New Phytologist

Plant nutrient-acquisition strategies drive topsoil microbiome structure and function

Journal:	*New Phytologist*
Manuscript ID	NPH-MS-2020-32438.R2
Manuscript Type:	MS - Regular Manuscript
Date Submitted by the Author:	31-Mar-2020
Complete List of Authors:	Bahram, Mohammad; University of Tartu, Ecology and Earth Sciences Netherway, Tarquin; Swedish University of Agricultural Sciences, Department of Ecology Hildebrand, Falk; Quadram Institute Bioscience, Gut Microbes and Health Pritsch, Karin; Helmholz Zentrum München, Environmental Sciences Drenkhan, Rein; Estonian University of Life Sciences, Institute of Agricultural and Environmental Sciences Loit, Kaire; Estonian Academy of Sciences, Institute of Agricultural and Environmental Sciences Anslan, Sten; University of Tartu, Institute of Botany and Ecology; Tartu Ulikooli Okoloogia ja Maateaduste Instituut, Department of Botany Bork, Peer; European Molecular Biology Laboratory, Structural and Computational Biology Tedersoo, Leho; University of Tartu, Institute of Botany and Ecology;
Key Words:	Mycorrhizal symbiosis, Metagenomics, Functional guilds, Fungal-bacterial interactions, Microbial communities

Plant nutrient-acquisition strategies drive topsoil microbiome structure and function

Mohammad Bahram[1,2*], Tarquin Netherway[2], Falk Hildebrand[3,4], Karin Pritsch[5], Rein Drenkhan[6], Kaire Loit[7], Sten Anslan[8], Peer Bork[9,10], Leho Tedersoo[1,11]

[1]Department of Botany, Institute of Ecology and Earth Sciences, University of Tartu, 40 Lai St, 51005, Tartu, Estonia. [2]Department of Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 756 51 Uppsala, Sweden. [3]Gut Microbes and Health, Quadram Institute Bioscience, Norwich, Norfolk, UK. [4]Digital Biology, Earlham Institute, Norwich, Norfolk, UK. [5]Institute of Soil Ecology, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany. [6]Institute of Forestry and Rural Engineering, Estonian University of Life Sciences, Fr.R. Kreutzwaldi, 5, 51006 Tartu, Estonia. [7]Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Fr.R. Kreutzwaldi, 5, 51006 Tartu, Estonia. [8]Zoological Institute, Technische Universität Braunschweig, Mendelssohnstrasse 4, 38106 Braunschweig, Germany. [9]Structural and Computational Biology, European Molecular Biology Laboratory, Heidelberg, Germany. [10]Molecular Medicine Partnership Unit, University of Heidelberg and European Molecular Biology Laboratory, Heidelberg, Germany. [11]Natural History Museum, University of Tartu, 14a Ravila, 50411 Tartu, Estonia.

* Correspondence to*: Mohammad Bahram*

Email: *bahram@ut.ee*

Keywords

Biotic interactions, mycorrhizal type, metagenomics, plant-soil feedback, symbiosis.

Total word count: 5248 (Introduction: 870; Material & Methods: 2076; Results: 768; Discussion: 1538); **Number of references**: 75.

Summary

- Plant nutrient acquisition strategies drive soil processes and vegetation performance, but their effect on the soil microbiome remains poorly understood. This knowledge is important to predict the shifts in microbial diversity and functions to increasing changes in vegetation traits under global change.
- Here we documented the topsoil microbiomes of 145 boreal and temperate terrestrial sites in the Baltic region that broadly differed in vegetation type and nutritional traits, such as mycorrhizal types and symbiotic nitrogen-fixation.
- We found that sites dominated by arbuscular mycorrhizal (AM) vegetation harbor relatively more AM fungi, bacteria, fungal saprotrophs, and pathogens in the topsoil compared with sites dominated by ectomycorrhizal (EM) plants. These differences in microbiome composition reflect the rapid nutrient cycling and negative plant-soil feedback in AM soils. Lower fungal diversity and bacteria-to-fungi ratios in EM-dominated habitats are driven by monodominance of woody vegetation as well as soil acidification by EM fungi, which are associated with greater diversity and relative abundance of carbohydrate-active enzymes.
- Our study suggests that shifts in vegetation related to global change and land use may strongly alter the topsoil microbiome structure and function.

Introduction

Climate change poses an increasing threat to biodiversity and carbon (C) stores in terrestrial ecosystems by shifting vegetation types, aboveground foliar traits, and belowground nutrient acquisition strategies (Stocker *et al.*, 2016; Jo *et al.*, 2019). Human-induced shifts in land use and pollution affect soil nitrogen (N) availability and terrestrial C cycling (Stocker *et al.* 2016; Douglas *et al.* 2018). While shifts in N cycling are more localized and related to sources of N pollution and fertilisation, soil C losses induced by warming and elevated CO_2 are globally more uniform (Stocker *et al.* 2016; Jo *et al.*

2019). However, shifts in soil C cycling depend on the costs of acquiring N, and thus on nutrient acquisition strategies of plants such as mutualistic root associations with mycorrhizal fungi and N-fixing bacteria (Norby 1987; Terrer *et al.* 2018).

Nearly 90% of plant families have evolved root symbioses with mycorrhizal fungi that benefit their hosts through enhanced nutrient and water uptake, protection against pathogens, and environmental stress (Smith & Read 2010; Brundrett & Tedersoo 2018). Based on the anatomy and taxonomic identity of the phyto- and mycobionts, mycorrhizas are broadly classified into four types: arbuscular mycorrhiza (AM), ectomycorrhiza (EM), ericoid mycorrhiza, and orchid mycorrhiza (Brundrett & Tedersoo 2018). AM and EM plants dominate in most natural and anthropogenic ecosystems (Soudzilovskaia *et al.* 2019), and they differ in belowground C allocation, capacity of organic nutrient acquisition, and impact on soil C and nutrient cycling (Phillips *et al.* 2013; Tedersoo & Bahram 2019). For example, EM systems have evolved relatively higher N-acquisition efficiency from organic material to cope with slower decomposition processes and lower litter quality (Smith & Read 2010; Terrer *et al.* 2018; Tedersoo & Bahram 2019). Thus, depending on limiting nutrients and plant nutrient acquisition strategies, global change may hamper or support C sequestration, with further implications on nutrient cycling and climate change (Averill *et al.* 2014; Soudzilovskaia *et al.* 2018; Jo *et al.* 2019). Together with contrasting nutrient dynamics, EM and AM systems have contrasting patterns of plant-soil feedbacks influencing plant community dynamics. In general neutral or positive plant-soil feedbacks prevail in EM systems, compared to negative plant-soil feedbacks prevailing in AM systems, likely resulting from contrasting effects of mycorrhizal type on soil properties and the activity of various microbial functional groups, namely antagonists such as soil-borne pathogens (Bennett *et al.* 2017; Teste *et al.* 2017; Kadowaki *et al.* 2018). In addition to mycorrhizal symbionts, root associated bacteria from Rhizobiaceae and Frankiaceae families fix atmospheric N_2 and sustain plant N nutrition. Yet, despite such general mycorrhizal type and N-fixing effects on soil processes and plant community dynamics, knowledge about direct and indirect effects of

plant nutrient-acquisition strategies on the diversity and composition of free-living soil microbes and microbial functions relevant to C and N cycling is still lacking.

