The ISME Journal (2014) 8, 2231–2242

www.nature.com/ismej

ORIGINAL ARTICLE

Arbuscular mycorrhizal fungal communities are phylogenetically clustered at small scales

Sebastian Horn[1,2], Tancredi Caruso[3], Erik Verbruggen[1,2], Matthias C Rillig[1,2] and Stefan Hempel[1,2]

[1]*Institut für Biologie—Ökologie der Pflanzen, Freie Universität Berlin, Berlin, Germany;* [2]*Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), Berlin, Germany and* [3]*School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, Belfast, Northern Ireland, UK*

Next-generation sequencing technologies with markers covering the full Glomeromycota phylum were used to uncover phylogenetic community structure of arbuscular mycorrhizal fungi (AMF) associated with *Festuca brevipila.* The study system was a semi-arid grassland with high plant diversity and a steep environmental gradient in pH, C, N, P and soil water content. The AMF community in roots and rhizosphere soil were analyzed separately and consisted of 74 distinct operational taxonomic units (OTUs) in total. Community-level variance partitioning showed that the role of environmental factors in determining AM species composition was marginal when controlling for spatial autocorrelation at multiple scales. Instead, phylogenetic distance and spatial distance were major correlates of AMF communities: OTUs that were more closely related (and which therefore may have similar traits) were more likely to co-occur. This pattern was insensitive to phylogenetic sampling breadth. Given the minor effects of the environment, we propose that at small scales closely related AMF positively associate through biotic factors such as plant-AMF filtering and interactions within the soil biota.

The ISME Journal (2014) **8,** 2231–2242; doi:10.1038/ismej.2014.72; published online 13 May 2014
Subject Category: Microbial population and community ecology
Keywords: arbuscular mycorrhizal fungi; community assembly; phylogenetic clustering; small-scale; spatially explicit; 454-pyrosequencing

Introduction

Arbuscular mycorrhizal fungi (AMF) form symbiotic relationships with the majority of land plants, facilitating the uptake of nutrients from soil in exchange for plant-assimilated carbon (Smith and Read, 2008). They have important roles in ecosystem functioning through their influence on biogeochemical processes (van der Heijden *et al.*, 2008; Veresoglou *et al.*, 2012) and on the structure and productivity of plant communities (van der Heijden *et al.*, 1998; Jansa *et al.*, 2008). The question that affects predominate in structuring AMF communities remains only partially answered and detailed information on mechanisms is sparse in spite of recent advances (for example, Öpik *et al.*, 2009, 2010; Caruso *et al.*, 2012; Lekberg *et al.*, 2013).

Grasslands cover one-fourth of the Earth's land surface and harbor most of herbaceous plant diversity (Shantz, 1954). AMF are the dominant symbiotic fungi in these systems. Although several recent studies deal with AMF in grasslands (Karanika *et al.*, 2008; Verbruggen *et al.*, 2010; Stover *et al.*, 2012), most of these studies simply cataloged species or use molecular techniques that preclude in-depth characterization of AMF communities. The scale of most studies generally exceeds the likely extent of an individual AMF, and this hampers inferences about species interactions at a local scale (Gai *et al.*, 2012; González-Cortés *et al.*, 2012; Verbruggen *et al.*, 2012; Zubek *et al.*, 2012). Moreover, AMF niche space is likely to be complex because of small-scale heterogeneity of soil (Veresoglou *et al.*, 2013), and thus large-scale studies may overlook important drivers of local AMF community assembly.

Recent research has shown that niche-based processes and environmental filtering are the dominating factors in structuring AMF communities, while neutral dynamics have a minor role (Lekberg *et al.*, 2007; Dumbrell *et al.*, 2010a, 2010b). Yet, although AMF diversity in natural systems is typically measured by using taxon-based approaches, considering AMF phylogeny may provide additional information on processes impacting

Correspondence: S Hempel, Institut für Biologie—Ökologie der Pflanzen, Freie Universität Berlin, Altensteinstr. 6, Berlin 14195, Germany.
E-mail: hempel.stefan@gmail.com
Received 7 November 2013; revised 1 April 2014; accepted 2 April 2014; published online 13 May 2014

