable patterns and associations between plants and belowground microbes.

Given the myriad of interactions between plants and microbes and their well-known impacts on ecosystem function (van der Heijden *et al.* 2008), it is often assumed that the composition of belowground microbial communities and aboveground plant communities will reflect one another. However, evidence to support this assertion is scarce (Prescott & Grayston 2013), with some studies showing that plant community composition is a significant predictor of overall bacterial and/or fungal community composition at regional or continental-scale studies (Prober *et al.* 2015), and other studies unable to find the same relationship (Talbot *et al.* 2014). Even when studies are tightly controlled, a correlation between plant species identity and overall bacterial or fungal community composition might (Jiang *et al.* 2012) or might not (McGuire *et al.* 2012) be present.

Various methodological and ecological factors could obscure the detection of tight associations between plants and belowground microbial communities. First, methodological constraints in microbial detection and quantification may make it difficult to identify patterns between above and belowground communities (Fierer & Lennon 2011). Second, if the environmental factors structuring plant communities are distinct from those that structure the overall composition of belowground microbial communities, relationships between individual plant species and microbial community structure would be unlikely. For example, some studies show that soil pH can have a large influence on the composition of soil bacterial communities,

but little influence on plant species composition, contributing to an apparent disconnect between plant communities and the bacterial communities found in soil (Prober *et al.* 2015). Third, plants may influence belowground microbial communities via specific plant traits or attributes that are not predictable from plant species identity alone such as litter C:N ratio, root exudate production, or litter carbon chemistry. There is a growing body of evidence suggesting that functional traits or the evolutionary relatedness of plants can often be more useful predictors of plant effects on belowground biota than species identity (De Deyn & Van der Putten 2005). Finally, soil microbial communities may not respond to aboveground plant distributions, but rather to the fine-scale distributions of plant root networks, because the spatial scale of aboveground ground canopies and belowground distributions of plant species may not be equivalent (Jones *et al.* 2011).

In order to disentangle the associations between plant and belowground microbial communities we combined detailed microbial community analyses with integrated assessments of belowground and aboveground plant distributions that considered the taxonomic identities of plants, the phylogenetic relationships among plants, and plant functional traits. To accomplish these objectives we used marker gene sequencing to identify plant roots and characterize the bacterial and fungal communities found in each of 625 soil samples collected from a 50-ha tropical forest plot located on Barro Colorado Island, Panama, a plot that has been intensively studied by plant ecologists for decades and has been a focal point for tropical research investigating tree coexistence and the maintenance of plant diversity (Hubbell & Foster 1983). We used this dataset to determine the extent to which the variability in the composition of soil bacterial and fungal communities across the plot was predicted by above and belowground plant species distributions. We tested the prediction that the association between plants and belowground microbial communities would be stronger for belowground than aboveground plant distributions because roots are expected to have important influences on the development of soil microbial communities. In addition, we tested the hypothesis that phylogenetic relatedness and the functional traits of plants would better predict belowground microbial distributions than plant species identity alone since more closely related plant species would be expected to share morphological and functional traits that are important for structuring soil microbial communities (Cantarel *et al.* 2015).

METHODS

Soil Sampling. We collected 625 surface soil cores (6.25 cm diameter x 20 cm depth) from the 0.5 km^2 Barro Colorado Island (BCI) forest dynamics plot, Republic of Panama (9.15° N, 79.8° W) (Hubbell & Foster 1983) between October 14th and December 2nd of 2010. Cores were sampled 1 m west of the central 5 m grid marker in every other 20 m by 20 m quadrat. If there was an obstruction that precluded sampling in this location, we sampled at a nearby location. We sampled even columns on odd rows and odd columns on even rows of the 20 m grid system, resulting in approximately 28 m spacing between nearest neighbor cores. Although our sampling scheme provided far higher spatial resolution than nearly any comparable study, we acknowledge that, by collecting soil cores spaced 28 m from on another, we may still lack the spatial resolution to capture detailed above-belowground relationships given the heterogeneous nature of tropical forests and the soils at this site. We georeferenced the soil cores at all sites using a differential GPS parameterized affine function (Wolf *et al.* 2015). Locations of the cores on both the plot coordinate system and a geographic coordinate system (WGS 1984 UTM Zone 17N, EPSG: 32617) are available in Wolf *et al.* (2015).

From each sampled location, we collected mineral soil (0-20 cm), excluding litter and organic horizons, using bulk soil rather than just rhizosphere soil for subsequent molecular analyses of microbial communities. Immediately after collecting the soil core, we placed the entire core in a Ziploc bag, thoroughly mixed the soil, and subsampled approximately 15 mL of homogenized soil (no roots) into a Whirl-pak bag for microbial analyses. Soils were frozen at -20°C on the same day of collection and were subsequently transported to Barnard College, Columbia University where they were stored at -20°C until analysis. Prior to molecular analyses, we sieved soils through sterilized 2 mm sieves to homogenize the microbial community and remove any remaining rocks and non-soil fragments.

From the remaining soil, we collected roots for molecular analysis and subsampled 50 g of soil for air-drying and subsequent soil chemical analyses, which are described in Wolf *et al.* (2015). Briefly, we analyzed $BaCl_2$-extractable cations, and P from Mehlich-3 extractions, total carbon and total nitrogen, and soil pH in both H_2O and 0.1 M $CaCl_2$ solution.