Free-living soil microorganisms both affect and respond to shifts in rhizosphere processes because of their integral roles in plant nutrition, cycling of organic material, and regulation of plant communities (Bardgett & Wardle 2010). Bacteria drive most soil N-cycling processes such as N-fixation, nitrification, and denitrification (Philippot *et al.* 2007; Reed *et al.* 2011). Saprotrophic Basidiomycota have evolved efficient mechanisms for degrading lignin and soil organic complexes, whereas other microfungi and bacteria decompose less recalcitrant polymers such as cellulose, hemicellulose, and chitin but also by-products of the lignin degradation pathway. The relative importance of bacteria and fungi in decomposition processes depend on soil acidity and C/N ratio (Waring *et al.* 2013; Bahram *et al.* 2018), which are strongly driven by the dominant vegetation via litter input (Waring *et al.* 2015) and potentially by mycorrhizal type (Lin *et al.* 2017). Due to high plant C allocation to EM fungi and the use of similar substrates, EM fungi may outcompete saprotrophic fungi for soil organic N sources, which may result in hampered degradation activity termed the Gadgil effect (Gadgil & Gadgil 1975; Fernandez & Kennedy 2016). Soil-borne pathogens have a direct negative effect on plant performance and thus can drive patterns of plant diversity (Bardgett & Wardle 2010). EM fungi compared to AM fungi offer greater physical protection against soil-borne pathogens, but may also maintain unfavorable conditions for disease progression such as high C/N ratios and low pH (Tedersoo & Bahram 2019).

Here we performed a regional-scale investigation of topsoil microbes and their potential gene functions in 145 sites with contrasting edaphic conditions (Fig. S1), dominated either by AM or EM vegetation and/or N-fixing plants, to estimate the effects of these nutrient acquisition strategies on the free-living topsoil microbiome and its potential gene functions. We used metabarcoding and shotgun metagenomics techniques for

identification of taxa and gene functions, respectively. We tested a hypothesis that temperate/boreal EM, AM, and N-fixing dominated systems have contrasting soil microbial community compositions and potential functions, driven by direct interactions and indirectly by altering the soil conditions of C/N ratio and pH. More specifically we predicted that with EM vegetation dominance EM fungi become the dominant soil organism group, favouring fungi to bacteria driven processes, as reflected in lower bacteria-to-fungi ratios in EM compared to AM systems. We further hypothesized that EM fungi suppress soil-borne pathogens and saprotrophs in EM systems, with negative consequences for plant community diversity and turnover (Mariotte *et al.* 2018),

Materials and Methods

Sampling

We performed sampling and documented vegetation parameters (including plants identity, functional type and basal area) in 145 plots (5800 subsamples) representing various soil types and various plant nutrient acquisition strategies across the northern Baltic region in Estonia and Latvia (Table S2). The sites were selected to represent various vegetation types, contrasting soil conditions and relative abundance of individual EM, AM, and N-fixing plant species but keeping climatic variation minimal (Fig. S1). These sites included 73, 16, 50 and six 2500-m^2 plots in natural forests dominated by EM deciduous woody plants (altogether 16 species), EM coniferous trees (Pinaceae; 3 species), AM deciduous woody plants (48 species), and AM coniferous trees (Cupressaceae; 2 species), respectively. Altogether 12 EM plots (*Alnus* incana, *A. glutinosa*) and four AM plots (*Caragana arborescens, Hippophae rhamnoides*) were dominated by N-fixing plants. Grasslands and fields comprising no woody plants included 6 and 8 plots, respectively. To determine plant composition, we recorded the relative abundance of plant species based on basal area (for tree) or relative cover (for grasses and crops) in each plot. We used a broad classification of AM and EM types, as

dual mycorrhizas as well as other mycorrhizal types such as ericoid mycorrhiza were underrepresented in our study. Time since last fire, vegetation age, and the proportion of each woody plant species (basal area basis) were also recorded. We downloaded climate data from the WorldClim database (www.worldclim.org). Samples were taken mostly during growing seasons from 2011 to 2016 (Table S1). Mean annual temperature and precipitation for all samples ranges from 4.7-7.0 °C and 549-745 mm, respectively. We excluded climatic variables from statistical analyses, because of little climate variation across the 400x400 km study area and little climate effect on microbes in our study based on preliminary analyses. In each site, 40 soil subsamples (5 cm diam. to 5 cm depth – no distinction was made between organic and mineral layers) were collected, followed by air-drying within 12 h of collection and homogenization. Total C, ^{13}C, total N and ^{15}N content were measured using an elemental analyzer (Eurovector, Milan, Italy) and an isotope ratio mass spectrometer (MAT 253; Thermo Electron, Bremen, Germany), following Tedersoo *et al.* (2012). Total P and K concentration were measured in ammonium lactate (Tecator ASTN 9/84; AOAC 1990). Concentrations of Ca and Mg were measured in 1 M ammonium acetate (Tecator ASTN 90/92; Page *et al.* 1982).

Molecular analysis

To determine the taxonomic and functional gene composition of soil, we used an amplicon-based approach for soil prokaryotes and eukaryotes as well as a shotgun metagenomics approach for gene-encoded functions. Total DNA was extracted from 2.0 g sample material (soil and fine root powder from homogenized soil samples) using the PowerMax Soil DNA Isolation Mini kit (MoBio) following the manufacturer's instructions (Tedersoo et al. 2014). Total DNA concentration was used as a proxy for microbial biomass, though it may also represent microbial necromass (Torti *et al.* 2015). The entire soil metagenome was sequenced from 5 μg extracted DNA that was sonicated to fragments of 300-400 bases and further ligated to adaptors using the TruSeq Nano DNA HT Library Prep Kit (Illumina Inc., San Diego, CA, USA). DNA libraries were sequenced on three runs in Illumina HiSeq 2500 platform (2 × 250 bp paired-end chemistry, rapid run mode). The DNA samples were subjected to metabarcoding of