AMF community structure and functioning (Lekberg *et al.*, 2013). Indeed, a greenhouse study showed that the phylogenetic breadth of an AMF community can positively affect co-existence, and thus the resulting AMF species richness and plant performance (Maherali and Klironomos, 2007). It has also been shown that phylogenetic relatedness among AMF is positively associated with coexistence (Roger *et al.*, 2013). However, the study of the predictive power of phylogeny relative to spatial and environmental determinants of fungal community structure is in its infancy, although mechanisms such as facilitation or the bidirectional interaction between plant and AMF in forming the symbiosis may be uniquely signaled by phylogenetic patterns. In fact, phylogenetic distance can reflect trait convergence or displacement if traits are phylogenetically conserved, which implies that nonrandom phylogenetic patterns in species distribution can reflect underlying processes such as competition, interactions with the soil biota or habitat filtering (Kembel and Hubbel, 2006; Kembel *et al.*, 2010; HilleRisLambers *et al.*, 2012). For example, phylogenetic dispersion (that is, segregation) is expected to occur under competitive processes, while trait selection processes may lead to phylogenetic clustering (that is, aggregation).

AMF species identification has historically been based on spore morphology, which suffers from some significant shortcomings (Hempel *et al.*, 2007; Taylor *et al.*, 2013). Classical cloning and Sanger sequencing is costly, often preventing in-depth sequencing of environmental samples. The development of next-generation sequencing techniques (Margulies *et al.*, 2005) facilitates the assessment of AMF communities at the species level in environmental samples (Öpik *et al.*, 2009, 2010), overcoming limitations inherent to morphological or genetic fingerprinting-based identification. The development of a new AMF-specific primer-set (Krüger *et al.*, 2009, 2012) allows access to an unprecedented amount of AMF diversity data in the field, as this primer set is both highly specific to AMF and amplifies all taxa within this group (Krüger *et al.*, 2009).

Here, we assessed the role of different factors that shape the AMF community in a semi-natural grassland. We had three main questions: (1) Do environmental factors structure the AMF community? (2) How much influence do distance-based effects and stochastic events have on AMF community structure? (3) Is the AMF community phylogenetically structured?

Our hypotheses regarding the community effects of each of the three components under investigation are further described in Figure 1. In order to disentangle the explanatory power of each of these three known factors shaping community composition, we extensively sampled the dominant plant species and used a variance partitioning approach to estimate variance explained by these factors while controlling for potential co-variation.

