Leff, Jonathan W.; Bardgett, Richard D.; Wilkinson, Anna; Jackson, Benjamin G.; Pritchard, William J.; De Long, Jonathan R.; Oakley, Simon; Mason, Kelly E.; Ostle, Nicholas J.; Johnson, David; Baggs, Elizabeth M.; Fierer, Noah.
2018. **Predicting the structure of soil communities from plant community taxonomy, phylogeny, and traits**. *ISME Journal*, 12 (7). 1794-1805.
https://doi.org/10.1038/s41396-018-0089-x

Title: Predicting the structure of soil communities from plant community taxonomy, phylogeny, and traits

Short running title: Associations between soil and plant communities

Authors: Jonathan W. Leff[1,2], Richard D. Bardgett[3], Anna Wilkinson[3], Benjamin G. Jackson[4], William J. Pritchard[3], Jonathan R. De Long[3], Simon Oakley[5], Kelly E. Mason[5], Nicholas J. Ostle[6], David Johnson[3], Elizabeth M. Baggs[7], and Noah Fierer[1,2*]

Affiliations:

[1]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA

[2]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO 80309, USA

[3]School of Earth and Environmental Sciences, Michael Smith Building, The University of Manchester, Oxford Road, Manchester M13 9PT, UK.

[4]School of Geosciences, Grant Institute, The King's Buildings, James Hutton Road, Edinburgh, EH9 3FE, UK

[5]Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster, LA1 4AP, UK

[6]Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK

[7]The Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush Campus, Midlothian, EH25 9RG Buildings, UK

***Corresponding author:**

Noah Fierer

University of Colorado

Cooperative Institute for Research in Environmental Sciences

UCB 216, CIRES Bldg. Rm. 318

Boulder, CO 80309-0216 USA

E-mail: Noah.Fierer@colorado.edu

Sources of financial support: The UK Biotechnology and Biological Sciences Research Council and the U.S. National Science Foundation.

Subject category: Microbial population and community ecology

Conflicts of interest: The authors declare no conflict of interest.

Abstract

There are numerous ways in which plants can influence the composition of soil communities. However, it remains unclear whether information on plant community attributes, including taxonomic, phylogenetic, or trait-based composition, can be used to predict the structure of soil communities. We tested, in both monocultures and field-grown mixed temperate grassland communities, whether plant attributes predict soil communities including taxonomic groups from across the tree of life (fungi, bacteria, protists, and metazoa). The composition of all soil community groups was affected by plant species identity, both in monocultures and in mixed communities. Moreover, plant community composition predicted additional variation in soil community composition beyond what could be predicted from soil abiotic characteristics. In addition, analysis of the field aboveground plant community composition and the composition of plant roots suggests that plant community attributes are better predictors of soil communities than root distributions. However, neither plant phylogeny nor plant traits were strong predictors of soil communities in either experiment. Our results demonstrate that grassland plant species form specific associations with soil community members and that information on plant species distributions can improve predictions of soil community composition. These results indicate that specific associations between plant species and complex soil communities are key determinants of biodiversity patterns in grassland soils.

Introduction

The interactions between plants and soil organisms can have important ramifications for ecosystem functioning and plant community dynamics, but the extent to which these interactions influence the spatial distributions of soil communities remains poorly understood. Knowing how plants control the spatial variation in belowground communities is important for building a predictive understanding of the heterogeneity in soil communities and contributing to pre-existing research that has identified how certain site and abiotic soil properties can influence the spatial variation in soil communities across large geographic scales (Fierer et al., 2009; Bates et al., 2013; Tedersoo et al., 2014; Kaiser et al., 2016). Further, this information will aid our ability to probe the undescribed and likely diverse ways in which soil organisms interact with plants since comparatively few plant-microbe interactions are well understood (Van der Putten et al., 2013).

Certain soil organisms are known to form close associations with particular plant species (Wardle et al., 2004; Bardgett and Wardle, 2010). Mycorrhizal relationships, for instance, involve a direct exchange of nutrients between plants and symbiotic soil fungi, and these relationships can influence plant-soil diversity linkages (van der Heijden et al., 1998; Hiiesalu et al., 2014). Indirect mechanisms, such as the release of root exudates and microbial attraction to those exudates, can also drive associations between specific microbes and plant species (Singh et al., 2004). However, these described interactions are likely only a small fraction of the numerous interactions among plants and soil organisms in a given ecosystem. Thus, it is uncertain whether the composition of soil communities as a whole is associated with plant community attributes under field conditions.

It has long been known that individual plant species can exert a powerful influence on

soil microbial communities (Grayston *et al.*, 1998; Berg and Smalla, 2009; Bardgett *et al.*, 1999), and there is evidence that divergence in soil bacterial and fungal communities is broadly linked to plant community composition at landscape (de Vries *et al.*, 2012; Grayston *et al.*, 2001) and global scales (Prober et al., 2015). Additionally, correlational analyses have revealed associations between individual plant species and soil fungal (Lekberg and Waller, 2016), bacterial (Berg, 2009), nematode (Bezemer et al., 2010), and arthropod (St. John et al., 2006) communities. However, it is unclear whether these relationships are driven by shared environmental preferences or by the direct effects of locally dominant plant species on soil communities. While plant invasions can elicit shifts in soil community structure (Hawkes *et al.*, 2005; Gibbons *et al.*, 2017), the effects of plant species identity on the overall composition of belowground communities are often weak or difficult to quantify, with several studies having failed to identify strong links between changes in plant assemblages and corresponding changes in soil communities (Porazinska *et al.*, 2003; Bezemer *et al.*, 2006; Tedersoo *et al.*, 2015; Lekberg and Waller, 2016; Carey *et al.*, 2015). As such, the existence of a general relationship between plants and soil communities remains uncertain and difficult to predict a priori.

There are multiple plant community attributes that could potentially be used to predict variation in soil communities. Plant species identity could be a strong predictor of variation in soil communities (Berg and Smalla, 2009; Bezemer et al., 2010; Lekberg and Waller, 2016), as could evolutionary history (i.e. the phylogeny) of plants, given the potential for more closely related plants to be associated with more similar belowground communities (Barberán et al., 2015b). Such patterns could arise as a product of coevolution between plants and soil microbes or if phylogenetic relatedness corresponds to other plant attributes that affect soil organisms (De Deyn and Van Der Putten, 2005). It has also been proposed that plant functional traits could be

104 used to predict plant-microbe associations a priori given that plant species' distributions and
105 community diversity are generally predictable based on their traits (Ben-Hur et al., 2012; Adler
106 et al., 2013), and soil communities can form associations with plants based on these traits
107 (Wardle et al., 2004). Although previous studies have shown that plant traits can explain
108 variation in soil microbial processes involved in C and N cycling (Orwin et al., 2010; Grigulis et
109 al., 2013; Cantarel et al., 2015; Moreau et al., 2015; Legay et al., 2016), it remains unclear
110 whether variation in soil community composition is directly caused by, or merely associated
111 with, differences in plant traits. Further, past studies show that links between plant traits and the
112 composition of soil communities are not always observed (Barberán et al., 2015b) and when they
113 have been found, they are often based on crude assessments of microbial community
114 composition, such as the relative abundance of fungi and bacteria (Orwin et al., 2010; de Vries et
115 al., 2012). Likewise, most previous work has focused on the relationships between soil biota and
116 aboveground plant traits, despite increasing evidence that root traits are likely to play a more
117 important role in structuring belowground communities (Bardgett et al., 2014; Legay et al., 2014;
118 Thion et al., 2016).

119 Here we provide the first in-depth evaluation of the predictive power of plant community
120 attributes, alongside abiotic factors, for explaining spatial (i.e. horizontal) variation in soil
121 communities at the individual plant and community-scale. While previous work has investigated
122 effects of plant species and community attributes on soil communities, we are not aware of any
123 previous study that has comprehensively assessed these effects across such a wide range of
124 functionally important belowground taxonomic groups. Specifically, we address the overarching
125 question: Can plant community attributes (i.e. taxonomic composition, phylogenetic
126 composition, and plant functional traits) be used to predict spatial variability in soil community

composition? To address this question, we sampled soils from both monocultures of 21 common temperate grassland plant species spanning eight families and a range of life history strategies, and we sampled an adjacent field experiment where grassland community composition had been manipulated through plant species additions to create a gradient of plant species and plant functional diversity. We used DNA sequencing-based approaches to target soil fungal, bacterial, protistan, and metazoan (faunal) communities. We first assessed whether the identity, phylogenetic history, and/or functional traits of individual plant species (both leaf and root traits) could be used to explain variation in soil communities. Next, we determined whether observations made at the individual plant scale correspond to similar trends in mixed plant communities in the field.