Bacteria and Archaea using primers 515FB and 926R to target the ribosomal rRNA 16S gene V4 region (Walters *et al.* 2016). All eukaryotes including fungi were targeted based on partial 18S rRNA gene (V9 subregion) and full-length Internal Transcribed Spacer (ITS) region using the primers ITS9MUNngs and ITS4ngsUni (Tedersoo & Lindahl 2016). Samples were amplified using one of 115 primers pairs, in which both primers were tagged with a unique 10-base Golay index (at least four differences to each other; all starting with A; AT/GC ratio between 0.4 and 0.6). The 25 µl PCR mix consisted of 16 µl sterilized H_2O, 5 µl 5× HOT FIREPol Blend MasterMix polymerase (Solis Biodyne, Tartu, Estonia), 0.5 µl each primer (200 nM) and 3 µl DNA extract. Thermal cycling conditions were the following: 95 ˚C for 15 min, 30 cycles of 95 ˚C for 30 s, 50 ˚C 45 s and 72 ˚C for 1 min, with a final extension step at 72 ˚C for 10 min. PCR products of two technical replicates were pooled and their relative quantity was estimated by running 2 µl DNA on 1% agarose gel for 15 min. DNA samples producing no visible band or an overly strong band were re- amplified using 35 and 25 cycles, respectively. The amplicons were pooled and purified using FavorPrep™ Gel/PCR Purification Kit (Favorgen, Vienna, Austria) and shipped for library preparation in sequencing service providers' laboratories in the Estonian Biocenter, University of Tartu (Illumina MiSeq and HiSeq platforms) and Oslo University Sequencing Center (PacBio). For fungal identification, we used PacBio platform that outperforms other platforms in terms of distinguishing species, taxonomic precision and (lower) proportion of artefacts (Tedersoo et al. 2018). Negative and positive controls were used throughout the experiment including sequencing. All metagenomics and metabarcoding sequences have been deposited in the European Bioinformatics Institute Sequence Read Archive database: PRJEB24121 (ERP105926); 16S and ITS metabarcoding data of global soil samples, accession numbers PRJNA598043.

Bioinformatics analysis

Metagenomics

Metagenomics analysis followed the procedure described in (Bahram *et al.* 2018). Briefly, reads were quality-filtered by removing those with <70% max length of the run (i.e. 150bp), an accumulated error >2 or an estimated accumulated error >2.5, with a probability of ≥0.01 and >1 ambiguous position. Reads were trimmed if a quality window of 15 bases dropped <20 at the 3' end. This resulted in 980,810,505 reads. Quality-filtered reads were merged using FLASH (Magoč & Salzberg 2011) and mapped against the reference databases eggnog (Huerta-Cepas *et al.* 2015) and custom modification of CAZY (using DIAMOND in blastx mode; parameters k=5, $e=e^{-4}$) to quantify the abundance of functional genes. The scores of two unmerged query reads that mapped to the same target were combined to avoid double counting reads. To combine DIAMOND hit scores on target proteins, we summed the hit score of the forward and reverse reads matching a given target, using custom Perl scripts. To calculate the corresponding e-values independently from sequence as well as database parameters, we selected the lower e-value of either forward and reverse reads. We note that this approach reflects the potential function of organisms, which are not necessarily expressed at the transcript and enzyme level. Compared to bacterial and fungal pathogens, forest soil fungi are underrepresented in genome databases, which will underpower assignment of functions to specific fungal guilds.

Metabarcoding

The LotuS 1.462 pipeline was used for 16S amplicon sequence processing (Hildebrand *et al.* 2014). Reads were demultiplexed with modified quality-filtering procedure to trim reads to 170 bp and rejecting substandard reads: accumulated error ≤2; presence of unique reads >7 times in one, >3 times in two or >2 times in three samples. In total 22,335,463 of 31,286,576 reads passed the quality control and these were clustered with UPARSE (Edgar 2013) at 97% sequence similarity. Chimeric OTUs were detected and removed based on both reference-based and *de novo* chimera checking, using the RDP reference (http://drive5.com/uchime/rdp_gold.fa) in UCHIME (Edgar *et al.* 2011). Since our intention was to focus on the most abundant taxa, low-abundance OTUs with <4

sequences were removed by UPARSE, as implemented in LotuS by default, to minimize artefactual taxa in our data sets. The bacterial OTU abundance matrix was filtered from sequences of eukaryotic and chloroplastic origin, and rarefied to the lowest number of shared OTUs to remove the effect of sequencing depth across samples. Bacterial OTUs were assigned into different functional groups using Faprotax (Louca et al., 2016).

PacBio amplicon sequences were processed using PipeCraft (Anslan *et al.* 2017), resulting in 394,067 quality-filtered reads. Clustering at 97% sequence similarity was used for calculating OTUs. Representative sequences of OTUs were blastN-queried against the UNITE 7.1 reference data set. Taxonomic assignments were performed at 70%, 75%, 80%, 85% and 90% sequence similarity to roughly match phylum, class, order, family and genus level, respectively. Taxa with sequence similarity <70% to any taxon or match e-value $>e^{-50}$ were considered unidentified at the kingdom level. Fungal taxa were functionally assigned to principal guilds (Nguyen *et al.* 2016; Tedersoo & Smith 2017).

Data analysis

OTU abundance matrices were rarefied once to an equal number of reads (20,000 for bacteria and 500 for fungi) per sample to reduce the effect of variation in terms of sequenced reads using the function rrarefy in *vegan* package (Oksanen *et al.* 2007) of R (R-Core-Team). Alternatively, residuals of models with square root of total read abundance were used for analyzing richness (Tedersoo *et al.* 2014). The results were very similar between two approaches, and thus we only report the results of the latter approach (McMurdie & Holmes 2014). To deal with compositionality of abundance matrices (Gloor *et al.* 2017), we transformed the abundance-based compositional dataset by using centered log ratio transformation (CLR) in *mixOmics* package (Rohart *et al.* 2017). Five samples were removed from the analyses, because there were either from anoxic habitats

and hence outliers in community analyses, or contained a few bacterial reads and were therefore excluded from both the fungal and bacterial dataset.

The B/F ratio was calculated based on the proportion of bacterial to fungal metagenomic rRNA genes, as shown previously to correlate strongly to phospholipid-derived fatty acids PLFA based B/F ratio (Bahram *et al.* 2018; Fig. 2a). To include plant community composition in our univariate analysis, we used the first two axes from a principal coordinates analysis (PCoA) as implemented in the *ape* package (Rohart *et al.* 2017) (Fig. S2). For the PCoA analysis, a Bray-Curtis distance matrix for the plant community was generated based on Hellinger transformed abundance data in *vegan*. Plant diversity was calculated based on Shannon diversity index in *vegan*.

For univariate analysis, the best predictors of microbial richness and relative abundances were identified and included in a final model selection procedure using a machine learning approach as implemented in the *randomForest* function of *Randomforest* package (Liaw & Wiener 2002). This approach estimates variable importance while training the random forest (Breiman 2001). For this analysis, we included 25 biotic and abiotic variables, including vegetation and soil parameters as well as latitude, longitude and altitude (Table S1). We tested spatial autocorrelation in our data in Random Forest analysis by including spatial distance (Principal Coordinates of Neighbourhood Matrix (PCNM) vectors; Borcard & Legendre 2002), generated based on a matrix of geographic distances among samples in *vegan*. To further test direct and indirect effects of variables, we built Structural equation modelling (SEM) models in the AMOS software (SPSS) by including predictors based on their importance in the Random Forest models. In a prior model, all indirect and direct links between variables were established based on their correlations. Then, we removed non-significant links and variables or created new links between error terms until a significant model fit was achieved. Differences between the relative abundance of the main taxonomic and functional groups across mycorrhizal types

were tested using a non-parametric Wilcoxon rank-sum test, with Benjamini-Hochberg multiple testing correction. To model these based on EM vegetation cover (Fig. 1), we used beta regression for proportions (Ferrari & Cribari-Neto 2004) as implemented in the *betareg* package (Zeileis *et al.* 2016) of R.