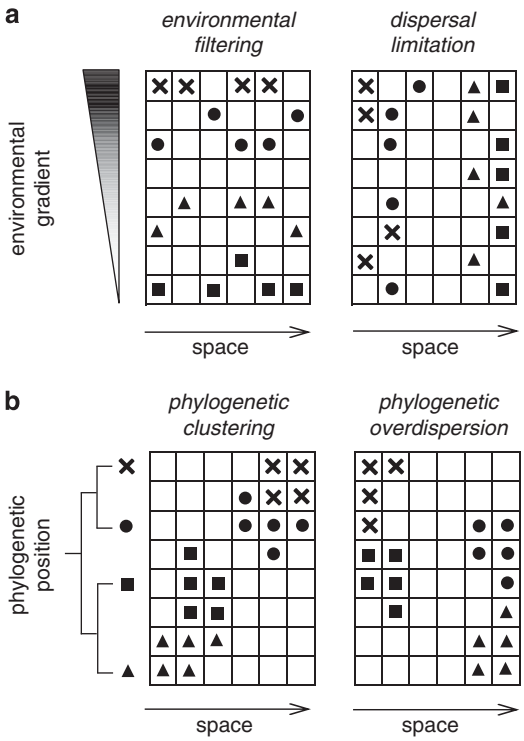

Figure 1 Proposed relationship between AMF community structure and the environmental, spatial and phylogenetic components. The symbols represent AMF communities of varying species composition. We expect the spatially structured environment (**a**) to be influencing AMF community composition either by environmental filtering or spatial processes. Environmental filtering will lead to fungal species aggregating along the environmental gradient. AMF communities in similar environments will be more similar to each other, no matter the spatial position (left diagram). Spatial processes like dispersal limitation will cause AMF communities to be more similar that are closer to each other, independent of the environmental properties in each sample (right diagram). The phylogenetic component (**b**) is expected to either cause segregation (overdispersion) or aggregation (clustering) of the AMF species co-occurrence in a sampled community. Phylogenetic clustering is expected when particular phylogenetically conserved trait values are selected in one sample over the other, and will show closely related species occurring more frequently (left diagram). Overdispersion is expected when AMF with increasingly different traits are increasingly more likely to co-occur, for example, through limiting trait similarity and/or niche partitioning, and will show less closely related species occurring more frequently (right diagram). We expect the actual data to be composed of a mixture of all the depicted effects, which will be disentangled by variance partitioning.

Materials and methods

Study area and sample collection

The grassland we studied is situated in a natural reserve at Mallnow, Lebus (Brandenburg, Germany, 52°27.778′ N, 14°29.349′ E). The reserve is known for its different types of species-rich dry grassland and has been managed by low-intensity sheep grazing for at least 500 years (Ristow *et al.*, 2011). At the beginning of October 2010, we sampled a larger plot of 15×15 m (henceforth called 'macroplot') located on the slope of a hillside. The uphill–downhill axis of the hillside where the macroplot is located is characterized by a steep soil textural

gradient from highly sandy (downhill) to sandy–loamy (uphill) soils. Geochemical analysis revealed that other soil parameters highly relevant for AMF communities, namely pH, carbon, nitrogen and plant available phosphorous (Kivlin *et al.*, 2011), strongly varied along the soil texture gradient (Supplementary Table S1). Specifically, pH, which is known to have important effects on AMF (Dumbrell *et al.*, 2010b), varied from 5.5 to 8.3.

We assessed the local AMF community in the roots and surrounding soil of *Festuca brevipila* R. Tracey. *F. brevipila* is by far the most abundant species in our macroplot (coverage >60% and in some case >80%) as well as throughout the grassland of the study area. This approach standardized the observed AMF community on an organism of wide prevalence, allowing a precise, yet general definition of the community unit: the AMF associated with the rhizosphere of the dominant grass. We used nine plots of 3×3 m equally distributed across the macroplot in order to reduce the amount of sampling necessary for capturing the whole extent of the gradient. Despite this sub-partitioning of the macroplot, we analyzed the samples across sampling locations (that is, the roots of an individual grass and its soil form the community unit), rather than based on plots. The sampling was replicated by taking soil cores (5 cm radius, 15 cm deep) centered on six randomly chosen *F. brevipila* individuals per plot, resulting in 54 (6 plants $\times$ 9 plots) sampling locations (henceforth called 'samples') in total. This sampling allowed us to detect spatial patterns within and between plots with intervals ranging from about 1 to nearly 15 m. Each soil core was thoroughly homogenized before subsampling for the different analyses. Roots were washed in Millipore water before subsequent analysis. Soil variables were measured according to the protocol provided in the Supplementary Information. To assess AMF colonization, a subsample of the roots was stained with Trypan blue, assessed at $\times$200 using the magnified gridline intersect method (McGonigle *et al.*, 1990).

DNA extraction, 454-pyrosequencing and operational taxonomic unit (OTU) delineation