Materials and Methods

Mesocosms experiment

To evaluate effects of individual plant species, their phylogeny, and their functional traits on soil communities, mesocosms containing plants grown in monoculture were established in a fenced enclosure at Colt Park within the Ingleborough National Nature Reserve in England (54°11'38.7"N 2°20'54.4"W). Mesocosms were constructed from polypropylene pots (38 x 38 x 30 cm) filled with 10 cm of rinsed gravel and 20 cm sieved and homogenized top soil (pH ~5.8; 8.9 C%; 0.92 N%). Top soil was a brown earth sourced from the adjacent grassland, a mesotrophic temperate grassland under extensive agricultural management, which involved light grazing by sheep and cattle from autumn to spring, but no grazing during the growing season when an annual hay crop was taken, and an occasional light dressing of farmyard manure or mineral fertilizer (~25 kg ha^{-1} N) in early spring (De Deyn *et al.*, 2011). Twenty-one grassland

plant species (Fig. 1) were germinated and grown in a greenhouse from commercial seed (Emorsgate Seeds, Norfolk, PE34 4RT, UK) or from seed collected at the site. Mesocosms were planted and arranged in a randomized block design with four blocks. Plants were actively weeded and harvested annually. Plant biomass and soil was collected in July, approximately two years following planting, during the height of the growing season and before seed filling. Eight to 20 leaves from at least three individuals per mesocosm were clipped and stored in sealed plastic bags at 4 °C prior to processing. A representative 6.8 cm diameter soil core was taken from the complete soil column of each mesocosm, and soil subsamples were frozen and shipped on dry ice to the University of Colorado for molecular soil community analysis. The remainder of the soil was immediately passed through a 4-mm sieve. All root material not passing through the sieve was retained and stored at 4 °C before being washed free of soil prior to processing for root trait measurements.

Field plots design and sampling

Experimental field plots were established 2 km from the mesocosm enclosure at Selside Shaw, within the Ingleborough National Nature Reserve. The plots were established in 2012, in a mesotrophic grassland with similar management, vegetation and soil to the meadow at Colt Park. The soil was characterized as a clayey brown earth soil with 60% clay, <1% silt, 39% sand, 5.7±0.4 pH (mean ± standard deviation), 4.9±1.4 %C, and 0.46±0.13 %N. Native grassland species were added to the existing plant communities in 6 m × 6 m field plots with the aim of creating a gradient of plant communities of increasing functional diversity and complexity. Over two years the plots were seeded (2014-2015) and planted with seedlings (2013-2015) of species belonging to one of three plant functional groups, namely the grasses (*Cynosurus cristatus, Dactylis glomerata, Festuca rubra, Poa trivialis* and *Briza media),* forbs (*Achillea millefolium,*

Geranium sylvaticum, Geum rivale, Leucanthemum vulgare, Plantago lanceolata, Prunella vulgaris, Hypochaeris radicata, Leontodon hispidus, Filipendula ulmaria, and *Centaurea nigra*), and legumes (*Lathyrus pratensis, Lotus corniculatus, Trifolium pretense* and *Trifolium repens*) or their respective two- and three-way combinations. These species are typical of species-rich mesotrohic meadow communities (UK National Vegetation Classification MG3b; Rodwell, 1992), the target plant community for biodiversity (Smith *et al.*, 2003). Together with unmodified control communities, this created a total of eight plant community treatments with five replicates of each arranged in a randomized design (n = 40 plots). Details on species added, seedling densities, and sowing rates across all treatments are given in Table S1. We note that most, but not all, of the species contained in the mesocosms were represented in the field plots.

We sampled vegetation and soil from four of the eight treatments (control, forb addition, legume addition, and grass-forb-legume addition) in July 2015. To sample vegetation and soil, 30 cm diameter sampling rings were placed at representative locations within plots (n = 4 per plot with 5 plots per treatment; i.e. n = 20 per treatment), and aboveground plant biomass was harvested from within each sampling ring. One 6.8 cm x 10 cm soil core was collected from within the center of each sampling ring and processed identically to the mesocosm soil samples. Root material was processed as above for use in the root-based assessment of plant community composition.

Soil community composition

Fungal, bacterial, protistan, and metazoan communities were assessed in soil samples following molecular marker gene sequencing protocols as described in Prober *et al*. (2015) and Ramirez *et al*. (2014). Briefly, DNA was extracted from each sample, and ribosomal marker genes were amplified using PCR with barcoded primers unique to each sample. We used the ITS1F/ITS2

and the 515f/926r primer pairs for fungi and bacteria, respectively, and the 1391f/EukBr primer set for protists and metazoa. Amplicon pools were sequenced on an Illumina MiSeq instrument using 2x251 bp sequencing kits at the BioFrontiers sequencing facility at the University of Colorado. Appropriate controls were used throughout the laboratory process to ensure there were no contaminants. Raw sequence data are available at figshare.com using the following digital object identifiers (DOIs): [DOIs will be provided prior to publication].

Raw sequences were processed using the DADA2 pipeline (Callahan *et al.*, 2016), which is designed to resolve exact biological sequences from Illumina sequence data and does not involve sequence clustering. Raw sequences were first demultiplexed by comparing index reads to a key, and paired sequences were trimmed to uniform lengths. Sequences were then dereplicated, and the unique sequence pairs were denoised using the 'dada' function with 'err=NULL' and 'selfConsist = TRUE'. Potential primers and adapters were then screened and removed using a custom script (https://github.com/leffj/dada2helper). Next, paired-end sequences were merged and chimeras were removed. Taxonomy assignments were determined using the RDP classifier trained on the UNITE (Abarenkov *et al.*, 2010), Greengenes (McDonald *et al.*, 2012), or PR2 databases (Guillou *et al.*, 2013) for fungi, bacteria, and protists and metazoa, respectively. *Zygomycota* classifications were changed to *Mucoromycota* as per Spatafora et al. (2016). 16S rRNA gene sequences identified as chloroplasts, mitochondria, or *Archaea* were removed. To account for differences in sequencing depths, samples were rarefied to 5,300, 1,300, 2,400, and 1,250 sequences per sample for fungi, bacteria, protists, and metazoa, respectively. Putative fungal functional groups were identified using FUNGuild (Nguyen *et al.*, 2015).

Plant community composition

Plant community composition in the field plot samples was assessed in four ways: (1) by sorting the aboveground biomass to species and measuring the biomass (dry weight) of each species, (2) by molecular analysis of the aboveground biomass, (3) by molecular analysis of the roots contained in the soil cores, and (4) by molecular analysis of DNA extracted from the soil samples. For visual inspection, harvested aboveground biomass was identified the same day as collection, and tissue from each species was dried and weighed. For molecular assessments, aboveground and root biomass samples were freeze-dried, ground, and homogenized prior to DNA extraction. We prepared DNA for sequencing following a protocol similar to Kartzinel *et al*. (2015). We identified the genus-level plant community composition by targeting both the P6 loop of the *trnL* gene and the rRNA ITS region. We extracted DNA using the PowerSoil DNA Isolation Kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA), and soil samples were diluted 1:10 prior to amplification. The primer set trnL(UAA)c/trnL(UAA) with included Illumina sequencing adapters was used to amplify the *trnL*-P6 marker following a PCR protocol of: denaturing at 94 °C for 2 min followed by 36 cycles of 94 °C for 1 min, 55 °C for 30 s, and 72 °C for 30 s, with a 5-min final extension at 72 °C. To amplify the ITS region, we used the forward primer, ITS1-F, and included two reverse primers, ITS1Ast-R and ITS1Poa-R (Kartzinel *et al.*, 2015), to specifically target *Asteraceae* and *Poaceae* species. All primers included appropriate Illumina adapters, and PCR reactions were carried out as for *trnL* amplification. Each PCR was done in duplicate and the amplification product was combined. All products for each sample were combined in equal volumes and cleaned using the UltraClean PCR Clean-Up Kit (Mo Bio Laboratories, Inc.). Illumina Nextera barcodes were added to the amplicons using an 8-cycle PCR, amplicons were cleaned and pooled using the SequalPrep kit (Invitrogen,

Carlsbad, CA, USA), and sequenced on an Illumina MiSeq instrument with a 2x151 bp kit at the University of Colorado BioFrontiers sequencing facility.

We processed raw plant sequences in a similar manner as for soil community sequences described above. We used the DADA2 pipeline (Callahan *et al.*, 2016) to trim forward and reverse paired reads to 145 and 130 bp, respectively. Following the denoising step, Illumina adapters were removed, paired, end reads were merged, and chimeras were filtered. We assigned taxonomy to each sequence using BLAST searches against the GenBank NR database. Sequences were assigned taxonomy only if ≥ 80% of the sequence aligned to a reference sequence and they matched the reference sequence with ≥ 95% identity. If a sequence had multiple best matches to reference sequences, a common genus and/or family name was assigned if one existed. Otherwise, sequences were assigned as 'unknown'. Taxonomy assignments were manually checked and verified in reference to species known to exist at the site. Separate taxa tables were created based on *trnL* amplicons and each of the *Asteraceae* and *Poaceae* ITS amplicons. Samples with fewer than 550, 1000, and 100 sequences were removed from taxa tables based on *trnL*, *Asteraceae* ITS, and *Poaceae* ITS amplicons, respectively. We calculated the relative abundance of individual plant genera in each sample using the *trnL* sequence counts. Because the *trnL* gene yields limited taxonomic resolution for the *Asteraceae* and *Poaceae*, we replaced the total relative abundances of taxa (mostly unknown genera) within these two families with normalized relative abundances of genera determined using the ITS sequence data.