Multivariate models were constructed in permutational multivariate analysis of variance (PERMANOVA) (Anderson 2005) with Adonis function of vegan (using 10^3 permutations), following variable selection in forward selection mode based on F_{pseudo}-values. PERMANOVA was performed to test discrimination of the relative abundance of different bacterial phyla and functional categories across mycorrhizal types. We further visualized taxonomic (operational taxonomic unit; OTU) and functional (orthologous gene; OG) composition of bacteria using global nonmetric multidimensional scaling (GNMDS) in vegan package based on the following options: two dimensions, initial configurations = 100, maximum iterations = 200, and minimum stress improvement in each iteration =10^{-7}. For constructing OG and OTU distance matrices, the Bray-Curtis dissimilarity was calculated between each pair of samples. Spatial autocorrelation as well as correlation in composition of different organism groups was calculated using Mantel test. Furthermore, to determine the relative importance of soil, vegetation and spatial variables (PCNMs) in shaping the composition of microbial taxa and functions, variation partitioning analysis was used as implemented in *varpart* function of *vegan*. To infer direct and indirect effects of mycorrhizal type at the multivariate level, we used the two first axes from a PCoA analysis of plant communities as explanatory or response variables in SEM.

Results

Microbial richness and abundance

Our analyses revealed 10,325 fungal Operational Taxonomic Units (OTUs) based on 304,248 reads and 29,813 bacterial OTUs (7,022,893 reads) as well as 74,298 bacterial (521,747,093 reads). Among abiotic variables, soil pH and $\delta^{15}N$ (an integrator of the N cycle) showed the strongest correlation with the EM/AM plant abundance ratio (based on basal area). Soil C/N ratio was strongly correlated with the proportion of coniferous EM trees (Fig. S3). Plant diversity correlated to coniferous EM abundance ratio ($r=0.423$, $p<0.001$) but not to the EM/AM plant abundance ratio ($p>0.05$). Total microbial biomass was positively associated with tree species richness ($r=0.471$, $p<0.001$) and the proportion of deciduous EM trees ($r=0.203$, $p=0.023$) but negatively with the proportion of coniferous trees (EM: $r=-0.240$, $p=0.007$; AM: $r=-0.341$, $p<0.001$, Fig. S3). Microbial biomass ($p>0.1$) was not different but bacteria-to-fungi abundance ratio (B/F ratio; $R^2_{adj}=0.197$, $p<0.001$) was relatively lower in EM ecosystems, especially in coniferous sites (Fig. 1). Random Forest analysis revealed that plant diversity, functional and mycorrhizal traits were the major determinants of absolute microbial biomass, whereas the B/F ratio was positively related to soil $\delta^{15}N$ and pH (Fig. 2).

Bacterial taxonomic richness was most strongly determined by soil pH (unimodal association; peak at pH 5-6; $R^2_{adj}=0.455$, $p<0.001$; Fig. 2), with the effect of vegetation traits remaining of secondary importance ($R^2_{adj}=0.027$; Fig. 2). Nevertheless, bacterial richness was highest in habitats dominated by N-fixing plants independently of soil pH ($R^2_{adj}=0.076$, $p= 0.0005$) and lowest in coniferous EM-dominated ecosystems ($R^2_{adj}=0.218$, $p<0.001$; Figs. 2, S4). Similarly to fungi, plant pathogenic bacteria were relatively more abundant in AM-dominated plots (0.358, $p<0.001$; Fig. S5). Nevertheless, N-fixing bacteria were more abundant in EM habitats ($r=0.527$, $p<0.001$), where soil N is mostly bound in organic material. N-fixing plants had no effect on relative abundance of N-fixing or denitrifier soil bacteria ($p>0.1$). SEM models indicated that EM relative abundance (basal area basis) had a direct negative effect but additional indirect negative effects on fungal richness through enhancing soil C/N ratio (Fig. S4). Of fungal functional groups, the relative abundance and richness of all guilds (except EM fungi)

declined with increasing EM dominance (Fig. 1). When excluding mycorrhizal fungi (both AM and EM) from the analysis, differences in these groups became weaker but remained significant (Fig. S6). As most fungal pathogens in current databases (e.g. UNITE) mostly comprise agricultural plant pathogens, we excluded croplands from these analyses, but the strong negative correlation between pathogens and EM/AM ratio remained ($r=-0.626$, $p<0.001$), indicating that the observed pattern for plant pathogens (Fig. 1) was not solely driven by this potential bias. Spatial vectors had little effect compared with other variables on microbial richness and relative abundance (Fig. S7).

Microbial composition

Bacterial community composition was also strongly driven by soil pH (Permanova: $F_{1,143}=61.8$, $R^2_{adj}=0.302$, $p=0.001$; Figs. 2,S8) but not by mycorrhizal type or other environmental variables. In contrast, fungal composition was mainly determined by mycorrhizal type both directly (Permanova: $F_{1,143}=7.37$; $R^2_{adj}=0.049$, $p=0.001$; Fig. S4) and indirectly via reduced soil pH in EM habitats (Figs. S3, S4). Except for Umbelopsidomycetes (Mucoromycota molds) and Agaricomycetes (Basidiomycota), relative abundance of most fungal higher-level taxa increased with AM tree dominance (Fig. 3). SEM analyses suggested that EM fungal relative abundance and richness drive those of bacteria and saprotrophs (Figs. 2,S3). Mantel test and variation partitioning analysis revealed weak spatial autocorrelation in our data. (Fig. S9).

Microbial functions

Plant nutrient-acquisition strategies affected the relative abundance and composition of microbial functional genes. Bacterial and fungal carbohydrate-active enzyme (CAZyme) profiles showed remarkable differences among habitats with different mycorrhizal types for both bacteria ($F_{1,143}=23.6$, $R^2_{adj}=0.116$, $p=0.001$) and fungi ($F_{1,143}=16.2$, $R^2_{adj}=0.090$, $p=0.001$). These differences among mycorrhizal types were particularly pronounced in

coniferous EM plots compared with any AM plots (Figs. 3,S4). In spite of reduced taxonomic richness, the diversity of bacterial (Mantel r=0.445, p=0.001); and fungal (r= 0.586, p<0.001) CAZyme genes increased significantly with increasing EM dominance (Figs 2A, S10). Bacterial functional gene composition was mainly determined by soil pH ($F_{1,143}$= 61.81, R^2_{adj}=0.302, p=0.001) and dominance of conifers ($F_{1,143}$=23.65, R^2_{adj}=0.141, p=0.001), with lower importance of mycorrhizal type ($F_{1,143}$= 13.77, R^2_{adj}=0.088, p=0.001). Bacterial OGs related to *inorganic ion transport and metabolism* were relatively more abundant in EM-dominated plots (R^2_{adj}=0.276, p<0.001, Fig. 3). Unlike in bacterial OGs (p>0.1), plant mycorrhizal type was the strongest determinant of fungal OG composition ($F_{1,143}$=20.2, R^2_{adj}=0.124, p=0.001; Figs. 2, S5), with $\delta^{15}N$ ($F_{1,143}$=27.3, R^2_{adj}=0.165, p=0.001) as an important edaphic predictor. Fungal OG richness was significantly higher in EM- than AM-dominated plots (R^2_{adj}= 0.201, p<0.001). Of the main fungal OG categories, *inorganic ion transport and metabolism* significantly increased with AM dominance, whereas *replication and recombination* increased with EM dominance (Fig. 3).