Molecular analyses of microbial communities. Microbial diversity was assessed using high-throughput sequencing methods to characterize the variation in taxonomic marker gene

sequences. For bacterial analyses, we sequenced the V4 hypervariable region of the 16S rRNA gene using the 515-F (GTGCCAGCMGCCGCGGTAA) and 806-R (GGACTACHVGGGTWTCTAAT) primer pair (Fierer *et al.* 2012). Although this primer pair also captures Archaea, the number of 16S rRNA reads from Archaea was very low in this dataset (<2.2% of the total number of phylotypes and <1.9% of total 16S rRNA sequences). For the fungal analyses, we sequenced the first internal transcribed spacer (ITS1) region of the rRNA operon using the ITS1-F (CTTGGTCATTTAGAGGAAGTAA) and ITS2 (GCTGCGTTCTTCATCGATGC) primer pair (McGuire *et al.* 2013). The primers included the appropriate Illumina adapters with the reverse primers also having an error-correcting 12-bp barcode unique to each sample to permit multiplexing of samples. PCR products from all samples were quantified using the PicoGreen dsDNA assay, and pooled together in equimolar concentrations for sequencing on an Illumina MiSeq instrument running the 2x150bp chemistry. All sequencing runs were conducted at the University of Colorado Next Generation Sequencing Facility.

The forward reads were demultiplexed using a custom Python script (https://github.com/leffj/helper-code-for-uparse), with quality filtering and phylotype clustering conducted using the UPARSE pipeline (Edgar 2013). For quality filtering, we used a maxee value of 0.5 (indicating that on average a maximum of 0.5 nucleotides were incorrectly assigned in every sequence). Sequences were also dereplicated and singleton sequences were removed prior to phylotype determinations. Representative sequences from the phylotypes that were not ≥75% similar to sequences contained in either the Greengenes 13_8 database (McDonald *et al.* 2012) or the UNITE May, 2014 database (Abarenkov *et al.* 2010) for 16S and ITS rRNA sequences, respectively, were discarded. Raw sequences were then mapped to phylotypes at the 97% similarity threshold. Phylotype taxonomy was determined using the RDP classifier with a confidence threshold of 0.5 (Wang *et al.* 2007) trained on the respective databases for 16S and ITS rRNA sequences. Sequences representing any phylotypes classified as mitochondria or chloroplast were removed. In order to reduce potential amplicon sequencing biases, we first removed samples with less than 10,000 sequences and then we normalized the sequence counts using a cumulative-sum scaling approach (Paulson *et al.* 2013). The total number of samples included in downstream analyses was 556 for bacteria and 480 for fungi. Representative sequences, phylotype abundance tables, and corresponding sample information are publicly

available in FigShare (http://figshare.com/articles/Soil_microbial_communities_Barro_Colorado/1449286).

Determination of plant distributions. Aboveground stem distributions were determined from the 2010 BCI census. Every woody and palm stem ≥ 1.0 cm diameter at breast height (DBH) was sampled. Tree community composition was determined for five neighborhood sizes (i.e. at 2.5, 5, 10, 15, and 20 m radii from the soil cores sampling points). Lianas were excluded from subsequent analyses. Palms were included in all analyses and, while we recognize that they are monocots and not woody trees, we use the term "tree" throughout the paper for ease of distinguishing canopy and subcanopy plant species from understory plant species (which were not included in the current study).

We assessed belowground root distributions so we could simultaneously determine the relationships between plant aboveground or belowground distributions and soil microbial distribution patterns. We did this by conducting molecular analyses of bulk fine root samples that were separated from individual soil cores after washing soil through a 0.1 mm sieve. We stored the fine roots from individual soil cores in a Tris-Low EDTA (TLE) buffer solution until dehydration by oven drying at 37 °C and homogenization by mortar and pestle with liquid nitrogen. We used approximately 0.1 g of homogenized root tissue to extract DNA using a modified CTAB DNA extraction technique (Li *et al.* 2007). We removed PCR inhibitors from DNA extracts using a standard bead-based DNA cleanup with 1.8x AMPure substitute (Rohland & Reich 2012). Following DNA cleanup, we amplified root DNA using sequence-tagged *rbcL* primers (Kress *et al.* 2009). We visually confirmed amplification success by running fragments on 1.5% agarose and we normalized amplicons across plates by adding 10 µL of PCR product to SequalPrep Normalization Kits (Invitrogen, Inc.) using the standard protocol. After normalization, we combined 6 µL from each normalized sample in a single tube, concentrated the combined DNA in a SpeedVac, and rehydrated the concentrated DNA with 16 µL 10 mM Tris-HCl. To apply sequence tags and sequencing adapters to each amplicon, we end-repaired and adenylated the rehydrated PCR amplicons by adding 16 µL of each amplicon pool to 1 µL End Prep Enzyme Mix (New England Biolabs), 2.5 µL reaction buffer, and 5.5 µL ddH_2O. We incubated this reaction for 25 minutes at 25 °C followed by 20 minutes at 72 °C. We added 1 µL of 25 mM 454 sequencing adapters (indexed) to the pool of normalized amplicons along with 1

µL Quick T4 DNA ligase (New England Biolabs), and we incubated the ligation reaction for 10 minutes at 25 °C followed by 10 minutes at 65 °C. Following ligation, we added 38 µL of ddH_2O to each reaction along with 48.6 AMPure XP beads (Beckman Coulter Inc.), and cleaned reactions following the standard AMPure protocol. We used qPCR to quantify each pool of amplicons, and we prepared an equimolar mixture of 1.0E8 copies by combining pools of amplicons together, and checked the size distribution of amplicons using an Agilent BioAnalyzer prior to sequencing. We sequenced the equimolar pools using either the Roche 454 FLX (1/8 lane) or the 454 Jr. (2 runs) at the UCLA Genotyping Core.