Plant traits

All leaf and root traits were measured using standard protocols (Cornelissen *et al.*, 2003). Briefly, we measured specific leaf area, specific root length, leaf dry matter content and root dry matter content by weighing and scanning the fresh leaf and root samples. The samples were then

oven dried at 60 °C for 48 h and their dry weights measured. The scanned digital images were analyzed in WinRhizo (Reagent Instruments Inc., Ville de Québec, QC, Canada) to determine leaf areas, root lengths and root diameters. Shoot and root N and C contents from the mesocosm-grown plants and the field sample plant communities were measured on an Elementar Vario elemental analyzer (Langenselbold, Germany). In both cases, plant material was freeze-dried and thoroughly homogenized prior to measurement.

Soil characteristics

Soil characteristics were measured as in Orwin *et al.* (2010). pH was measured using a ratio of 1 g fresh soil: 2.5 ml dH_2O. Dissolved inorganic N, individual ions (NO_3-N , NH_4-N), and net N mineralization were assessed using 1 M KCl extracts, and dissolved organic N was assessed using water extracts as in Bardgett *et al*. (2003). Total soluble N was determined following oxidation of these extracts using potassium persulphate (Bardgett *et al.*, 2003). Extracted mineral fractions were quantified using standard spectrophotometric protocols on a AA3 segmented flow analyser (SEAL Analytical Inc., Mequon, WI, USA). Total C and N of dried and ground subsamples were measured using an Elementar Vario EL elemental analyzer.

Statistical analyses

All statistical analyses were performed in R (R Core Team, 2016) using specific packages where noted, and the package 'mctoolsr' (http://leffj.github.io/mctoolsr/) was used to facilitate data manipulation and analyses. To represent differences in community composition, we calculated Bray-Curtis dissimilarities using square-root transformed relative abundances. Permutational analysis of variance (PERMANOVA), as implemented in the 'adonis' function from the 'vegan' package, was used to test for differences in soil community composition across factors. We

compared the relative abundances of taxa from control (i.e. unplanted) mesocosm communities to the relative abundances of taxa from planted mesocosms using linear mixed effects models based on rank-transformed data with block included as a random effect. *P* values were corrected for multiple comparisons using false discovery rate corrections, and zeros were replaced with an estimate of the lower detection limit (1×10^{-5}) when creating Fig. S3 to avoid infinite fold changes. To test for differences in soil community composition across mesocosm plant species, we used PERMANOVA and included block identity as a random factor in the model. Network analysis plots were created using the 'igraph' package with multidimensional scaling to distribute points. Soil taxa were considered present if their mean relative abundance was ≥ 0.1%, and only taxa with a relative abundance > 0.5% that associated with ≥ 1 plant species are shown. We identified particular soil taxa that associated with specific plant species using indicator analyses (Dufrêne and Legendre, 1997). 'Cosmopolitan' soil taxa were defined as those taxa associated with all plant species (i.e. had a mean relative abundance ≥ 0.1% across replicates for each species), 'intermediate' as taxa associated with only 2 to 20 plant species, and 'specialized' as taxa that associated with only a single plant species.

To test the relationship between the composition of soil communities and plant species relatedness in the mesocosms, we used the phylogeny from Durka and Michalski (2012). Relationships between difference in soil community composition and plant phylogenetic distances were evaluated using Mantel tests with Spearman correlations. We tested for a phylogenetic signal in the relative abundance of individual protist taxa using the phylosig function in the 'phytools' package, where the statistic, *K*, represents the strength of the signal (Blomberg *et al.*, 2003). We calculated multivariate dissimilarities in trait values by normalizing and standardizing individual trait values and calculating Euclidian distances. We tested the

relationship between Euclidian trait distances and community composition dissimilarities using Mantel tests.

For the field samples, we calculated differences in the phylogenetic structure of plant communities (i.e. phylogenetic dissimilarity) using UniFrac (Lozupone *et al.*, 2011) as implemented in the package, 'picante'. We used the plant phylogenetic tree as reported in Durka and Michalski (2012), and plants not identified to the genus level were removed. We assessed the relationship between phylogenetic dissimilarity and the Bray-Curtis dissimilarities in soil community composition using Mantel tests with Spearman correlations.

To assess whether differences in plant community composition predicted variation in soil community composition beyond the explanatory power of soil characteristics, we built models of soil community composition dissimilarity using multiple regression on distance matrices (MRM) as implemented in the 'ecodist' package and compared the explanatory power of the model with and without the addition of plant community dissimilarity as a predictor variable. In these models, each soil variable was transformed using log or inverse transformations where necessary to approximate a normal distribution, and they were standardized prior to calculating Euclidian distances. MRM was implemented with rank (i.e. Spearman) correlations, and the "best" models containing only soil variables were derived by first including all soil variables and using backwards elimination until all predictors explained significant levels of variation in the response dissimilarities.

Results and Discussion

The effect of plant species identity on soil communities

Overall, the mesocosm soils contained expectedly diverse communities (Fig. S1A). Soil fungal

communities were primarily composed of *Ascomycota* [43% of internal transcribed spacer (ITS) sequence reads, on average], *Basidiomycota* (31%), and *Mucoromycota* (21%); bacterial communities were primarily composed of *Acidobacteria* (31% of 16S rRNA gene reads, on average), *Proteobacteria* (20%), and *Verrucomicrobia* (16%); protistan communities were primarily composed of *Rhizaria* (26%), *Amoebozoa* (25%), *Alveolata* (22%), and *Stramenopiles* (16%); and metazoan communities were primarily composed of *Nematoda* (33%), *Arthropoda* (28%), and *Annelida* (15%; Fig. S1B). The structure of these communities was similar to those found in other temperate grasslands (Leff et al., 2015; Bates et al., 2013; Wu et al., 2011).

Plant species identity explained differences in the overall composition of soil fungal (R^2 = 0.33; $P < 0.001$), bacterial ($R^2 = 0.27$; $P = 0.02$), protistan ($R^2 = 0.32$; $P < 0.001$), and metazoan ($R^2 = 0.31$; $P < 0.001$) communities (Fig. 1A). Further, these plant species effects were driven by differences among multiple plant species rather than one or a small number of plant species associating with distinct belowground communities (Fig. 1B, Fig. S2). Certain fungal, protistan, and metazoan taxa tended to be strongly associated with individual plant species, while others tended to have more general associations (Fig. 1C, Fig. S3). For example, the fungal taxa identified as *Olpidium brassicae* and *Phoma* sp. associated with *Achillea millefolium*, while several *Ascomycota*, *Basidiomycota*, and *Mucoromycota* taxa were associated with all plant species (Fig. S4). We used an indicator analysis approach to identify those taxonomic groups that were most strongly associated with each of the individual plant species and found that many of the plant species formed specific associations (Fig. S4). Since there are likely to be different traits associated with more specialized versus more cosmopolitan soil taxa (Lennon et al., 2012), we investigated whether soil taxa unique to individual plant species tended to represent different taxonomic groups when compared to taxa that were more ubiquitous across plant species.

Cosmopolitan taxa were represented by a higher proportion of *Mucoromycota*, *Acidobacteria*, *Rhizaria*, and *Nematoda*, while more specialized taxa were represented by a greater proportion of *Glomeromycota*, *Planctomycetes*, *Alveolata*, and *Rotifera* (Fig. 1D). Additionally, cosmopolitan fungal taxa represented a greater proportion of putative saprotrophs compared to more specialized taxa, which had a greater proportion of pathogens and mutualists (Fig. 1E). This suggests that, in temperate grasslands, pathogens and mutualists tend to be more strongly limited to individual plant species, while saprotrophs are more cosmopolitan and less influenced by plant species identity. This finding is in concordance with a previous study conducted in an Amazon rainforest showing stronger plant-soil linkages for pathogenic and mycorrhizal fungi compared to saprotrophs (Peay et al., 2013).

Can the effect of plant species identity be explained by plant phylogeny or functional traits?

We next sought to assess whether plant species identity effects could be explained by plant phylogeny or leaf and root functional traits, two attributes that could potentially be used to predict plant associations with belowground communities a priori. The mesocosm plant species represented eight families including *Poaceae*, *Asteraceae*, and *Fabaceae*, providing an opportunity to evaluate the influence of a wide-ranging phylogeny on the composition of soil communities. Plant phylogenetic distances were not significantly related to differences in fungal, bacterial, or metazoan community composition ($P > 0.1$ in all cases; Fig. 2A). Differences in protistan community composition were related to plant phylogenetic distance, but this relationship was relatively weak (rho = 0.29, $P = 0.002$; Fig. 2A). Nonetheless, the relative abundance of *Stramenopiles* was significantly related to plant species phylogeny ($K = 0.51$, $P = 0.004$; Fig. S5). We might expect plant phylogenetic differences to be associated with the structure of belowground communities due to coevolution with mutualists or pathogens (De

376 Deyn and Van Der Putten, 2005; Anacker et al., 2014); however, this did not appear to be the
377 case for most soil taxonomic groups. Further, the general lack of a relationship between plant
378 phylogeny and belowground communities found in our study is consistent with studies of plant-
379 soil feedbacks, which likewise have shown no relation to plant phylogeny (Mehrabi and Tuck,
380 2015).