We used 250 mg of each soil and washed root material per core to extract DNA using the PowerSoil DNA Isolation Kit (MoBio Laboratories Inc., Carlsbad, CA, USA) and the procedures in the manufacturer's manual. Then we created 454-pyrosequencing amplicon pools using AMF-specific primers (Krüger *et al.*, 2009) following the protocol in the Supplementary Information, which involved three PCRs of 30 cycles each. We used a primer mixture, which increased competition for target molecules, delays exponential growth of products and therefore justifies an increased cycle number, but should theoretically also lead to a reduced PCR bias toward more abundant species. As our results are consistent with the findings of other diversity studies on AMF concerning the representation of genera (Öpik *et al.*, 2010; Maherali and Klironomos, 2012), we assume no bias despite the high number of PCR cycles. Sequencing of the samples was done using the primer set LR3 and LR0R (Hofstetter *et al.*, 2002). These two primers span an area of about 900–950 bp, including the variable D1–D2–D3 region of the large sub-unit and are conserved among eukaryotes (Vilgalys and Hester, 1990), therefore preserving the diversity obtained by the AMF-specific primers. LR3 was tagged with adapter B, LR0R was fused with the adapter A and error-correcting barcode sequence (Hamady *et al.*, 2008) for the 454 run, so we sequenced from the 3′-end of our target sequence toward the start of the large sub-unit. 454-Pyrosequencing was done at Duke University Sequencing core facility (Durham, NC, USA) using the Titanium chemistry from Roche (Basel, Switzerland).

Resulting sequence sets were subjected to a denoising and clustering pipeline. Sequences were denoised using the PyroNoise approach implemented in Mothur (Schloss *et al.*, 2009). This denoising approach removes bad quality sequences, creates sequence clusters and removes chimera sequences, while being based on clustering flowgrams rather than sequences alone, which allows 454 errors to be modeled and removed naturally (Quince *et al.*, 2009, 2011). We used the standard parameters for Titanium sequencing as suggested by Quince *et al.* (2009), with a minimum flow amount of 360 and a maximum of 720. After denoising, the sequences of roots and soil were clustered using CROP. The program uses a Bayesian clustering algorithm, which addresses species delineation uncertainty better than hierarchical clustering methods because of its flexible cutoff and therefore creates significantly fewer artificial OTUs than other fixed cutoff clustering approaches (Hao *et al.*, 2011). Runtime parameters and source code from the analysis in R described below are provided in the Supplementary Information.

Owing to the nature of pyrosequencing, there were differences in read numbers for every sampling location, so we resampled the read numbers to equal amounts of 700 reads per sampling location, 350 each for root and soil subsample, using a bootstrap approach. Sampling locations with considerably lower read numbers than the resampling value (<260 per root or soil sample) were discarded (9 soil and no root samples). Based on read numbers, rarefaction curves were calculated for each root or soil sample per location. As all species accumulation curves were leveling off, no sample was excluded. Singletons were removed from all samples. The resulting OTUs were represented by an abundance matrix.

Phylogenetic tree calculation

Tree calculation was done with RAxML (Stamatakis, 2006) and BEAST (Drummond and Rambaut, 2007). We added representative sequences of an small

sub-unit–internal transcribed spacer–large sub-unit AMF reference alignment (Krüger *et al.*, 2012) to our data set to determine the phylogenetic position of our OTUs. Using the position of an OTU in a phylogenetic tree relative to reference sequence creates more reliable species estimation than just using database queries (Ross *et al.*, 2008). In order to compare the quality of our tree and to add more description to the OTUs, we annotated them using the BLAST hit with the lowest E-value. The reference alignment was first trimmed to the region present in our sequences and then used as a template in Mothur to align our OTU sequences. The two alignments were combined and further refined in MUSCLE (Edgar, 2004). We used phylogenetic trees to further refine our OTU set and remove sequences that clustered outside the Glomeromycota and are therefore likely to be erroneous or non-AMF sequences. We used two different tree calculation approaches to validate the accuracy of the obtained phylogeny. Using RAxML, we calculated 1000 rapid bootstrap trees and subsequently applied a thorough maximum likelihood analysis. BEAST was run with the Extended Bayesian Skyline Plot as a tree model (Minin *et al.*, 2008) with a chain length of 10 million generations, from which the best tree was chosen for evaluation. Trees were then inspected and edited using FigTree (Rambaut, 2012).