Prior to classifying sequences to species, we created a composite reference database by combining sequences from the database of BCI trees from Kress *et al.* (2009). We supplemented the composite reference database with sequences from Genbank for those species that occur on BCI but were absent from Kress et al. (2009). To classify sequence reads to woody plant species, we first removed reads shorter than 400 bp from the dataset. Then we removed core samples having fewer than 30 total sequence reads, which we believe represented *rbcl* amplification failures. Following this filtering, we performed a BLAT search of all sequence reads to all species in the composite database using a 98% similarity cutoff. We assigned a sequence read to a given reference species by minimizing the number of base pair mismatches between sequence read and reference data. Because *rbcL* is not completely diagnostic for all woody plant species on the BCI plot (Kress et al. 2009), we assigned ambiguous sequences (i.e. sequences with more than one matching species and the same number of base pair mismatches) to species in the reference database by computing the relative abundance of the potential matches in the plant neighborhood within a distance of 15 m from the location of the soil core and assigning species identity based upon the potential match having the highest relative abundance (sum of total basal area of all stems > 1 cm of the candidate species). We selected 15 m as our radius because this value is the maximum estimated crown radius for individuals in the neighborhood around each sampling point (see below). If none of the potential matches were present within the 15 m neighborhood, we removed the ambiguous sequence from the analysis. We also removed singletons (i.e. species with only one sequence present in a core) from the data set. The total number of root samples included in downstream analyses was 536.

Plant functional traits and phylogeny. We obtained the following plant functional traits for 248 of the 288 tree species (all woody and palm species ≥ 1.0 cm DBH) found in the 20 m neighborhoods surrounding the soil cores: wood density, fruit mass, seed mass, average DBH, growth rate, mortality rate, leaf morphology (leaf area, leaf thickness, leaf mass per area), and leaf elemental chemistry (concentration of aluminum, calcium, potassium, magnesium, phosphorous, nitrogen and carbon) (Wright *et al.* 2010). For the 248 tree species for which we had trait information, phylogenetic relationships among these species were estimated according to a maximum-likelihood phylogeny (Kress *et al.* 2009). After multiple sequence alignment using MAFFT (Katoh & Standley 2013) and trimming of poorly aligned positions using GBLOCKS (Castresana 2000), phylogenetic relationships from root samples were estimated with the FastTree approximate maximum-likelihood algorithm (Price *et al.* 2010) and the mid-point method for rooting.

The distance between plant assemblages as determined from their trait distributions was calculated using the abundance weighted mean pairwise distance (MPD). MPD is defined as the average functional distance separating two species drawn at random from different assemblages (Webb 2000). Plant phylogenetic assemblage distance was calculated using the abundance weighted UniFrac metric, which is defined as percent of branch length unique to any pair of assemblages (Lozupone & Knight 2005). Functional trait and phylogenetic analyses were carried out in the R environment (www.r-project.org) using the ape (http://ape-package.ird.fr/) and picante (http://picante.r-forge.r-project.org/) packages.

Statistical analyses. Patterns in plant and microbial community similarity were represented by non-metric multidimensional scaling (NMDS) using the Bray-Curtis distance metric after Hellinger standardization. We used partial Mantel tests controlling for the potential confounding effects of topography (slope) and soil pH as well as quantile regressions as implemented in the R package quantreg (https://cran.r-project.org/web/packages/quantreg/) to determine the relationships between plant and microbial community distance matrices. To estimate the explanatory power of individual soil and topographical variables on soil microbial communities and tree assemblages, we used permutational multivariate analysis of variance (PERMANOVA). Multivariate statistical analyses were implemented using the R packages vegan (http://vegan.r-forge.r-project.org/) and ecodist (http://cran.r-project.org/web/packages/ecodist/).

RESULTS AND DISCUSSION

Characteristics of the soil microbial communities at BCI

A total of of 33,480 bacterial and 24,610 fungal phylotypes were detected across all samples. Each soil sample collected from the 50-ha plot contained an average of 2,600 bacterial phylotypes and 600 fungal phylotypes (Fig. S1). For trees, the average number of aboveground species associated with each soil sample was 6, 17, 42, 64, and 83 at increasing neighborhood distances (2.5, 5, 10, 15 and 20 m radius, respectively) away from each soil sampling location, with a total of 288 trees species counted across all samples. On average, we identified roots from nine tree species in each soil core, with a total of 203 tree species detected in all soil cores (Fig. S1). We observed no significant relationship between patterns in plant and belowground microbial richness levels across the 0.5 km^2 plot (Fig. S2).