381 The measured leaf and root traits were highly variable across the mesocosm species.
382 Grassland plants vary in their ecological strategies. Exploitative species grow fast under high
383 nutrient conditions and have characteristically high specific leaf areas and N contents while
384 conservative species are selected to survive under lower nutrient conditions and have opposite
385 traits (Lavorel and Garnier, 2002; Roumet et al., 2016). For each plant species in the mesocosms,
386 we measured the plant traits that are known to be indicative of the tradeoffs in these life history
387 strategies (Fig. S6A, Table S2). For example, the *Fabaceae* species tended to have a greater
388 shoot and root N and C content, while *Poaceae* species tended to have high leaf dry matter
389 contents (Fig. S6B). Yet, there were no strong or significant relationships (i.e., Bonferroni
390 corrected $P < 0.05$) between belowground community composition and individual leaf or root
391 traits (Fig. 2C). Furthermore, multivariate dissimilarity in leaf and root traits of plant species was
392 not predictive of differences in communities of any of the soil taxonomic groups ($P > 0.1$ in all
393 cases; Fig. 2B).

394 These results suggest that the plant traits we measured are not effective indicators of the
395 specific relationships plants form with belowground communities. Previous studies have
396 detected relationships between plant traits and coarse measures of microbial community
397 composition (Orwin et al., 2010; de Vries et al., 2012) or specific microbial groups, such as
398 ammonia oxidizers (Thion et al., 2016). However, our findings are in line with other studies. For

example, Porazinska *et al*. *(2003)* found that certain soil communities were linked to individual plant species in a prairie grassland, but they were unable to identify traits that could predict soil communities. Likewise, Barberán et al. (2015a) demonstrated that plant species identity is more predictive of soil communities than plant traits. Nonetheless, it is possible that the plant-soil organism associations we observed could have been driven by unmeasured plant traits given that certain plant characteristics must explain the species identity effects we observed. For example, variations in the quantity and quality of root exudates can influence soil community composition (Haichar et al., 2008). Likewise, leaf litter chemistry has been shown to be related to coarse measures of soil microbial community composition in a manner broadly consistent with the leaf economic spectrum (Orwin et al., 2010). Also, while we did not observe relationships between plant traits and the overall composition of soil communities, it is possible that specific soil organisms do respond to plant traits, including those taxa directly involved with N cycling (Legay et al., 2014; Moreau et al., 2015; Thion et al., 2016). Other potential reasons exist for our failure to detect strong associations between soil communities and plant traits or phylogeny. First, it is possible that if the experiment had a longer duration, additional effects on soil communities would become evident, and these effects would more strongly correspond to differences in plant traits and/or phylogeny. Second, soil can contain DNA from cells that are no longer viable (Carini *et al*., 2016), and this 'relic' DNA could obscure ecological relationships among organisms.

Are soil communities in the field predictable based on plant community attributes?

The results from the mesocosm study demonstrated that plant species identity is a more important determinant of soil community composition than plant phylogeny or plant traits. Given this, we would hypothesize that knowledge of the species composition of mixed plant

communities in the field should be an effective predictor of soil communities. We tested this hypothesis by analyzing plant and soil samples from a series of experimental plots established at a grassland site close to the mesocosm experiment, where grassland community composition had been manipulated for three years to create a gradient of plant species composition and diversity. Plant community composition was assessed using marker gene sequencing of plant DNA extracted from dried and ground representative samples of plant biomass collected immediately above each soil sample, and this molecular approach was verified for efficacy by comparing it to visual assessments of aboveground biomass (Fig. S7).

Differences in the composition of each soil taxonomic group were related to differences in plant community composition ($P < 0.05$ in all cases). By comparing the compositions of the plant communities across experimental plots (using the first principal coordinate score based on aboveground assessments), we could identify specific plant genera that drove variation in soil community composition across the samples (Fig. 3A, Table S3). For instance, some samples had comparatively high relative abundances of *Lolium* spp. while other samples had high relative abundances of *Agrostis* spp. These differences in plant community composition were related to the relative abundance of certain groups of soil taxa, including the *Ascomycota*, *Mucoromycota*, *Acidobacteria*, *Amoebozoa, Stramenopiles*, and *Arthropoda* (Fig. 3A). These specific associations between plant and soil taxa can ultimately be used to predict the composition of soil communities from plant species abundances. For example, our results suggest that plant communities dominated by *Agrostis* spp. are likely to have greater relative abundances of *Ascomycota* and lower relative abundances of *Acidobacteria* in the soils in which they grow.

We also evaluated whether the phylogenetic structure or community-aggregated plant traits (de Vries et al., 2012; Grigulis et al., 2013) could explain relationships between plants and

soil communities. We did this by testing whether plant communities containing genera with more similar phylogenetic histories or trait values were associated with more similar soil communities. However, plant community phylogenetic structure was not significantly related to the composition of any of the soil taxonomic groups ($P > 0.3$ in all cases), suggesting that phylogenetic relatedness is not predictive of soil community composition. This finding is in agreement with the monoculture mesocosm study described above and a field study conducted in a tropical rainforest that failed to find a strong effect of tree species phylogenetic relationships on soil communities (Barberán et al., 2015b). Furthermore, differences in community-aggregated trait values, including leaf and root N and C content, also did not significantly relate to the composition of any of the soil taxonomic groups ($P > 0.1$ in all cases). The trait values we measured were not predictive of soil community composition in mixed grassland communities, results that are consistent with those from the mesocosm experiment of individual plant species.

In addition to assessing relationships between the composition of soil taxonomic groups and plant communities based on aboveground biomass, we evaluated plant community composition in two other ways: using root DNA and plant DNA in soil. We used these approaches because roots of different species are intermingled and difficult to identify visually, and assessing plant communities via soil DNA provides an alternate approach to determine which plant species have occupied a given location currently or in the past (Yoccoz et al., 2012). Roots might also might be more strongly associated with soil community structure than aboveground tissue (Orwin *et al.*, 2010). As with the aboveground plant biomass-based analysis, differences in the compositions of each of the soil taxonomic groups were related to differences in plant community composition assessed using the plant DNA extracted from soil ($P < 0.05$ in all cases). However, the differences in the composition of soil communities were not

significantly related to differences in plant community composition assessed using root DNA (P > 0.1 in all cases; Fig. 3B). It is possible that the composition of plant communities as assessed via roots were unrelated to soil communities because much of the root biomass consisted of dormant plants or dead tissue (Hiiesalu et al., 2012). Further, it is possible that root distributions are so variable over time that they obscure plant species effects on belowground communities.

Differences in aboveground plant community composition were unrelated to differences in root community composition (P = 0.11), but they were related to differences in the plant community composition as assessed using plant DNA in soil (rho = 0.2; P < 0.001; Fig. 3C). This shows that shoot and root biomass in a given location do not represent the same plant community, as also found in a tropical rainforest (Barberán et al., 2015b). Additionally, these results suggest that plant DNA in soil can be used as a proxy for the community composition of the aboveground biomass (Yoccoz et al., 2012). This has implications for future research since it is often logistically easier to obtain a representative sample of surface soils rather than sampling and homogenizing aboveground plant biomass.

Are the associations between plant and soil communities driven by soil characteristics?

We aimed to assess whether relationships between soil communities and plant communities in the field plots were attributable to the direct effects of the plants, shared environmental drivers, or intermediary effects of the plants on soil properties. Therefore, we evaluated whether plant community composition contributed additional explanatory power to the observed variation in soil community composition given differences in edaphic characteristics. Shifts in the composition of soil communities across the field plots were significantly correlated with multiple, individual edaphic properties (Table S4), and combinations of these properties explained 13 – 29% of the variation in soil community composition (P = 0.001 in all cases; Fig.

S8A). For example, soil N content and pH were typically predictive of the composition of the four taxonomic soil groups. Only differences in fungal community composition could be predicted more accurately when information on aboveground plant community composition was added to the models containing only soil characteristics as predictor variables ($P = 0.01$; Fig. S8). When soil DNA-based plant community composition information was used instead of aboveground plant community composition, fungal, bacterial, and protistan community composition could all be predicted more accurately with the addition of information on plant community composition (R^2 increased 9 – 24%; $P < 0.02$ in all cases; Fig. S8). These results suggest that shifts in aboveground community composition likely influence soil communities in ways not accounted for in commonly measured soil properties, and indicate that the structure of complex soil communities in grasslands is controlled by a combination of plant and soil characteristics (Berg and Smalla, 2009; Harrison and Bardgett, 2010).