Phylogenetic community structure

We addressed community structure by analyzing phylogenetic diversity between the AMF communities. Using the picante package (Kembel *et al.*, 2010) in R 3.0.2 (R Core Team, 2013), we obtained two estimates of phylogenetic diversity: (1) the standardized effect size of mean pair wise distance (SES-MPD), which measures alpha-diversity, and (2) inter-community mean pair wise distance (IC-MPD), which measures beta-diversity. Phylogenetic distances between OTUs serving as input for the estimates of phylogenetic diversities were calculated using the Needleman–Wunsch implementation of Esprit (Sun *et al.*, 2009). The SES-MPD was calculated using the phylogenetic distance matrix of the OTUs plus the abundance matrix of the OTUs per sample and applied a null model randomization. The result was an net relatedness index (NRI) value for each sample, which is defined as $(-(\mathrm{MPD} - \mathrm{MPD}_{\mathrm{null}})/_{\mathrm{SD}}(\mathrm{MPD}_{\mathrm{null}}))$, where MPD is the mean pairwise phylogenetic distance among species in the sample (Kembel and Hubbel, 2006). The mean values of the NRIs of all samples of roots and soil, respectively, were then used as the alpha-diversity measure to judge the clustering or segregation of the overall AMF community. Positive NRI values are correlated with clustering, negative values with overdispersion. The null model algorithm we used was 'independentswap' with 999 randomized null communities. 'Independentswap' retains column and row totals for null model analysis of species co-occurrence (Gotelli, 2000). This approach is particularly suited to problems that concern differences in species composition, because it accounts for variations in other community attributes such as diversity and richness. Significance of the calculated NRIs was tested using *t*-test.

IC-MPD calculates pairwise phylogenetic distances of the samples, based on pairwise genetic distances between OTUs and yields a sample $\times$ sample distance matrix as a measure of beta-diversity. In order to include the IC-MPD information in a subsequent variance partitioning analysis (Legendre and Legendre, 1998), the distance matrix was subjected to a principal coordinate analysis, a commonly used tool to reduce dimensionality, which provides a measure of the amount of variance explained in the a few independent principal axes (Legendre and Legendre, 1998). The first two principal coordinate analysis axes, which represented a major of amount of total phylogenetic variation, were extracted and used as the phylogenetic explanatory variables.

Variance partitioning

The analysis of patterns in community structure was conducted in R, using the vegan (Oksanen *et al.*, 2012) and the SPACEMAKER (Dray, 2011) package and the abundance matrix obtained from processing the sequences as response matrix. Spatial information (pairwise distance between samples), log-transformed environmental data (sample pH and C, N, P, and water content, and C/N ratio) and the estimates of phylogenetic beta-diversity were used as explanatory variables.

The OTU abundance matrix was Hellinger transformed and subjected to a multivariate analysis to test for effects of spatial, environmental and phylogenetic variables influencing community variation. In variance partitioning, 'space' stands indeed for spatial autocorrelation: moran eigenvector mapping was used to factor in spatial autocorrelation at multiple scales in community variance partitioning (Dray *et al.*, 2006). This method represents a general, more powerful version of the widely used principal coordinate analysis of neighbor matrix (Borcard *et al.*, 2004), which allows testing for several types of spatial structure. Several competing spatial models are possible and the most parsimonious model is selected using a multivariate extension of the Akaike Information Criterion (Akaike, 1973). This model provides the best linear combination of eigenvectors accounting for spatial autocorrelation at multiple spatial scales and each eigenvector represents a certain scale (Dray *et al.*, 2006). We used redundancy analysis and variance partitioning to resolve the contribution of each of the factors to the total variance (Legendre and Legendre, 1998). As this was an observational study, we applied a conservative logic: unequivocal evidences of the effect of a certain factor are obtained only when

controlling for the effect of other factors. For example, shared variation between spatial and environmental factors may depend on the fact that we sampled along an environmental gradient. However, this correlation does not imply causation as linear changes in community composition can also be observed along directions where there is little environmental variation or the gradient may not structure the community directly (Legendre and Legendre, 1998). Thus, a non-spatially structured effect of the environment would suggest that communities are similar if their environments are similar, regardless of their spatial position. This is more robust evidence of independent effects of the environment in the framework of observational studies. This is also the reason why spatial autocorrelation at multiple scales needs to be accounted for in order to perform a robust test of factors that are spatially structured (Legendre and Legendre, 1998; Borcard *et al.*, 2004). In this sense, it is not the spatial variation in itself that is under investigation because this variation can be due to several possible and collinear factors that neither have been measured nor can be disentangled from measured factors. Given this logic, variance partitioning is the appropriate tool to quantify the unique contribution of the three factors investigated in this study (Borcard *et al.*, 1992). Factors were tested using a partial redundancy analysis-based permutation approach, which tests for the focal factors by controlling for the other factors (Oksanen *et al.*, 2012).