At the phylum level, the soil bacterial communities were dominated by Proteobacteria, Acidobacteria, and Verrucomicrobia (34%, 21%, and 9% of 16S rRNA sequences, respectively). The fungal communities were predominately composed of taxa within the Ascomycota and Basidiomycota phyla (66% and 27% of the fungal ITS1 sequences, respectively) (Fig. S3). Tree species from the orders Gentianales (20%), Malpighiales (18%) and Sapindales (15%) dominated aboveground samples, while Gentianales (23%), Rosales (16%), Fabales (14%) and Malpighiales (10%) dominated belowground samples (Fig. S3).

Soil microbial community composition is correlated with aboveground tree distributions

For both soil bacteria and fungi, community similarity was weakly related to the geographic location of the sample (Fig. S4). As has been demonstrated previously (John *et al.* 2007), tree assemblage patterns were correlated with topography and soil characteristics (Fig. S5), and some of these same factors were also predictive of fungal and bacterial community composition (Fig. S5). In particular, tree community composition was associated with slope (PERMANOVA: R^2 = 0.05, $P < 0.001$), while microbial community composition was associated with soil pH. Bacterial communities showed a stronger relationship with soil pH than the fungal communities (PERMANOVA: $R^2 = 0.16$, $P < 0.001$ and $R^2 = 0.06$, $P < 0.001$, respectively; Fig. S5). These results are in line with other studies suggesting that the composition of soil bacterial

247 communities, and to a lesser degree soil fungal communities, can be strongly influenced by
248 differences in soil pH (Prober *et al.* 2015).
249 Both bacterial and fungal community distance patterns were positively associated with
250 aboveground tree assemblage distance patterns (spatial visualization of ordination axes in Fig. 1,
251 Fig. 2, Fig. S6). We observed this positive relationship between soil microbial communities and
252 tree assemblages after controlling for the potential confounding effects of shared environmental
253 and habitat correlations by including topography (slope) and soil pH as covariates (Fig. 2). Thus,
254 the correlations between tree distributions and belowground microbial community composition
255 were not just driven by shared responses to changes in soil characteristics and topography across
256 the plot. Rather, our data suggest that the observed correlations between the plant and soil
257 microbial communities are related to aboveground-belowground feedbacks (Bever 2003),
258 acknowledging that unmeasured environmental parameters or other biotic interactions might also
259 be contributing to the observed patterns. When we re-ran these same models using only the
260 largest trees in the plot (only those 27 tree species that had DBH values >750 mm), we found
261 weaker associations ($r = 0.13$ and $r = 0.16$ for the largest trees at 20 m neighborhood compared
262 to $r = 0.18$ and $r = 0.27$ for the whole dataset at 20 m neighborhood, for bacteria and fungi
263 respectively). These results demonstrate that the correlations between soil microbial
264 communities and aboveground tree distributions are not just driven by the largest trees in the
265 neighborhood analyses.
266 Fungal community composition showed a stronger relationship with aboveground tree
267 distributions than the bacterial communities (Fig. 2, Fig. S6), which is likely due to physiological
268 and ecological differences between these groups (Waring *et al.* 2013). Fungi tend to depend
269 more directly on plant products such as structural leaf litter compounds and root exudates
270 (Broeckling *et al.* 2008) and are key decomposers of plant necromass (Boddy *et al.* 2008). Some
271 groups of fungi also display resource-use specialization on individual organic C and N
272 compounds found in leaf litter (McGuire *et al.* 2010). Surprisingly, the relationship between
273 arbuscular mycorrhizal fungi (Glomeromycota) and aboveground tree community composition
274 was weaker than the one observed for overall fungal community composition ($r = 0.11$, $P < 0.01$
275 for Glomeromycota compared to $r = 0.27$, $P < 0.01$ for all fungi).
276 When we compared the aboveground plant assemblage data at different spatial
277 neighborhood sizes (2.5, 5, 10, 15 and 20 m), we found that the correlations between

aboveground tree community composition and bacterial or fungal community composition were strongest at the largest neighborhood size (20 m) (Fig. 2, Fig. S6). In other words, if we want to predict what types of bacteria or fungi we will find in an individual soil sample, it is most useful to know what trees are within a 20 m radius of that sample. It was somewhat surprising that the identity of trees very close to the soil sample was not nearly as useful for predicting microbial composition, because we would expect leaf litter inputs to be highest by trees closer to the sampling site. Our finding that the 20 m neighborhood was more closely related to microbial composition suggests that plant community information collected at that scale may more effectively capture those plants that could be influencing soil at a given site. It may also be that non-additive litter mixture effects due to canopy overlap structure soil microbes in ways that are not predictable from single-species litter effects (Chapman *et al.* 2013). Additionally, differences in phenology (in particular, litter fall) may be another potential explanation for the observed relationships between tree assemblages and soil microbial community composition. For example, a recent study found that variation in litter inputs explained a significant proportion of soil C:N variability and that particular tree species dominated the litter inputs with seasonal differences (Uriarte *et al.* 2015). While litter phenology has not been explicitly evaluated on BCI, variability in species-specific litter phenology could explain the lack of a tight correlation between plant and microbial composition at smaller neighborhood scales, particularly if litter from a more productive, neighboring tree species was more influential on microbial composition at a given sampling point and time.