Conclusions

We demonstrate that plant community composition is an effective predictor of the structure of complex grassland soil communities, especially when combined with information on soil abiotic properties. Furthermore, we show that plant community composition is particularly effective for predicting distributions of certain groups of soil organisms, such as fungal symbionts and pathogens. Importantly, we found that plant species identity, rather than plant phylogeny or functional traits, was the best predictor of soil community composition at both the individual plant and community scale. This is significant because it raises questions about the effectiveness of phylogenetic and trait-based approaches for explaining spatial variation in soil community composition at a local scale. Such approaches are increasingly being used to predict how changes in plant community composition impact soil properties and functions (Bardgett et al., 2014;

Laliberté, 2017), but our findings indicate that, at a local scale in temperate grassland, they are ineffective for explaining variation in soil communities. Finally, it is important to note that much of the variation in soil community composition could not be explained by the measured soil characteristics or plant community attributes, highlighting the difficulty of predicting complex soil communities in situ and the need to build a mechanistic understanding of which specific plant attributes are responsible for driving plant species effects on the biodiversity of soil. Combined, our findings provide new evidence that associations between specific plant species and complex soil communities, associations that are not explained by plant phylogeny or commonly measured plant traits, act as key determinants of spatial patterns of biodiversity in grassland soils.

Acknowledgements

This research was supported by a grant from the UK Biotechnology and Biological Sciences Research Council (BBSRC) (Grant BB/I009000/2), initiated and led by RDB, a BBSRC International Exchange Grant (BB/L026406/1) between RDB and NF, and a grant from the U.S. National Science Foundation (NSF) (DEB 1542653) awarded to NF. We thank Emily Morgan for her assistance with the microbial community analyses, Colin Newlands of Natural England for permission to use the field sites, and Marina Semchenko for comments on a previous version of this manuscript. We also thank all who helped establish the field experimental site and collect and process the plant and soil samples: Aurore Kaisermann, Debora Ashworth, Angela Straathof, Imelda Uwase, Marina Semchenko, Maatren Schrama, Melanie Edgar, Mark Bradford, Mike Whitfield, Rachel Marshall, and Andrew Cole.

Supplementary information is available at The ISME Journal's website.

References

Abarenkov K, Henrik Nilsson R, Larsson K-H, Alexander IJ, Eberhardt U, Erland S, *et al.* (2010). The UNITE database for molecular identification of fungi – recent updates and future perspectives. *New Phytol* **186**: 281–285.

Adler PB, Fajardo A, Kleinhesselink AR, Kraft NJB. (2013). Trait-based tests of coexistence mechanisms. *Ecol Lett* **16**: 1294–1306.

Anacker BL, Klironomos JN, Maherali H, Reinhart KO, Strauss SY. (2014). Phylogenetic conservatism in plant-soil feedback and its implications for plant abundance. *Ecol Lett* **17**: 1613–1621.

Barberán A, Dunn RR, Reich BJ, Pacifici K, Laber EB, Menninger HL, *et al.* (2015a). The ecology of microscopic life in household dust. *Proc R Soc B* **282**: 1–9.

Barberán A, Mcguire KL, Wolf JA, Jones FA, Wright SJ, Turner BL, *et al.* (2015b). Relating belowground microbial composition to the taxonomic, phylogenetic, and functional trait distributions of trees in a tropical forest. *Ecol Lett* **18**: 1397–1405.

Bardgett RD, Mawdsley JL, Edwards S, Hobbs PJ, Rodwell JS, Davies WJ. (1999). Plant species and nitrogen effects on soil biological properties of temperate upland grasslands. *Funct Ecol* **13**: 650–660.

Bardgett RD, Mommer L, De Vries FT. (2014). Going underground: Root traits as drivers of ecosystem processes. *Trends Ecol Evol* **29**: 692–699.

Bardgett RD, Streeter TC, Bol R. (2003). Soil microbes compete effectively with plants for organic-nitrogen inputs to temperate grasslands. *Ecology* **84**: 1277–1287.

Bardgett RD, Wardle DA. (2010). Aboveground-belowground linkages: biotic interactions,

559 ecosystem processes, and global change. Oxford University Press Oxford: New York, NY, USA.
560 Bates ST, Clemente JC, Flores GE, Walters WA, Parfrey LW, Knight R, *et al.* (2013). Global
561 biogeography of highly diverse protistan communities in soil. *ISME J* **7**: 652–659.
562 Ben-Hur E, Fragman-Sapir O, Hadas R, Singer A, Kadmon R. (2012). Functional trade-offs
563 increase species diversity in experimental plant communities. *Ecol Lett* **15**: 1276–1282.
564 Berg G. (2009). Plant-microbe interactions promoting plant growth and health: Perspectives for
565 controlled use of microorganisms in agriculture. *Appl Microbiol Biotechnol* **84**: 11–18.
566 Berg G, Smalla K. (2009). Plant species and soil type cooperatively shape the structure and
567 function of microbial communities in the rhizosphere. *FEMS Microbiol Ecol* **68**: 1–13.
568 Bezemer TM, Fountain MT, Barea JM, Christensen S, Dekker SC, Duyts H, *et al.* (2010).
569 Divergent composition but similar function of soil food webs beneath individual plants: plant
570 species and community effects. *Ecology* **91**: 3027–3036.
571 Bezemer TM, Lawson CS, Hedlund K, Edwards AR, Brook AJ, Igual JM, *et al.* (2006). Plant
572 species and functional group effects on abiotic and microbial soil properties and plant-soil
573 feedback responses in two grasslands. *J Ecol* **94**: 893–904.
574 Blomberg SP, Garland T, Ives AR. (2003). Testing for phylogenetic signal in comparative data:
575 behavioral traits are more labile. *Evolution (N Y)* **57**: 717–745.
576 Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. (2016). DADA2 :
577 High resolution sample inference from amplicon data. *Nat Methods* **13**: 581–583.
578 Cantarel AAM, Pommier T, Desclos-Theveniau M, Diquélou S, Dumont M, Grassein F, *et al.*
579 (2015). Using plant traits to explain plant–microbe relationships involved in nitrogen acquisition.
580 *Ecology* **96**: 788–799.
581 Carey CJ, Michael Beman J, Eviner VT, Malmstrom CM, Hart SC. (2015). Soil microbial

582 community structure is unaltered by plant invasion, vegetation clipping, and nitrogen fertilization
583 in experimental semi-arid grasslands. *Front Microbiol* **6**. e-pub ahead of print, doi:
584 10.3389/fmicb.2015.00466.
585 Carini P, Marsden PJ, Leff JW, Morgan EE, Strickland MS, Fierer N. (2016). Relic DNA is
586 abundant in soil and obscures estimates of soil microbial diversity. *Nat Microbiol* **2**. e-pub ahead
587 of print, doi: 10.1038/nmicrobiol.2016.242.
588 Cornelissen JHC, Lavorel S, Garnier E, Díaz S, Buchmann N, Gurvich DE, *et al.* (2003). A
589 handbook of protocols for standardised and easy measurement of plant functional traits
590 worldwide. *Aust J Bot* **51**: 335–380.
591 De Deyn GB, Van Der Putten WH. (2005). Linking aboveground and belowground diversity.
592 *Trends Ecol Evol* **20**: 625–633.
593 De Deyn GB, Shiel RS, Ostle NJ, Mcnamara NP, Oakley S, Young I, *et al.* (2011). Additional
594 carbon sequestration benefits of grassland diversity restoration. *J Appl Ecol* **48**: 600–608.
595 Dufrêne M, Legendre P. (1997). Species assemblages and indicator species: The need for a
596 flexible asymmetrical approach. *Ecol Monogr* **67**: 345–366.
597 Durka W, Michalski SG. (2012). Daphne: a dated phylogeny of a large European flora for
598 phylogenetically informed ecological analyses. *Ecology* **93**: 2297.
599 Fierer N, Strickland MS, Liptzin D, Bradford M a., Cleveland CC. (2009). Global patterns in
600 belowground communities. *Ecol Lett* **12**: 1238–1249.
601 Gibbons SM, Lekberg Y, Mummey DL, Sangwan N, Ramsey PW, Gilbert JA. (2017). Invasive
602 Plants Rapidly Reshape Soil Properties in a Grassland Ecosystem. *mSystems* **2**: e00178-16.
603 Grayston SJ, Griffith GS, Mawdsley JL, Campbell CD, Bardgett RD. (2001). Accounting for
604 variability in soil microbial communities of temperate upland grassland ecosystems. *Soil Biol*

Biochem **33**: 533–551.

Grayston SJ, Wang S, Campbell CD, Edwards AC. (1998). Selective influence of plant species on microbial diversity in the rhizosphere. *Soil Biol Biochem* **30**: 369–378.

Grigulis K, Lavorel S, Krainer U, Legay N, Baxendale C, Dumont M, *et al.* (2013). Relative contributions of plant traits and soil microbial properties to mountain grassland ecosystem services. *J Ecol* **101**: 47–57.

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, *et al.* (2013). The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* **41**: 597–604.