Results

Study area and sample collection

The study area was characterized by steep gradients in all measured environmental components, following roughly the uphill–downhill direction (Supplementary Figure S1). Plant available phosphorus concentration was low in most of the plots, with a range from 10.9 soil to 85.0 mg kg^{-1} (median 30.1 mg kg^{-1}). Soil C to N ratios varied between 13:1 and 43:1 (median 19:1), pH showed a range from 5.5 to 8.3 (median 7.4) and root colonization by AMF ranged from 5% to 99% (median 77.5%). The colonization was significantly correlated with an increase in loam content of the soil, which linearly correlated with water content and organic content (Supplementary Figure S2; for all environmental data see Supplementary Table S1). We did not find a correlation between root colonization and phylogenetic distance. Correlation analysis shows that most of the environmental variables were correlated with each other, confirming our prediction of a single linear environmental gradient, with the exception of soil phosphorus (Supplementary Table S3).

454-Pyrosequencing and OTU delineation

After resampling and removal of singletons and non-AMF sequences, the root data set consisted of 54 OTUs and the soil data set of 73 OTUs, with a total of 74 OTUs. Almost half of the detected OTUs (32 of 74) belonged to the genus *Glomus sensu* (Schüßler and Walker, 2010), and in most samples this was the most abundant taxon. The others were spread across all major AMF groups, spanning 10 different AMF genera (Figure 2), suggesting that our methods are capable of detecting all major lineages within AMF. The dominance of *Glomus* was also found when comparing the read numbers of each of the AMF genera instead of OTU numbers (Figure 2). The highest abundance of sequence reads to any of the OTUs was attributed to a member of the *Rhizophagus* genus, which based on BLAST annotations is

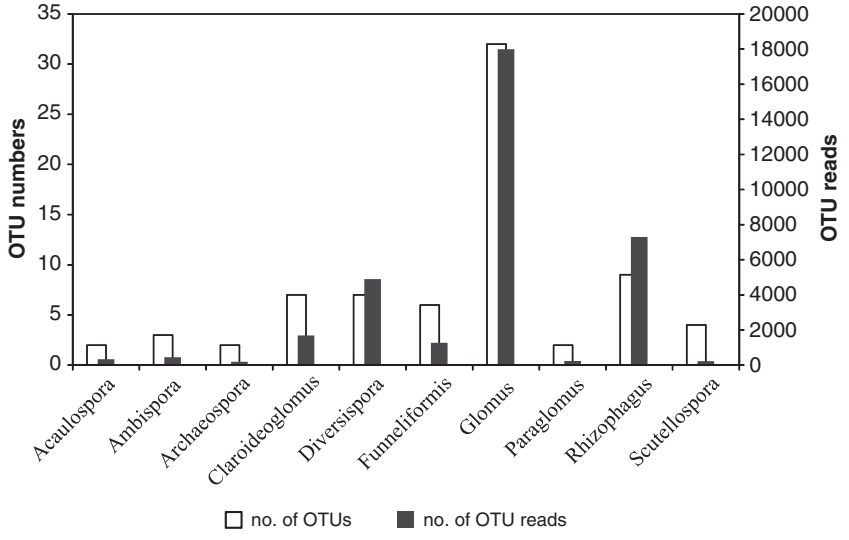

Figure 2 Number of OTUs and number of sequence reads per AMF genus. OTU numbers are represented by white bars (left y axis). OTU sequence reads are represented by dark-gray bars (right y axis).

likely the cosmopolitan *Rhizophagus irregularis* found in high abundance in several studies (Öpik *et al.*, 2006; Lekberg *et al.*, 2013).