While phylogenetic relationships have been useful for explaining community assembly patterns for a variety of organisms (Cavender-Bares *et al.* 2009), and phylogenetic information can help explain some of the variation in some leaf-associated microbial communities (Kembel *et al.* 2014), incorporating phylogenetic information did not improve our ability to predict soil microbial distributions better than knowing species identity, alone (Fig. 3, Fig. S7). One possible explanation for this finding is that the high taxonomic diversity of trees in the plot obscured the relative effects of phylogenetic versus taxonomic distance, since these indices were highly correlated ($r = 0.86$, $P < 0.001$). Alternatively, the plant traits that shape belowground microbial communities may not be predictable from plant phylogeny due to convergence in selected phenotypic characteristics (Kursar *et al.* 2009).

Root distributions are poor predictors of microbial community composition

While we found fairly strong relationships between aboveground tree composition and soil microbial communities, there were only weak correlations between root distributions and microbial community composition (Fig. 1, 2, Fig. S6), which conflicts with our original prediction. This weak correlation was maintained when we conducted phylogenetic analyses of the root communities ($r = 0.061$, $P < 0.001$; $r = 0.128$, $P < 0.001$, for bacteria and fungi, respectively). Focusing solely on the composition of Glomeromycota did not improve the correlation with root composition ($r = 0.11$, $P = 0.11$). This result may be due to the fact that the distributions of aboveground trees and their roots are not well correlated at this site (Jones *et al*. 2011), which implies that either aboveground or belowground species-specific effects are spatially decoupled or that species-specific root zones may not actually exist like they do for aboveground neighborhoods. There is some evidence for the latter hypothesis from a recent temperate forest study which found that roots from multiple tree species migrated to nutrient-enriched soil patches and diminished belowground species segregation across the forest (Valverde-Barrantes *et al.* 2015).

Another explanation for the weak relationship between root and microbial composition is that roots are highly mobile and root distributions change quickly over time. Thus, temporal shifts in microbial communities may be uncoupled from the corresponding temporal dynamics of roots at this site. The average rate of fine root turnover for tree species in an adjacent forest was ~120 days, although significant species-specific variation likely exists (Yavitt *et al.* 2009). In addition, the microbes that are likely to be most affected by root dynamics may be rhizosphere specialists living on or around root surfaces rather than the microbes residing in bulk soil. Numerous studies have found that rhizosphere microbial communities are distinct from bulk soil microbial communities (Prescott & Grayston 2013), so species-specific effects of tree roots may be realized at much smaller spatial scales in the rhizosphere than could be detected from the bulk soil analyses conducted here. More manipulative studies are clearly necessary to unravel these alternative explanations, as most studies tend to have an aboveground bias and very little work has been done on the distributions of tropical tree roots (Iversen 2014).

Are tree traits predictive of soil microbial community composition?

While the taxonomic and phylogenetic metrics of tree assemblage composition were correlated with belowground bacterial and fungal communities, the aboveground leaf and stem traits were not good predictors of microbial community composition. Overall functional assemblage distance based on traits using the abundance weighted mean pairwise functional distance (Webb 2000) was not correlated with soil microbial community composition (Fig. 3, Fig. S7). Given that there is a growing body of literature suggesting that plant trait information is useful for predicting community assembly and ecosystem services (De Deyn & Van der Putten 2005; Cornwell *et al.* 2008), we were surprised to find that most plant trait distributions were not well-correlated with microbial distributions across the plot. However, it must be noted that most of the work linking microbial communities and microbially-mediated processes to plant traits has been done in temperate grassland ecosystems (Cantarel *et al.* 2015), and it is possible that the patterns and processes documented in herbaceous ecosystems may not be observed in diverse tropical forests. It is also possible that other unmeasured traits, such as litter carbon chemistry or root traits (Cantarel *et al.* 2015), are more important drivers of microbial composition than the traits included in our analyses.

We also explored specific relationships between individual assemblage-weighted functional traits and soil bacterial or fungal community composition. Although we did not observe significant relationships between the overall, aggregated aboveground traits of the plant assemblages and the composition of the soil microbial communities, several individual traits showed weak correlations (Fig. S8). Specifically, Al content in leaves and tree mortality rate were correlated with both bacterial and fungal community composition. The relationship with Al may reflect the toxicity of that element to microbial cells (Pina & Cervantes 1996). For fungal communities, carbon and Ca content in leaves were also correlated with community composition (Fig. S8), which may be related to the effects of these plant traits on soil C and litter input quantities. The relationship of leaf Ca concentrations and fungal composition likely reflects the important role of Ca in fungal growth, nutrient uptake, and mycorrhizal symbiosis (Pera & Callieri 1997).

CONCLUSIONS

Despite the high tree diversity, we found significant correlations between tree composition, soil pH, and the composition of soil microbial communities at the local scale. We also found

correlations between microbial composition and a subset of assemblage-weighted plant traits. The lack of strong relationships between soil microbial composition and root distributions was surprising, but may be related to the ephemeral and mobile nature of fine roots. Since aboveground tree composition is less dynamic than belowground fine root turnover, the integrated effects of decadal litter inputs may be a more significant factor structuring soil microbial communities than root inputs.