Haichar FEZ, Marol C, Berge O, Rangel-Castro JI, Prosser JI, Balesdent J, *et al.* (2008). Plant host habitat and root exudates shape soil bacterial community structure. *ISME J* **2**: 1221–1230.

Harrison KA, Bardgett RD. (2010). Influence of plant species and soil conditions on plant-soil feedback in mixed grassland communities. *J Ecol* **98**: 384–395.

Hawkes C V., Wren IF, Herman DJ, Firestone MK. (2005). Plant invasion alters nitrogen cycling by modifying the soil nitrifying community. *Ecol Lett* **8**: 976–985.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T, *et al.* (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**: 69–72.

Hiiesalu I, Öpik M, Metsis M, Lilje L, Davison J, Vasar M, *et al.* (2012). Plant species richness belowground: Higher richness and new patterns revealed by next-generation sequencing. *Mol Ecol* **21**: 2004–2016.

Hiiesalu I, Pärtel M, Davison J, Gerhold P, Metsis M, Moora M, *et al.* (2014). Species richness of arbuscular mycorrhizal fungi: Associations with grassland plant richness and biomass. *New*

Phytol **203**: 233–244.

St. John MG, Wall DH, Behan-Pelletier VM. (2006). Does plant species co-occurrence influence soil mite diversity? *Ecology* **87**: 625–633.

Kaiser K, Wemheuer B, Korolkow V, Wemheuer F, Nacke H, Schöning I, *et al.* (2016). Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. *Sci Rep* **6**: 33696.

Kartzinel TR, Chen P a., Coverdale TC, Erickson DL, Kress WJ, Kuzmina ML, *et al.* (2015). DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proc Natl Acad Sci* **112**: 8019–8024.

Laliberté E. (2017). Below-ground frontiers in trait-based plant ecology. *New Phytol* **213**: 1597–1603.

Lavorel S, Garnier E. (2002). Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Funct Ecol* **16**: 545–556.

Leff JW, Jones SE, Prober SM, Barberan A, Borer ET, Firn JL, *et al.* (2015). Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proc Natl Acad Sci U S A* **112**: 10967–10972.

Legay N, Baxendale C, Grigulis K, Krainer U, Kastl E, Schloter M, *et al.* (2014). Contribution of above- and below-ground plant traits to the structure and function of grassland soil microbial communities. *Ann Bot* **114**: 1011–1021.

Legay N, Lavorel S, Baxendale C, Krainer U, Bahn M, Binet M-N, *et al.* (2016). Influence of plant traits, soil microbial properties, and abiotic parameters on nitrogen turnover of grassland ecosystems. *Ecosphere* **7**: 1–17.

Lekberg Y, Waller LP. (2016). What drives differences in arbuscular mycorrhizal fungal

communities among plant species? *Fungal Ecol* 10–13.

Lennon JT, Aanderud ZT, Lehmkuhl BK, Schoolmaster DR. (2012). Mapping the niche space of soil microorganisms using taxonomy and traits. *Ecology* **93**: 1867–1879.

Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. (2011). UniFrac: an effective distance metric for microbial community comparison. *ISME J* **5**: 169–72.

McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, *et al.* (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* **6**: 610–618.

Mehrabi Z, Tuck SL. (2015). Relatedness is a poor predictor of negative plant-soil feedbacks. *New Phytol* **205**: 1071–1075.

Moreau D, Pivato B, Bru D, Busset H, Deau F, Faivre C, *et al.* (2015). Plant traits related to nitrogen uptake influence plant-microbe competition. *Ecology* **96**: 2300–2310.

Nguyen NH, Song Z, Bates ST, Branco S, Tedersoo L, Menke J, *et al.* (2015). FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol* **20**: 241–248.

Orwin KH, Buckland SM, Johnson D, Turner BL, Smart S, Oakley S, *et al.* (2010). Linkages of plant traits to soil properties and the functioning of temperate grassland. *J Ecol* **98**: 1074–1083.

Peay KG, Baraloto C, Fine PVA. (2013). Strong coupling of plant and fungal community structure across western Amazonian rainforests. *ISME J* **7**: 1852–61.

Porazinska DL, Bardgett RD, Blaauw MB, Hunt HW, Parsons AN, Seastedt TR, *et al.* (2003). Relationships at the Aboveground-Belowground Interface: Plants, Soil Biota, and Soil Processes. *Ecol Monogr* **73**: 377–395.

Prober SM, Leff JW, Bates ST, Borer ET, Firn J, Harpole WS, *et al.* (2015). Plant diversity

674 predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecol Lett* **18**:
675 85–95.
676 Van der Putten WH, Bardgett RD, Bever JD, Bezemer TM, Casper BB, Fukami T, *et al.* (2013).
677 Plant-soil feedbacks: The past, the present and future challenges Hutchings M (ed). *J Ecol* **101**:
678 265–276.
679 R Core Team. (2016). R: A language and environment for statistical computing.
680 Ramirez KS, Leff JW, Barberán A, Bates ST, Betley J, Crowther TW, *et al.* (2014).
681 Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to
682 those observed globally. *Proc R Soc B Biol Sci* **281**: 1–9.
683 Rodwell JS. (1992). British plant communities. Volume 3. Grassland and montane communities.
684 3rd ed. Cambridge University Press.
685 Roumet C, Birouste M, Picon-Cochard C, Ghestem M, Osman N, Vrignon-Brenas S, *et al.*
686 (2016). Root structure-function relationships in 74 species: Evidence of a root economics
687 spectrum related to carbon economy. *New Phytol*. e-pub ahead of print, doi: 10.1111/nph.13828.
688 Singh BK, Millard P, Whiteley AS, Murrell JC. (2004). Unravelling rhizosphere–microbial
689 interactions: opportunities and limitations. *Trends Microbiol* **12**: 386–393.
690 Smith RS, Shiel RS, Bardgett RD, Millward D, Corkhill P, Rolph G, *et al.* (2003). Soil microbial
691 community, fertility, vegetation and diversity as targets in the restoration management of a
692 meadow grassland. *J Appl Ecol* **40**: 51–64.
693 Spatafora JW, Chang Y, Benny GL, Lazarus K, Smith ME, Berbee ML, *et al.* (2016). A phylum-
694 level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia*
695 **108**: 1028–1046.
696 Tedersoo L, Bahram M, Cajthaml T, Põlme S, Hiiesalu I, Anslan S, *et al.* (2015). Tree diversity

and species identity effects on soil fungi, protists and animals are context dependent. *ISME J* 1–17.

Tedersoo L, Bahram M, Polme S, Koljalg U, Yorou NS, Wijesundera R, *et al.* (2014). Global diversity and geography of soil fungi. *Science (80-)* **346**: 1256688–1256688.

Thion CE, Poirel JD, Cornulier T, De Vries FT, Bardgett RD, Prosser JI. (2016). Plant nitrogen-use strategy as a driver of rhizosphere archaeal and bacterial ammonia oxidiser abundance. *FEMS Microbiol Ecol* **92**. e-pub ahead of print, doi: 10.1093/femsec/fiw091.

de Vries FT, Manning P, Tallowin JRB, Mortimer SR, Pilgrim ES, Harrison KA, *et al.* (2012). Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities. *Ecol Lett* **15**: 1230–1239.

Wardle DA, Bardgett RD, Klironomos JN, Setälä H, van der Putten WH, Wall DH. (2004). Ecological linkages between aboveground and belowground biota. *Science (80-)* **304**: 1629–33.

Wu T, Ayres E, Bardgett RD, Wall DH, Garey JR. (2011). Molecular study of worldwide distribution and diversity of soil animals. *Proc Natl Acad Sci* **108**: 17720–17725.

Yoccoz NG, Bråthen KA, Gielly L, Haile J, Edwards ME, Goslar T, *et al.* (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Mol Ecol* **21**: 3647–3655.