Discussion

Microbiome diversity

Our study demonstrates that the composition of microbial taxa and functional genes consistently differs among ecosystems dominated by plants with different nutrient-acquisition strategies, especially mycorrhizal type. Plant-soil feedbacks involving soil-inhabiting microorganisms contribute to the substantial differences in ecosystem processes such as soil C and nutrient cycling among forest stands dominated by different tree species (Waring *et al.* 2015) and mycorrhizal types (Phillips *et al.* 2013; Tedersoo & Bahram 2019). In particular, we found that non-fungal eukaryotes and saprotrophic fungi are more enriched in AM habitats. This is in line with higher decomposition rates in AM ecosystems (Tedersoo & Bahram 2019) and implies strong competitive interactions of EM fungi with free-living saprotrophs (Bödeker *et al.* 2016), bacteria and potentially

other soil microbes. Of decomposer organisms, only saprotrophic Agaricomycetes were relatively more common in EM forests, particularly EM coniferous ecosystems, compared with AM-dominated habitats (Figs. 1, S6), which reflects the production of low-quality litter by conifers (Cornelissen *et al.* 2001).

EM plants generally accumulate recalcitrant litter with a high C/N ratio (Lin *et al.* 2017; Sasse *et al.* 2017), which has strong direct effects on the soil microbiome (Bahram *et al.* 2018). Our model suggests that coniferous EM plants are the major drivers of an increase in soil C/N ratio and decrease in pH, which in turn reduce the richness of bacteria, archaea, several protist groups and the B/F ratio (Fig. S3). The strong effect of pH on soil bacterial richness corroborates previous reports and validates our analyses (Rousk *et al.* 2010; Bahram *et al.* 2018). Due to differences in physiology, capacity to withstand H^+ stress and nutrient stoichiometry (Rousk et al. 2010), bacterial decomposition pathways dominate over those of fungi in soils with high pH and low C/N ratio (Waring *et al.* 2013; Bahram *et al.* 2018).

While bacterial functional and taxonomic composition was relatively insensitive to factors other than pH, both functional and taxonomic diversity of fungi responded strongest to mycorrhizal type and plant diversity. Biotrophic fungal guilds such as plant pathogens, AM and EM mutualists are intimately associated with living plants and thus exhibited stronger and more specific plant interactions compared with most bacteria. Our analyses also indicate that the effect of mycorrhizal type on soil communities may strongly depend on other vegetation parameters as reflected by large differences between deciduous and coniferous EM-dominated ecosystems.

Functional genes

The distribution of overall functional gene categories was similar in EM and AM ecosystems, except ligninolytic CAZymes that were more abundant and family-rich in EM habitats, which we attribute to the accumulation of recalcitrant litter and high abundance of saprotrophic Agaricomycetes with efficient ligninolytic weaponry (Nagy *et al.* 2017). Although other OGs were generally represented at similar abundance, EM and AM ecosystems differed substantially in gene composition of bacterial *vs.* fungal origin. Most gene families are fundamental to signaling, cellular growth and both metabolic and anabolic processes that are equally required for functioning by various functional guilds of the microbiome. Conversely, bacterial and fungal degradation and nutrient cycling pathways exhibit great differences in emission of gases (NH_4, NO_3, CH_4) and intracellular *vs.* extracellular biopolymer degradation (Bugg *et al.* 2011; Frey-Klett *et al.* 2011).

The relatively greater B/F ratio in AM habitats explains the higher proportion of bacterial genes related to nutrient and C cycling, which is consistent with the more dynamic and leaky nutrient cycling in AM ecosystems (Tedersoo & Bahram 2019). Although our metagenomics data cannot be used to infer production and efficiency of specific enzymes, we advocate that bacterial and fungal pathways of the seemingly redundant functions may differentially affect soil nutrient fluxes and C cycling (Caspi *et al.* 2011). *In situ* measurements of soil processes integrated with proteome and transcriptome analyses will provide deeper insights into quantitative functional differences among plant nutrient acquisition strategies (Tedersoo & Bahram 2019). Overall, the relative abundance and diversity of N-fixing plants and N-fixing bacteria showed weak association with the diversity and composition of bacterial and fungal gene functions, indicating the more prominent role of mycorrhizal types and other vegetation parameters in driving these patterns.

Pathogens and plant community dynamics

According to the Janzen-Connell hypothesis, density-dependent accumulation of plant species-specific antagonists regulates species abundance and promotes diversity (Mariotte *et al.* 2018; Bagchi *et al.* 2014). Yet, EM fungi may interfere with the general Janzen-Connell model (Dickie *et al.* 2005; Chen *et al.* 2019), as reflected in often conspecific monodominant EM-dominated systems compared to more species rich mixed AM-dominated systems. EM associations could counteract negative-density dependent mechanisms through positive plant-soil feedbacks that favour the aggregation of conspecific individuals rather than a diverse plant community (Bever *et al.* 2010; Peh *et al.* 2011; Johnson *et al.* 2018). In line with this, recent experimental studies indicate that AM trees experience greater antagonisms from their associated soil microbiota compared with EM trees (Bennett *et al.* 2017; Teste *et al.* 2017; Kadowaki *et al.* 2018). Our metabarcoding results complement the idea of pathogen protection and suppression as a key mechanism driving positive plant-soil feedbacks in EM systems, demonstrating that putative plant pathogens are on average 2.6-fold more abundant in AM-dominated ecosystems, especially in AM deciduous forests (Fig. 1). However, we did not find support for the effect of this mechanism on plant community and diversity, as EM basal area showed rather weak, or even positive correlation in the case of coniferous EM, to plant diversity. Other mechanisms may also be important in promoting positive EM plant-soil feedbacks, such as extensive common EM mycelial networks redistributing nutrients and promoting EM seedlings (McGuire 2007; Kadowaki *et al.* 2018), and EM fungi trapping EM-dominated systems in a N-limitation feedback loop that reinforces the dominance and obligatory nature of the EM symbiosis (Franklin *et al.* 2014). Our results support the importance of the latter to some extent, specifically in coniferous EM dominated systems that were related to higher C/N and lower B/F ratios, soil $\delta^{15}N$, and pH (Fig. S3), as well as the greater relative abundance of EM fungi representing the dominant fungal functional guild (Fig. S4). Although our results need more empirical support, we suggest that the negative plant-soil feedback in AM systems compared to positive or neutral soil feedback in EM systems (Bennett *et al.* 2017; Teste *et al.* 2017; Kadowaki *et al.* 2018) may be attributed to four non-exclusive mechanisms: species-specific damage by pathogens (Mariotte *et al.* 2018), a relatively greater ability of EM

fungi to physically protect their hosts from the soil environment (Kadowaki *et al.* 2018), direct antagonistic effects of EM mycelium against antagonists and competitors, and the maintenance by EM fungi of unfavorable acidic and low-nutrient soil conditions for many microbial groups including pathogens (Figs. 1,S3). In mixed EM-AM forests, these contrasting mechanisms driving a collection of positive, neutral, and negative plant-soil feedbacks may generate complex microsites supplying regeneration niches for various species and promoting overall plant diversity (Mariotte *et al.* 2018).