ACKNOWLEDGEMENTS

We thank the Center for Tropical Forest Science for logistical support and Jonathan W. Leff for help with the sequence processing. Soil sampling and export permits were granted by the U.S. Department of Agriculture and the Smithsonian Tropical Research Institute to K.M. This work was supported in part by grants from the U.S. National Science Foundation (to N.F., DEB-0953331). Root work was supported by a Scholarly studies grant to SPH and BCF. FAJ acknowledges support from Oregon State University and the National Science Foundation (DEB-1257976). A.B. was supported by a James S. McDonnell Postdoctoral Fellowship. The BCI Forest Dynamics Research Project was made possible by National Science Foundation grants to SPH: DEB-0640386, DEB-0425651, DEB-0346488, DEB-0129874, DEB-00753102, DEB-9909347, DEB-9615226, DEB-9615226, DEB-9405933, DEB-9221033, DEB-9100058, DEB-8906869, DEB-8605042, DEB-8206992, DEB-7922197, support from the Center for Tropical Forest Science, the Smithsonian Tropical Research Institute, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Small World Institute Fund, and numerous private individuals, and through the hard work of over 100 people from 10 countries over the past two decades. The plot project is part the Center for Tropical Forest Science, a global network of large-scale demographic tree plots.

REFERENCES

1. Abarenkov, K., Nilsson, R.H., Larsson, K.H., Alexander, I.J., Eberhardt, U., Erland, S. *et al.* (2010). The UNITE database for molecular identification of fungi - recent updates and future perspectives. *New Phytologist*, 186, 281-285.
2.

Bardgett, R.D., Bowman, W.D., Kaufmann, R. & Schmidt, S.K. (2005). A temporal approach to linking aboveground and belowground ecology. *Trends in Ecology & Evolution*, 20, 634-641.

3.

Bever, J.D. (2003). Soil community feedback and the coexistence of competitors: conceptual frameworks and empirical tests. *New Phytologist*, 157, 465-473.

4.

Bever, J.D., Broadhurst, L.M. & Thrall, P.H. (2013). Microbial phylotype composition and diversity predicts plant productivity and plant-soil feedbacks. *Ecology Letters*, 16, 167-174.

5.

Boddy, L., Frankland, J.C. & Van West, P. (2008). *Ecology of Saprotrophic Basidiomycetes*. Elsevier Ltd., London.

6.

Broeckling, C.D., Broz, A.K., Bergelson, J., Manter, D.K. & Vivanco, J.M. (2008). Root exudates regulate soil fungal community composition and diversty. *Applied and Environmental Microbiology*, 74, 738-744.

7.

Cantarel, A.A.M., Pommier, T., Desclos-Theveniau, M., Diquélou, S., Dumont, M., Grassein, F. *et al.* (2015). Using plant traits to explain plant–microbe relationships involved in nitrogen acquisition. *Ecology*, 96, 788-799.

8.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17, 540-552.

9.

Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters*, 12, 693-715.

10.

Chapman, S.K., Newman, G.S., Hart, S.C., Schweitzer, J.A. & Koch, G.W. (2013). Leaf litter mixtures alter microbial community development: mechanisms for non-additive effects in litter decomposition. *PLoS One*, 8, e62671.

11.
Cornwell, W.K., Cornelissen, J.H.C., Amatangelo, K., Dorrepaal, E., Eviner, V.T., Godoy, O. *et al.* (2008). Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecology Letters*, 11, 1065-1071.
12.
De Deyn, G.B. & Van der Putten, W.H. (2005). Linking aboveground and belowground diversity. *Trends in Ecology & Evolution*, 20, 625-633.
13.
Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods*, 10, 996-998.
14.
Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L. *et al.* (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 21390-21395.
15.
Fierer, N. & Lennon, J.T. (2011). The generation and maintenance of diversity in microbial communities. *American Journal of Botany*, 98, 439-448.
16.
Hines, J., Megonigal, J.P. & Denno, R.F. (2006). Nutrient subsidies to belowground microbes impact aboveground food web interactions. *Ecology*, 87, 1542-1555.
17.
Hubbell, S.P. & Foster, R.B. (1983). Commonness and rarity in a Neotropical forest: implications for tropical tree conservation. In: *Community Ecology* (eds. Diamond, JM & Case, TJ). Harper and Row New York.
18.
Iversen, C.M. (2014). Using root form to improve our understanding of root function. *New Phytologist*, 203, 707-709.
19.

461 Jiang, Y.M., Chen, C.R., Xu, Z.H. & Liu, Y.Q. (2012). Effects of single and mixed species forest
462 ecosystems on diversity and function of soil microbial community in subtropical China.
463 *Journal of Soils and Sediments*, 12, 228-240.
464 20.
465 John, R., Dalling, J.W., Harms, K.E., Yavitt, J.B., Stallard, R.F., Mirabello, M. *et al.* (2007). Soil
466 nutrients influence spatial distributions of tropical tree species. *Proceedings of the*
467 *National Academy of Sciences of the United States of America*, 104, 864-869.
468 21.
469 Jones, F.A., Erickson, D.L., Bernal, M.A., Bermingham, E., Kress, W.J., Herre, E.A. *et al.*
470 (2011). The roots of diversity: below ground species richness and rooting distributions in
471 a tropical forest revealed by DNA barcodes and inverse modeling. *PLoS One*, 6, e24506.
472 22.
473 Katoh, K. & Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7:
474 improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772-
475 780.
476 23.
477 Kembel, S.W., O'Connor, T.K., Arnold, H.K., Hubbell, S.P., Wright, S.J. & Green, J.L. (2014).
478 Relationships between phyllosphere bacterial communities and plant functional traits in a
479 neotropical forest. *Proceedings of the National Academy of Sciences of the United States*
480 *of America*, 111, 13715-13720.
481 24.
482 Kress, W.J., Erickson, D.L., Jones, F.A., Swenson, N.G., Perez, R., Sanjur, O. *et al.* (2009). Plant
483 DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama.
484 *Proceedings of the National Academy of Sciences of the United States of America*, 106,
485 18621-18626.
486 25.
487 Kursar, T.A., Dexter, K.G., Lokvam, J., Pennington, R.T., Richardson, J.E., Weber, M.G. *et al.*
488 (2009). The evolution of antiherbivore defenses and their contribution to species
489 coexistence in the tropical tree genus Inga. *Proceedings of the National Academy of*
490 *Sciences of the United States of America*, 106, 18073-18078.
491 26.