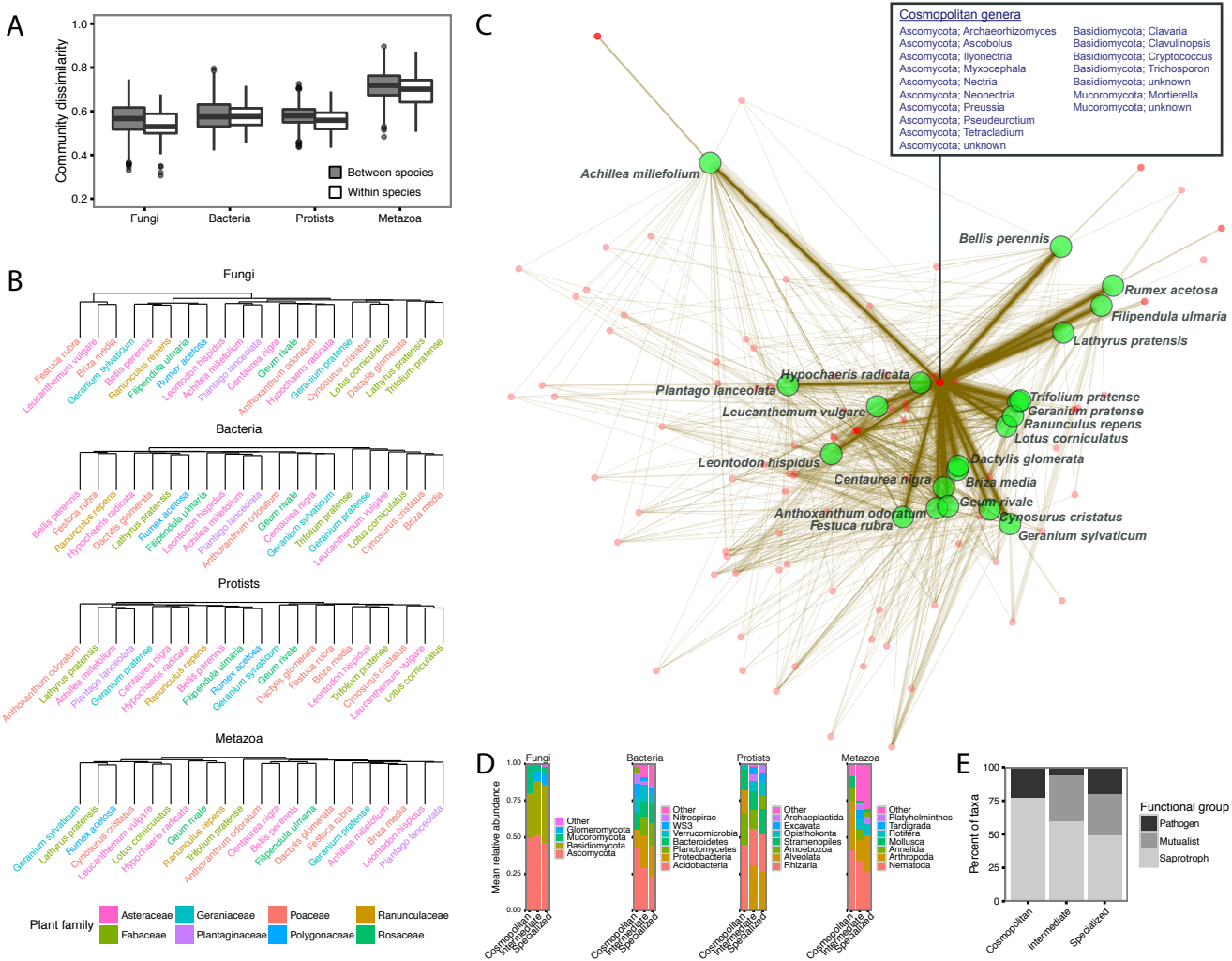
Figure legends

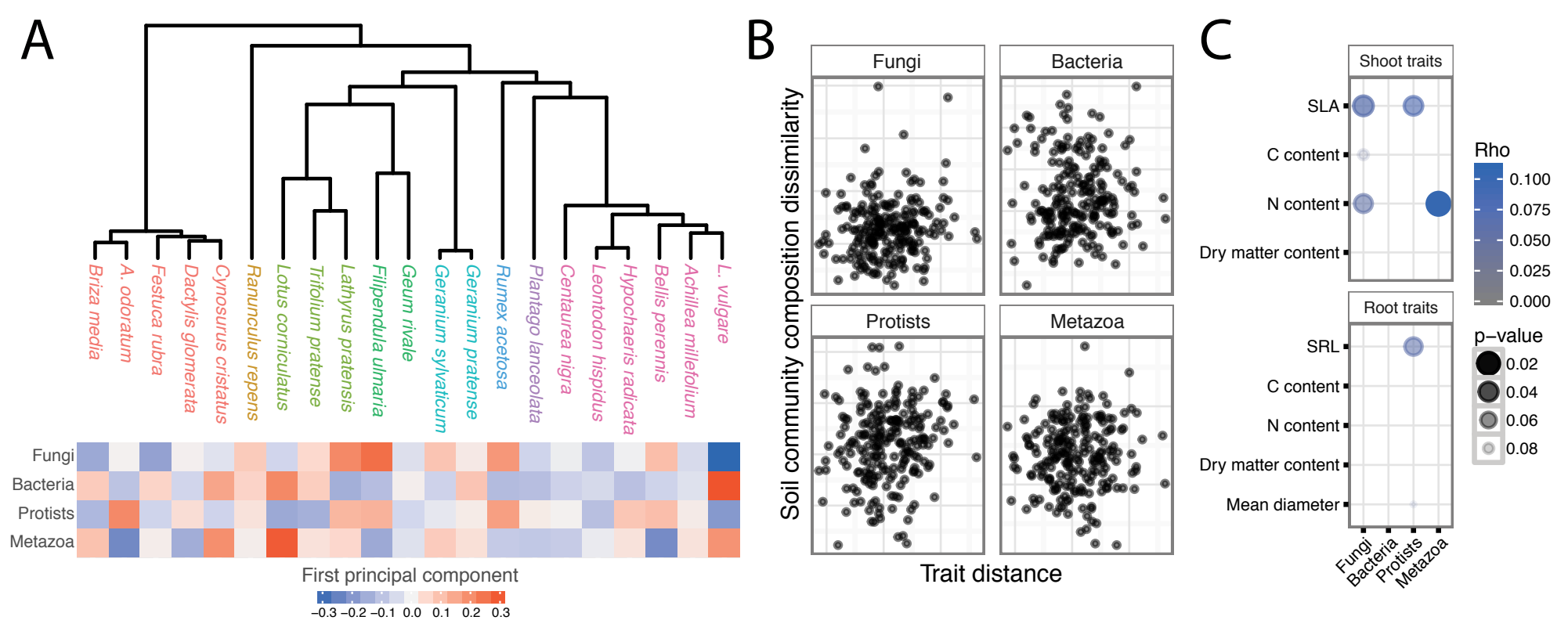
Figure 1. The effects of plant species identity on the composition of soil communities from mesocosms containing monocultures. Boxplots represent pairwise Bray-Curtis dissimilarities in community composition between vs. within soils from the same plant species (A). Hierarchical clustering diagrams based on mean dissimilarities across the plant species (B). Bipartite network diagram, where edges (lines) connect plant species (green circles) to fungal taxa (red points) that occurred in the same mesocosm (C). The composition of cosmopolitan soil taxa (those taxa associated with all plant species), intermediate (taxa associated with only 2 to 20 plant species), and specialized (taxa that associate with only a single plant species) (D). The composition of functional groups of fungal taxa identified as being cosmopolitan, intermediate, and specialized across plant species (E).

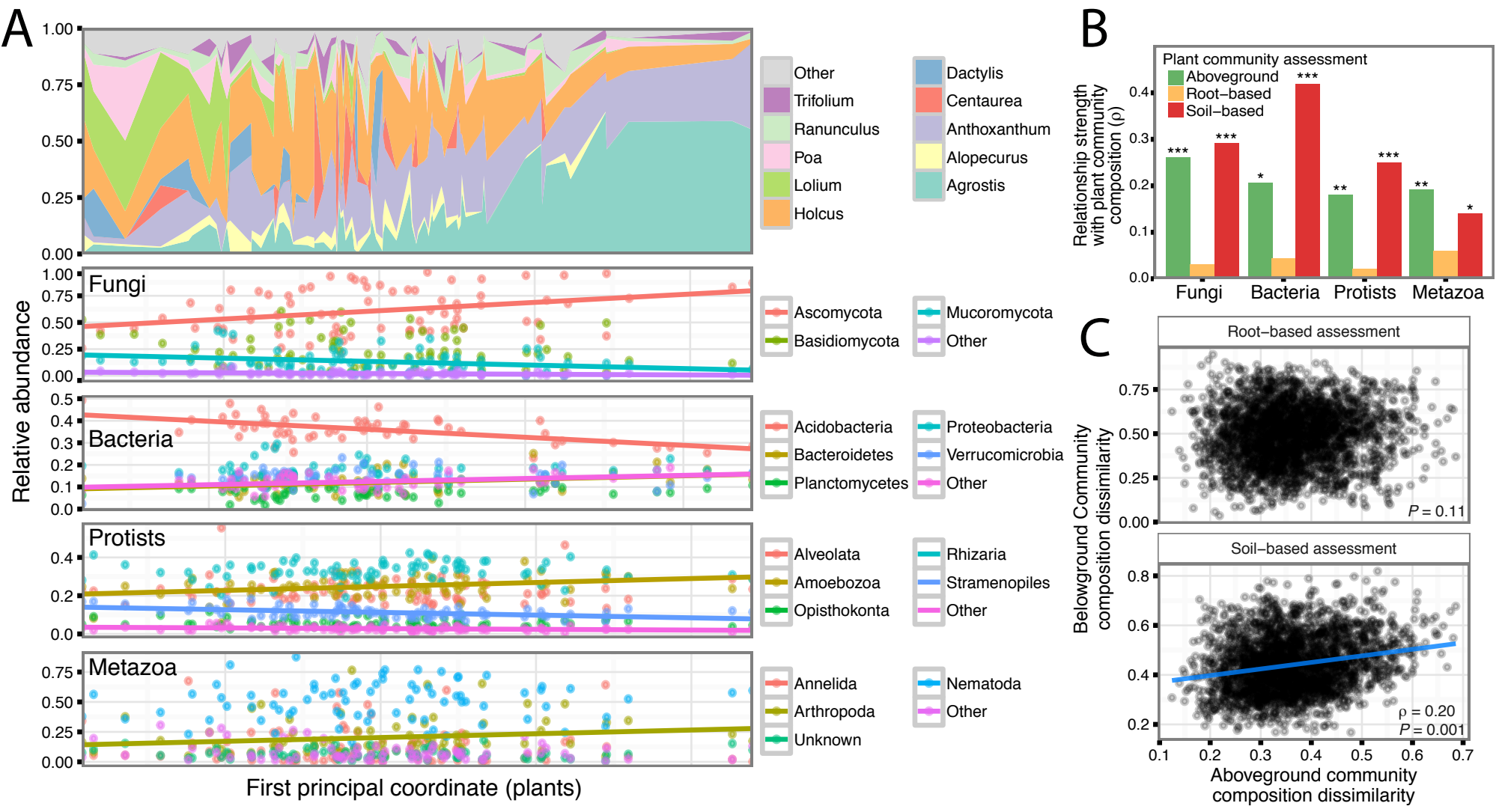
Figure 2. Relationships between plant species' relatedness and differences in the composition of soil communities. Panel A shows a plant phylogenetic tree with species names colored by family (key shown in Fig. 1) with the corresponding heatmap showing the dissimilarities in the composition of each soil community. Colors represent the first principal coordinate analysis axis calculated from Bray-Curtis dissimilarities (A). The relationship between differences in the composition of soil communities and plant trait distances (B). Euclidean trait distances were calculated using all the traits shown in panel C. The relationship between differences in the composition of soil communities and individual plant traits (C). Points represent Spearman correlation coefficients (rho) and Mantel test results (*P* value).