Global implications

We demonstrated that mycorrhizal type may be one of the strongest predictors of soil microbiome diversity and functioning across contrasting soil and vegetation types at the regional scale, specifically in temperate and boreal ecosystems of the Baltic region. At large geographic scales, the effects of vegetation parameters on bacterial and fungal composition are less pronounced, perhaps due to the interplay of other predictors such as climatic variables and historical factors including dispersal limitation (Tedersoo *et al.* 2014; Maestre *et al.* 2015; Bahram *et al.* 2018; Delgado-Baquerizo *et al.* 2018). Modelling from regional to global scales demonstrates that climatic factors and land use play additional important roles in determining the distribution of mycorrhizal types and that EM vegetation may enhance soil C storage (Soudzilovskaia *et al.* 2019). Alternatively, litter decomposition potential was proposed as a key driver of mycorrhizal type distribution globally (Steidinger *et al.* 2019). However, for our studied temperate/boreal region our SEM models suggest that mycorrhizal type affects both relative and absolute abundances of soil saprotrophs and bacteria as well as genes related to decomposition, not *vice versa*. The directionality is important, because both mycorrhizal fungi and pathogens determine plant establishment success on a landscape scale, which further shapes the habitat for particular saprotrophic groups. The local influential variables such as soil and vegetation parameters and as well as climatic variables on larger scales have strong implications for the potential effects of global change on vegetation type, soil microbial diversity and the processes governed by these

(Soudzilovskaia *et al.* 2019). EM vegetation, which is patchily distributed in tropical ecosystems and limited by rainfall, may suffer strongest from extended drought periods and atmospheric pollution (Terrer *et al.* 2016; Tedersoo 2017; Jo *et al.* 2019).

To conclude, our results suggest that shifts in the balance between EM-AM vegetation may alter the soil microbiome structure and function in temperate and boreal ecosystems. However, it remains unclear to what extent these functional differences among plant nutrient-acquisition strategies can be extrapolated to arctic and tropical ecosystems, because differences among mycorrhizal types on soil chemistry and ecosystem processes are somewhat weaker at low latitudes (Keller & Phillips 2019; Soudzilovskaia *et al.* 2019). We certainly need controlled experiments to test the interactions of soil pH and temperature with mycorrhizal type effects on soil microbiome structure and functioning and to what extent these effects are bidirectional.

Acknowledgments

We thank Rasmus Puusepp and Emma Johanna Sova for help with molecular analysis and Franz Buegger for help with stable isotope analyses. We also thank four anonymous reviewers for providing constructive comments on an earlier version of the manuscript. This study was funded by Estonian Research Council grant (PUT1317) and Swedish Research Council (Vetenskapsrådet, grant no: 2017-05019).

Author Contributions

MB, RD, KL and LT contributed data. MB, TN, FH, SA and LT analysed the data. FH conducted bioinformatic analysis. KP contributed to Isotopic analysis. MB and LT wrote the manuscript with input from FH, PB and other authors.

References

Anderson MJ. **2005**. Permutational multivariate analysis of variance. *Department of Statistics, University of Auckland, Auckland* **26**: 32–46.

AOAC. 1990. Official Methods of Analysis, 15th edn. AOAC, Kensington.

Anslan S, Bahram M, Hiiesalu I, Tedersoo L. **2017**. PipeCraft: flexible open-source toolkit for bioinformatics analysis of custom high-throughput amplicon sequencing data. *Molecular ecology resources*.

Averill C, Turner BL, Finzi AC. **2014**. Mycorrhiza-mediated competition between plants and decomposers drives soil carbon storage. *Nature* **505**: 543.

Bagchi R, Gallery RE, Gripenberg S, Gurr SJ, Narayan L, Addis CE, Freckleton RP, OT L. **2014**. Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature* **506**: 85–88.

Bahram M, Hildebrand F, Forslund SK, Anderson JL, Soudzilovskaia NA, Bodegom PM, Bengtsson-Palme J, Anslan S, Coelho LP, Harend H, *et al.* 2018. Structure and function of the global topsoil microbiome. *Nature* **560**: 233–237.

Bardgett RD, Wardle DA. **2010**. *Aboveground-belowground linkages: biotic interactions, ecosystem processes, and global change*. Oxford University Press.

Bennett JA, Maherali H, Reinhart KO, Lekberg Y, Hart MM, Klironomos J. **2017**. Plant-soil feedbacks and mycorrhizal type influence temperate forest population dynamics. *Science* **355**: 181.

Bever JD, Dickie IA, Facelli E, Facelli JM, Klironomos J, Moora M, Rillig MC, Stock WD, Tibbett M, Zobel M. **2010**. Rooting theories of plant community ecology in microbial interactions. *Trends in ecology & evolution* **25**: 468–478.

Bödeker IT, Lindahl BD, Olson Åke, Clemmensen KE. **2016**. Mycorrhizal and saprotrophic fungal guilds compete for the same organic substrates but affect decomposition differently. *Functional Ecology* **30**: 1967–1978.

Borcard D, Legendre P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modeling* **153**: 51–68.

Breiman, L. (2001). Random forests. *Machine learning* **45**: 5–32.

Brundrett MC, Tedersoo L. **2018**. Evolutionary history of mycorrhizal symbioses and global host plant diversity. *New Phytologist* **220**: 1108–1115.

Bugg TD, Ahmad M, Hardiman EM, Rahmanpour R. **2011**. Pathways for degradation of lignin in bacteria and fungi. *Natural product reports* **28**: 1883–96.

Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA. **2011**. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic acids research* **40**: D742–D753.

Chen L, Swenson NG, Ji N, Mi X., Ren H, Guo L, *et al.* 2019. Differential soil fungus accumulation and density dependence of trees in a subtropical forest. *Science* **366**: 124–128.

Cornelissen J, Aerts R, Cerabolini B, Werger M. 2001. Carbon cycling traits of plant species are linked with mycorrhizal strategy. *Oecologia* **129**: 611.

Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett RD, Maestre FT, Singh BK, Fierer N. **2018**. A global atlas of the dominant bacteria found in soil. *Science* **359**: 320–325.