492 Li, J., Yang, J., Chen, D., Zhang, X. & Tang, Z. (2007). An optimized mini-preparation method
493 to obtain high-quality genomic DNA from mature leaves of sunflower. *Genet Mol Res*, 6,
494 1064-1071.
495 27.
496 Lozupone, C. & Knight, R. (2005). UniFrac: a new phylogenetic method for comparing
497 microbial communities. *Applied and Environmental Microbiology*, 71, 8228-8235.
498 28.
499 McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A. *et al.*
500 (2012). An improved Greengenes taxonomy with explicit ranks for ecological and
501 evolutionary analyses of bacteria and archaea. *ISME Journal*, 6, 610-618.
502 29.
503 McGuire, K.L., Bent, E., Borneman, J., Majumder, A., Allison, S.D. & Treseder, K.K. (2010).
504 Functional diversity in resource use by fungi. *Ecology*, 91, 2324-2332.
505 30.
506 McGuire, K.L., Fierer, N., Bateman, C., Treseder, K.K. & Turner, B.L. (2012). Fungal
507 community composition in Neotropical rain forests: the influence of tree diversity and
508 precipitation. *Microb. Ecol.*, 63, 804-812.
509 31.
510 McGuire, K.L., Payne, S.G., Palmer, M.I., Gillikin, C.M., Keefe, D., Kim, S.J. *et al.* (2013).
511 Digging the New York City skyline: soil fungal communities in green roofs and city
512 parks. *PLoS One*, 8, e58020.
513 32.
514 Paulson, J.N., Stine, O.C., Bravo, H.C. & Pop, M. (2013). Differential abundance analysis for
515 microbial marker-gene surveys. *Nat. Methods*, 10, 1200-1202.
516 33.
517 Pera, L.M. & Callieri, D.A. (1997). Influence of calcium on fungal growth, hyphal morphology
518 and citric acid production in Aspergillus niger. *Folia Microbiologica*, 42, 551-556.
519 34.
520 Pina, R.G. & Cervantes, C. (1996). Microbial interactions with aluminium. *Biometals*, 9, 311-
521 316.
522 35.

Prescott, C.E. & Grayston, S.J. (2013). Tree species influence on microbial communities in litter and soil: Current knowledge and research needs. *Forest Ecology and Management*, 309, 19-27.

36.

Price, M.N., Dehal, P.S. & Arkin, A.P. (2010). FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS One*, 5, e9490.

37.

Prober, S.M., Leff, J.W., Bates, S.T., Borer, E.T., Firn, J., Harpole, W.S. *et al.* (2015). Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecology Letters*, 18, 85-95.

38.

Rohland, N. & Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research*, 22, 939-946.

39.

Talbot, J.M., Bruns, T.D., Taylor, J.W., Smith, D.P., Branco, S., Glassman, S.I. *et al.* (2014). Endemism and functional convergence across the North American soil mycobiome. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 6341-6346.

40.

Valverde-Barrantes, O.J., Smemo, K.A., Feinstein, L.M., Kershner, M.W. & Blackwood, C.B. (2015). Aggregated and complementary: symmetric proliferation, overyielding, and mass effects explain fine-root biomass in soil patches in a diverse temperate deciduous forest landscape. *New Phytologist*, 205, 731-742.

41.

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

42.

Uriarte, M., Thompson, J., Turner, B.L. & Zimmerman, J.K. (2015). Linking spatial patterns of leaf litterfall and soil nutrients in a tropical forest: a neighborhood approach. *Ecol Appl*, in press (http://dx.doi.org/10.1890/15-0112.1).