Figure 3. Soil community composition is related to plant community composition in the field.

Variation in plant community composition across the field samples ordered by the first principal coordinate score (i.e. the x-axis represents a gradient of plant community compositions where communities further apart are more dissimilar), and relationships between soil taxonomic group relative abundance and the plant first principal coordinate score (A). Linear trend lines were only plotted for groups that had a Pearson correlation $P \leq 0.05$. Relationship strength between dissimilarities in soil communities and dissimilarities in plant communities (* = $P < 0.05$, ** = $P < 0.01$, *** = $P = 0.001$; Mantel tests; B). Pairwise Bray-Curtis dissimilarities in plant community composition, as assessed using aboveground tissue, are not related to dissimilarities in plant community composition as assessed using root tissue, but they are related to dissimilarities in plant community composition as assessed using plant DNA in soil (C).


A
Community dissimilarity
1.0
0.8
0.6
0.4
0.2
Between species
Within species
Fungi
Bacteria
Protists
Metazoa
B
Fungi
Festuca rubra
Briza media
Leucanthemum vulgare
Geranium sylvaticum
Bellis perennis
Ranunculus repens
Filipendula ulmaria
Rumex acetosa
Leontodon hispidus
Achillea millefolium
Plantago lanceolata
Anthoxanthum odoratum
Hypochaeris radicata
Geranium pratense
Centaurea nigra
Geum rivale
Cynosurus cristatus
Lotus corniculatus
Dactylis glomerata
Lathyrus pratensis
Trifolium pratense
Bacteria
Bellis perennis
Festuca rubra
Ranunculus repens
Hypochaeris radicata
Dactylis glomerata
Lathyrus pratensis
Rumex acetosa
Filipendula ulmaria
Leontodon hispidus
Achillea millefolium
Plantago lanceolata
Anthoxanthum odoratum
Geum rivale
Centaurea nigra
Geranium sylvaticum
Trifolium pratense
Geranium pratense
Leucanthemum vulgare
Lotus corniculatus
Cynosurus cristatus
Briza media
Protists
Anthoxanthum odoratum
Lathyrus pratensis
Achillea millefolium
Plantago lanceolata
Geranium pratense
Centaurea nigra
Hypochaeris radicata
Ranunculus repens
Bellis perennis
Filipendula ulmaria
Rumex acetosa
Geranium sylvaticum
Geum rivale
Dactylis glomerata
Festuca rubra
Briza media
Leontodon hispidus
Trifolium pratense
Cynosurus cristatus
Leucanthemum vulgare
Lotus corniculatus
Metazoa
Geranium sylvaticum
Lathyrus pratensis
Rumex acetosa
Cynosurus cristatus
Leucanthemum vulgare
Lotus corniculatus
Hypochaeris radicata
Geum rivale
Ranunculus repens
Trifolium pratense
Anthoxanthum odoratum
Centaurea nigra
Bellis perennis
Filipendula ulmaria
Dactylis glomerata
Festuca rubra
Geranium pratense
Achillea millefolium
Briza media
Leontodon hispidus
Plantago lanceolata
Plant family
Asteraceae
Fabaceae
Geraniaceae
Plantaginaceae
Poaceae
Polygonaceae
Ranunculaceae
Rosaceae
C
Cosmopolitan genera
Ascomycota; Archaeorhizomyces
Ascomycota; Ascobolus
Ascomycota; Ilyonectria
Ascomycota; Myxocephala
Ascomycota; Nectria
Ascomycota; Neonectria
Ascomycota; Preussia
Ascomycota; Pseudeurotium
Ascomycota; Tetracladium
Ascomycota; unknown
Basidiomycota; Clavaria
Basidiomycota; Clavulinopsis
Basidiomycota; Cryptococcus
Basidiomycota; Trichosporon
Basidiomycota; unknown
Mucoromycota; Mortierella
Mucoromycota; unknown
Achillea millefolium
Bellis perennis
Rumex acetosa
Filipendula ulmaria
Lathyrus pratensis
Hypochaeris radicata
Plantago lanceolata
Leucanthemum vulgare
Trifolium pratense
Geranium pratense
Ranunculus repens
Lotus corniculatus
Leontodon hispidus
Dactylis glomerata
Centaurea nigra
Briza media
Geum rivale
Anthoxanthum odoratum
Festuca rubra
Cynosurus cristatus
Geranium sylvaticum
D
Fungi
Bacteria
Protists
Metazoa
Mean relative abundance
1.00
0.75
0.50
0.25
0.00
Other
Glomeromycota
Mucoromycota
Basidiomycota
Ascomycota
Other
Nitrospirae
WS3
Verrucomicrobia
Bacteroidetes
Planctomycetes
Proteobacteria
Acidobacteria
Other
Archaeplastida
Excavata
Opisthokonta
Stramenopiles
Amoebozoa
Alveolata
Rhizaria
Other
Platyhelminthes
Tardigrada
Rotifera
Mollusca
Annelida
Arthropoda
Nematoda
Cosmopolitan
Intermediate
Specialized
E
Percent of taxa
100
75
50
25
0
Functional group
Pathogen
Mutualist
Saprotroph
Cosmopolitan
Intermediate
Specialized


A
Briza media
A. odoratum
Festuca rubra
Dactylis glomerata
Cynosurus cristatus
Ranunculus repens
Lotus corniculatus
Trifolium pratense
Lathyrus pratensis
Filipendula ulmaria
Geum rivale
Geranium sylvaticum
Geranium pratense
Rumex acetosa
Plantago lanceolata
Centaurea nigra
Leontodon hispidus
Hypochaeris radicata
Bellis perennis
Achillea millefolium
L. vulgare
Fungi
Bacteria
Protists
Metazoa
First principal component
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
B
Fungi
Bacteria
Protists
Metazoa
Soil community composition dissimilarity
Trait distance
C
Shoot traits
SLA
C content
N content
Dry matter content
Root traits
SRL
C content
N content
Dry matter content
Mean diameter
Fungi
Bacteria
Protists
Metazoa
Rho
0.100
0.075
0.050
0.025
0.000
p−value
0.02
0.04
0.06
0.08


A
B
C
Relative abundance
First principal coordinate (plants)
Other
Trifolium
Ranunculus
Poa
Lolium
Holcus
Dactylis
Centaurea
Anthoxanthum
Alopecurus
Agrostis
Fungi
Ascomycota
Basidiomycota
Mucoromycota
Other
Bacteria
Acidobacteria
Bacteroidetes
Planctomycetes
Proteobacteria
Verrucomicrobia
Other
Protists
Alveolata
Amoebozoa
Opisthokonta
Rhizaria
Stramenopiles
Other
Metazoa
Annelida
Arthropoda
Unknown
Nematoda
Other
Plant community assessment
Aboveground
Root–based
Soil–based
Relationship strength with plant community composition (ρ)
Fungi
Bacteria
Protists
Metazoa
Root–based assessment
P = 0.11
Soil–based assessment
ρ = 0.20
P = 0.001
Belowground Community composition dissimilarity
Aboveground community composition dissimilarity