Dickie IA, Schnitzer SA, Reich PB, Hobbie SE. 2005. Spatially disjunct effects of co-occurring competition and facilitation. *Ecology Letters* **8**: 1191–1200.

584 **Douglas PM, Pagani M, Eglinton TI, Brenner M, Curtis JH, Breckenridge A,**
585 **Johnston K**. **2018**. A long-term decrease in the persistence of soil carbon caused by
586 ancient Maya land use. *Nature Geoscience* **11**: 9.

587 **Edgar RC**. **2013**. UPARSE: highly accurate OTU sequences from microbial amplicon
588 reads. *Nature Methods* **10**: 996–998.

589 **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R**. **2011**. UCHIME improves
590 sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200.

591 **Feng Y, Hu Y, Wu J, Chen J, Yrjälä K, Yu W. 2019.** Change in microbial
592 communities, soil enzyme and metabolic activity in a Torreya grandis plantation in
593 response to root rot disease. *Forest Ecology & Management* **432**: 932-941.

594 **Feng X, Uriarte M, González G, Reed S, Thompson J, Zimmerman JK, Murphy L..**
595 **2018.** Improving predictions of tropical forest response to climate change through
596 integration of field studies and ecosystem modeling. *Global change biology* **24**: e213-
597 e232.

598 **Fernandez CW, Kennedy PG**. **2016**. Revisiting the 'Gadgil effect': do interguild fungal
599 interactions control carbon cycling in forest soils? *New phytologist* **209**: 1382–1394.

600 **Franklin O, Näsholm T, Högberg P, Högberg MN. 2014.** Forests trapped in nitrogen
601 limitation–an ecological market perspective on ectomycorrhizal symbiosis. *New*
602 *Phytologist* **203**: 657-666.

603 **Frey-Klett P, Burlinson P, Deveau A, Barret M, Tarkka M, Sarniguet A**. **2011**.
604 Bacterial-fungal interactions: hyphens between agricultural, clinical, environmental, and
605 food microbiologists. *Microbiology & Molecular Biology Reviews* **75**: 583–609.

606 **Gadgil PD, RL G**. **1975**. *Suppression of litter decomposition by mycorrhizal roots of*
607 *Pinus radiata*. New Zealand Forest Service.

608 **Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. 2017.** Microbiome
609 datasets are compositional: and this is not optional. *Frontiers in Microbiology* **8**: 2224.

Hildebrand F, Tadeo R, Voigt AY, Bork P, Raes J. **2014**. LotuS: an efficient and user-friendly OTU processing pipeline. *Microbiome* **2**: 30.

Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, *et al.* 2015. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic acids research* **44**: D286-93.

Jo I, Fei S, Oswalt CM, Domke GM, RP P. **2019**. Shifts in dominant tree mycorrhizal associations in response to anthropogenic impacts. *Science Advances* **5**.

Johnson DJ, Clay K, Phillips RP. **2018**. Mycorrhizal associations and the spatial structure of an old-growth forest community. *Oecologia* **186**: 195–204.

Kadowaki K, Yamamoto S, Sato H, Tanabe AS, Hidaka A, Toju H. **2018**. Mycorrhizal fungi mediate the direction and strength of plant–soil feedbacks differently between arbuscular mycorrhizal and ectomycorrhizal communities. *Communications biology* **1**: 196.

Keller, A.B. & Phillips, R.P. (2019). Leaf litter decay rates differ between mycorrhizal groups in temperate, but not tropical, forests. *New Phytol.*, 222, 556–564.

Liao W, others. 2017. *Global climate change will increase the abundance of symbiotic nitrogen_fixing trees in much of North America* **23**: 4777–4787.

Liaw A, Wiener M. **2002**. Classification and regression by randomForest. *R news* **2**: 18–22.

Lin G, McCormack ML, Ma C, Guo D. **2017**. Similar below-ground carbon cycling dynamics but contrasting modes of nitrogen cycling between arbuscular mycorrhizal and ectomycorrhizal forests. *New Phytologist* **213**: 1440.

Maestre FT, Delgado-Baquerizo M, Jeffries TC, Eldridge DJ, Ochoa V, Gozalo B, Quero JL, García-Gómez M, Gallardo A, Ulrich W, *et al.* 2015. Increasing aridity

reduces soil microbial diversity and abundance in global drylands. *Proceedings of the National Academy of Sciences*: 201516684.

Magoč T, Salzberg SL. **2011**. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**: 2957–2963.

Mariotte P, Mehrabi Z, Bezemer TM, De Deyn GB, Kulmatiski A, Drigo B, Veen GC. Van der Heijden MG, Kardol P. 2018. *Plant soil feedback: bridging natural and agricultural sciences* **33**: 129–142.

McGuire KL. 2007. Common ectomycorrhizal networks may maintain monodominance in a tropical rain forest. *Ecology* **88**: 567-574.

McMurdie PJ, Holmes S. **2014**. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS computational biology* **10**: 4.

Nagy LG, Riley R, Bergmann PJ, Krizsan K, Martin FM, Grigoriev IV, Cullen D, DS H. **2017**. Genetic bases of fungal white rot wood decay predicted by phylogenomic analysis of correlated gene-phenotype evolution. *Molecular Biology and Evolution* **34**: 35–44.

Nguyen NH, Song Z, Bates ST, Branco S, Tedersoo L, Menke J, Schilling JS, Kennedy PG. **2016**. FUNGuild: an open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecology* **20**: 241–248.

Oksanen J, Kindt R, Legendre P, O'Hara B, Stevens MHH, Oksanen MJ, Suggests M. **2007**. The vegan package. *Community ecology package* **10**: 631–637.

Page AL, Miller RH, Keeney DR. 1982. *Methods of Soil Analysis. II. Chemical and Microbiological Properties*, 2nd edn. American Society of Agronomy, Madison.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

659 **Peh KS-H, Lewis SL, Lloyd J**. **2011**. Mechanisms of monodominance in diverse
660 tropical tree-dominated systems. *Journal of Ecology* **99**: 891–898.

661 **Philippot L, Hallin S, Schloter M**. **2007**. Ecology of denitrifying prokaryotes in
662 agricultural soil. *Advances in agronomy* **96**: 249–305.

663 **Phillips RP, Brzostek E, Midgley MG**. **2013**. The mycorrhizal-associated nutrient
664 economy: a new framework for predicting carbon–nutrient couplings in temperate
665 forests. *New Phytologist* **199**: 41–51.

666 **Reed SC, Cleveland CC, AR T**. **2011**. Functional ecology of free-living nitrogen
667 fixation: a contemporary perspective. *Annual review of ecology, evolution, and*
668 *systematics* **42**: 489–512.

669 **Reich P, Oleksyn J**. **2004**. Global patterns of plant leaf N and P in relation to
670 temperature and latitude. *Proceedings of the National Academy of Sciences USA* **101**:
671 11001.

672 **Norby RJ. 1987.** Nodulation and nitrogenase activity in nitrogen_fixing woody plants
673 stimulated by CO2 enrichment of the atmosphere. *Physiologia Plantarum* **71**: 77–82.