554 43.
555 Wang, Q., Garrity, G.M., Tiedje, J.M. & Cole, J.R. (2007). Naive Bayesian classifier for rapid
556 assignment of rRNA sequences into the new bacterial taxonomy. *Applied and*
557 *Environmental Microbiology*, 73, 5261-5267.
558 44.
559 Wardle, D.A., Bardgett, R.D., Klironomos, J.N., Setala, H., van der Putten, W.H., & Wall, D.H.
560 (2004). Ecological linkages between aboveground and belowground biota. *Science*, 304,
561 1629-1633.
562 45.
563 Waring, B.G., Averill, C. & Hawkes, C.V. (2013). Differences in fungal and bacterial
564 physiology alter soil carbon and nitrogen cycling: insights from meta-analysis and
565 theoretical models. *Ecology Letters*, 16, 887-894.
566 46.
567 Webb, C.O. (2000). Exploring the phylogenetic structure of ecological communities: an example
568 for rain forest trees. *American Naturalist*, 156, 145-155.
569 47.
570 Wolf, J.A., Hubbell, S.P., Fricker, G.A. & Turner, B.L. (2015). Geospatial observations on
571 tropical forest surface soil chemistry. *Ecology*, 96, 2313-2313.
572 48.
573 Wright, S.J., Kitajima, K., Kraft, N.J.B., Reich, P.B., Wright, I.J., Bunker, D.E. *et al.* (2010).
574 Functional traits and the growth-mortality trade-off in tropical trees. *Ecology*, 91, 3664-
575 3674.
576 49.
577 Yavitt, J.B., Harms, K.E., Garcia, M.N., Wright, S.J., He, F. & Mirabello, M.J. (2009). Spatial
578 heterogeneity of soil chemical properties in a lowland tropical moist forest, Panama.
579 *Australian Journal of Soil Research*, 47, 674-687.
580 **FIGURES**
581 **Figure 1.** Sampling sites, soil pH, and maps of community similarity (NMDS axis
582 scores). **(A)** Sampling sites for 625 soil cores (the 50 ha plot was divided into 20 x 20 m
583 quadrats). **(B)** Soil pH. **(C)** Soil bacterial communities. **(D)** Soil fungal communities. **(E)** Root

assemblages. **(F)** Tree assemblages. Interpolated values were calculated using inverse distance weighting with square-root distance.

Figure 2. Relationships between **(A)** soil bacterial communities and **(B)** soil fungal communities with tree assemblages at increasing spatial scales and root assemblages. Differences among assemblages and communities estimated using Bray-Curtis distances. Lines represent fitted linear regressions between tree assemblages and associated bacterial or fungal communities. Statistics correspond to rank-based partial Mantel tests controlling for differences in slope and soil pH with p-values corrected for multiple comparisons by False Discovery Rate (FDR). Note the difference in y-axis scales. Similar results were obtained using quantile regression analyses (Fig. S6).

Figure 3. Relationships between **(A)** soil bacterial communities and **(B)** soil fungal communities with tree assemblage distance at the 20 m neighborhood radius around each soil core based on taxonomic, phylogenetic and functional trait information. Only the species for which we had trait information were included in the analyses. Lines represent fitted linear regressions between assemblage distances and microbial community distances. Statistics correspond to rank-based partial Mantel tests controlling for differences in slope and soil pH with p-values corrected for multiple comparisons by False Discovery Rate (FDR). Dashed lines represent non-significant relationships. Note the difference in y-axis scales. Similar results were obtained using quantile regression analyses (Fig. S7).

SUPPLEMENTARY MATERIAL

Supplementary Figure 1. Sample-based species accumulation curves (mean and 95% confidence interval) for **(A)** plant assemblages and **(B)** soil microbial communities.

Supplementary Figure 2. Relationships between **(A)** soil bacterial community richness and **(B)** soil fungal community richness with tree assemblage richness at increasing spatial scales and richness of root assemblages. Note the difference in y-axis scales.

Supplementary Figure 3. Proportion of different taxonomic groups along the NMDS ordination axis for **(A)** bacterial phyla, **(B)** fungal phyla, **(C)** root plant orders, and **(D)** tree plant orders.

Supplementary Figure 4. Relationships between community dissimilarity and spatial distance for **(A)** soil bacterial communities ($r = 0.09$, $P < 0.001$) and **(B)** soil fungal communities ($r = 0.13$, $P < 0.001$). Note the difference in y-axis scales.

Supplementary Figure 5. Relationships among slope, soil pH, bacterial and fungal community similarity, and similarity in either root or tree assemblages. Similarity in communities and assemblages measured using the Bray-Curtis distance metric.

Supplementary Figure 6. Relationships between **(A)** soil bacterial communities and **(B)** soil fungal communities with tree assemblages at increasing spatial scales and root assemblages. Community and assemblage similarities measured using Bray-Curtis distances. Lines represent fitted quantile regressions between tree assemblages and associated bacterial or fungal communities. Note the difference in y-axis scales.

Supplementary Figure 7. Relationships between **(A)** soil bacterial communities and **(B)** soil fungal communities with tree assemblage distance at the 20 m neighborhood radius around each soil core based on taxonomic, phylogenetic and functional trait information. Only the tree species for which we had trait information were included in the analyses. Lines represent fitted quantile regressions between assemblage distance and microbial community distance. Note the difference in y-axis scales.

Supplementary Figure 8. Relationships between **(A)** soil bacterial communities and **(B)** soil fungal communities with difference in assemblage-weighted average plant traits. Only a few of the aggregated aboveground traits of the plant assemblages (wood density, fruit mass, seed mass, average DBH, growth rate, mortality rate, leaf area, leaf thickness, leaf mass per area, and leaf concentration of aluminum, calcium, potassium, magnesium, phosphorous, nitrogen and carbon) showed significant albeit weak relationships with the composition of the soil microbial communities. Lines represent fitted linear regressions. Statistics correspond to rank-based partial Mantel tests controlling for differences in slope and soil pH. In bold, p-values < 0.01 after correcting for multiple comparisons using False Discovery Rate (FDR). These relationships were assessed with the same subset of plant species (n = 93) for which complete trait and phylogenetic information was available. Note the difference in y-axis scales.