674 **Rohart F, Gautier B, Singh A. Lê Cao KA. 2017.** mixOmics: An R package for 'omics
675 feature selection and multiple data integration. *PLoS Computational Biology* **13**:
676 e1005752.

677 **Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG, Knight R,**
678 **Fierer N**. **2010**. Soil bacterial and fungal communities across a pH gradient in an arable
679 soil. *The ISME journal* **4**: 1340.

680 **Sasse J, Martinoia E, Northen T**. **2017**. Feed your friends: do plant exudates shape the
681 root microbiome? *Trends in Plant Science* **23**: 25–41.

682 **Smith SE, Read DJ**. **2010**. *Mycorrhizal symbiosis*. Academic press.

683 **Soudzilovskaia NA, van Bodegom PM, Terrer C, van't Zelfde M, McCallum I, Luke**
684 **McCormack M., *et al.*** 2019. Global mycorrhizal plant distribution linked to terrestrial
685 carbon stocks. *Nature Communications* **10**: 5077.

686 **Steidinger BS, Crowther TW, Liang J, Nuland MEV, Werner GDA, Reich PB,**
687 **Nabuurs GJ, de-Miguel S, Zhou M, Picard N, *et al.* 2019**. Climatic controls of
688 decomposition drive the global biogeography of forest-tree symbioses. *Nature* **569**: 404–
689 408.

690 **Stocker BD, Prentice IC, Cornell SE, Davies-Barnard T, Finzi AC, Franklin O,**
691 **Janssens I, Larmola T, Manzoni S, Näsholm T**. **2016**. Terrestrial nitrogen cycling in
692 Earth system models revisited. *New Phytologist* **210**: 1165–1168.

693 **Tedersoo L**. **2017**. Global biogeography and invasions of ectomycorrhizal plants: past,
694 present and future. In: Biogeography of Mycorrhizal Symbiosis. Springer, 469–531.

695 **Tedersoo L, Bahram M**. **2019**. Mycorrhizal types differ in ecophysiology and alter plant
696 nutrition and soil processes. *Biological Reviews*.

697 **Tedersoo L, Bahram M, Polme S, Koljalg U, Yorou NS, Wijesundera R, Ruiz LV,**
698 **Vasco-Palacios AM, Thu PQ, Suija A, *et al.* 2014**. Global diversity and geography of
699 soil fungi. *Science* **346**: 1256688–1256688.

700 **Tedersoo L, Lindahl B**. **2016**. Fungal identification biases in microbiome projects.
701 *Environmental microbiology reports* **8**: 774–779.

702 **Tedersoo L, Naadel T, Bahram M, Pritsch K, Buegger F, Leal M, Kõljalg U,**
703 **Põldmaa K**. **2012**. Enzymatic activities and stable isotope patterns of ectomycorrhizal
704 fungi in relation to phylogeny and exploration types in an afrotropical rain forest. *New*
705 *Phytologist* **195**: 832–843.

706 **Tedersoo L, Smith ME**. **2017**. Ectomycorrhizal fungal lineages: detection of four new
707 groups and notes on consistent recognition of ectomycorrhizal taxa in high-throughput
708 sequencing studies. In: Biogeography of mycorrhizal symbiosis. Springer, 125–142.

Terrer C, Vicca S, Hungate BA, Phillips RP, Prentice IC. **2016**. Mycorrhizal association as a primary control of the CO2 fertilization effect. *Science* **353**: 72–74.

Terrer C, Vicca S, Stocker BD, Hungate BA, Phillips RP, Reich PB, Finzi AC, IC P. **2018**. Ecosystem responses to elevated CO2 governed by plant soil interactions and the cost of nitrogen acquisition. *New Phytologist* **217**: 507.

Teste FP, Kardol P, Turner BL, Wardle DA, Zemunik G, Renton M, Laliberté E. **2017**. Plant-soil feedback and the maintenance of diversity in Mediterranean-climate shrublands. *Science* **355**: 173.

Torti A, Lever MA, Jørgensen BB. 2015. Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Marine Genomics*, 24, 185-196.

Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, Gilbert JA, Jansson JK, Caporaso JG, Fuhrman JA. **2016**. Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *Msystems* **1**: e00009–15.

Waring BG, Álvarez-Cansino L, Barry KE, Becklund KK, Dale S, Gei MG, Keller AB, Lopez OR, Markesteijn L, Mangan S, *et al.* 2015. Pervasive and strong effects of plants on soil chemistry: a meta-analysis of individual plant 'Zinke' effects. *Proceedings of the Royal Society B: Biological Sciences* **282**: 20151001.

Waring BG, Averill C, Hawkes CV. **2013**. Differences in fungal and bacterial physiology alter soil carbon and nitrogen cycling: insights from meta-analysis and theoretical models. *Ecology Letters* **16**: 887–894.

Zeileis A, Cribari-Neto F, Gruen B, Kosmidis I, Simas AB, Rocha AV, *et al.* 2016. Package 'betareg.' *R Package*.

Figures and Tables

Fig. 1. Functional guild composition of soil bacteria (% relative to fungi) and fungal functional guild relative abundance in relation to EM dominance. Data points show the relative abundance of bacteria and fungi (and fungal guilds) in each plot. EM dominance corresponds to the percentage of EM vegetation estimated on a basal-area basis. The relative abundances on the y-axes were scaled from 0 to 1 for better visualization.

Fig. 2. Microbial diversity and composition associate with biotic and abiotic factors. **A**) Random Forest heatmap indicates relationship of microbial taxa and functional groups to plant traits, edaphic and geographic variables. The size of circles corresponds to the variable importance (% of mean decrease accuracy estimated based on out-of-bag-CV); blue and red depict negative and positive Spearman correlations, respectively. Plant composition 1 & 2 are the first two PCA axes representing changes in the composition of plants across the plots. The top barplot shows the out-of-bag variance explained for each model with the dependent variables on the x-axis. **B**) Best-fitting structural equation model based on relationships retrieved in (A) for the relative abundance of bacteria and fungi. All relationships were significant ($P<0.05$) and model fits were acceptable according to Chi-Square test ($P>0.1$) and PCLOSE test ($p>0.1$). See Table S1 for statistical details. We tested both directions for the relationships between $\delta^{15}N$ or pH and the relative abundance of functional groups or Bacteria/Fungi ratio, and kept those that improved model fit (based on PCLOSE test).

Fig. 3. The distribution of microbial taxa and functional genes differ across habitats with different plant nutrient-acquisition strategies and the dominant vegetation type. The figure shows the relative abundance of major prokaryotic phyla (classes for Proteobacteria), eukaryotic phyla and functional gene categories in AM-dominant and EM dominant plots. Letters denote significant differences at the 0.05 probability level on the basis of Kruskal–Wallis tests corrected for multiple testing. Jittered points and bars represent individual relative abundances per sample, whereas bars represent the mean of

761 relative abundances per category. Values on the y-axes are the square root of relative
762 abundances.