After denoising of a total of 67 558 (roots) and 50 594 (soil) sequences with PyroNoise and the Bayesian clustering step in CROP, 301 OTUs were obtained. Further removal of OTUs was based on the elimination of singletons (164), the exclusion of OTUs that did not yield any BLAST result (33), resampling (13) and removal of non-AMF sequences identified from the trees. Our primers proved to be highly AMF specific, with only a few non-target OTUs from the Chytridiomycota phylum and other fungi (17).

If sampled sufficiently, the root community should ideally represent a subset of the soil community. We found only one OTU in the root data set, which was not part of the soil data set and which was very likely a sampling effect on the very rare OTU. The rarefaction curves (Supplementary Figure S3) showed that all the communities were leveling off or were very close to saturation. The sequences clustered well with Glomeromycota reference data (Figure 3) as published in Krüger *et al.* (2012). In general, phylogenetic position in the tree could assign many OTUs to genera that were only poorly annotated in the NCBI database (for example, 'uncultured Glomeromycetes'; Figure 3).

Phylogenetic community structure

The SES-MPD null model analysis showed significant differences from random distribution, when the abundance weighed data were used (Table 1). Mean NRIs for both root and soil data sets were positive with comparable sizes (0.27 and 0.26), which means that AMF communities contain taxa that are phylogenetically more related than expected by chance (that is, significantly clustered). In the non-abundance weighted SES-MPD indices, the trend toward clustering is still present, albeit not significant. As the number and relative abundance of OTUs was strongly biased toward the *Glomus* group (Figure 2), we split up the data into *Glomus* and non-*Glomus* OTUs and performed a separate analysis on each group. For both data sets, the significant phylogenetic clustering persisted suggesting the pattern is valid independent of whether closely or distantly related taxa are compared. In the *Glomus* data set, significance was independent of abundance with effect sizes being comparable in root and soil. In the non-*Glomus* OTU set, results were similar to the complete OTU set. The magnitude of the NRI was comparable in root and soil.

Variance partitioning and community clustering

The variance of the whole OTU set was significantly explained by spatial and phylogenetic patterns plus their combined effects (Figure 4, Supplementary Table S2). The phylogenetically structured environmental effect was very small (<0.0001) in all of the treatments, so this partition was omitted. The influence of spatial position was more important in soil than in roots when abundance data were used, while with presence–absence data phylogenetic composition was more important in soil than in roots. Effects of spatially structured environment as well as environment alone remained comparable among root and soil, as well as between abundance and presence–absence data, but in general abundance data increased the amount of variation explained.

For the *Glomus* OTU variation, the major explanatory components were again phylogeny and the spatial signal (Figure 4; Supplementary Table S2). Differences between root and soil indicated that environmental filtering is more selective in soil.

In the data set of all OTUs except *Glomus*, spatial and phylogenetic components were again the major variables contributing to explained variation (Figure 4; Supplementary Table S2) and major differences were found between root and soil. Phylogeny was a major explanatory variable, but it decreased significantly from root to soil. In the roots, the decrease in phylogenetic signal was also found in the joint effect of spatial structure and phylogeny. Finding comparable results when removing the most abundant taxon group shows that the patterns are not exclusively shaped by *Glomus* alone.

Discussion

In this study, we have been able to quantify the relative predictive power of different factors in explaining small-scale AMF community composition in a semi-natural grassland. The three main community factors under investigation were environmental drivers, spatial structure and phylogenetic distance and below we discuss each of them with regard to our three main questions.

Do environmental factors structure AMF communities?

Previous studies addressing AMF community structure and applying variance partitioning have shown the dependence of AMF on the environment (for example, Lekberg *et al.*, 2007; Dumbrell *et al.*, 2010b). In our study, the non-spatially structured environmental component explained only little of the variation in community structure. Despite our expectations that a gradient like the one we found in our field site should significantly shape a soil community, the environment was only found to be significant in the '*Glomus* only' and 'non-*Glomus*' subsets. Given that these two groups may respond slightly differently to environmental properties, this could then lead to diminished significance in the overall data set. However, environmental effects can definitely exert their effect along a gradient in a spatially structured manner, as indicated by the variance fraction accounting for spatially structured

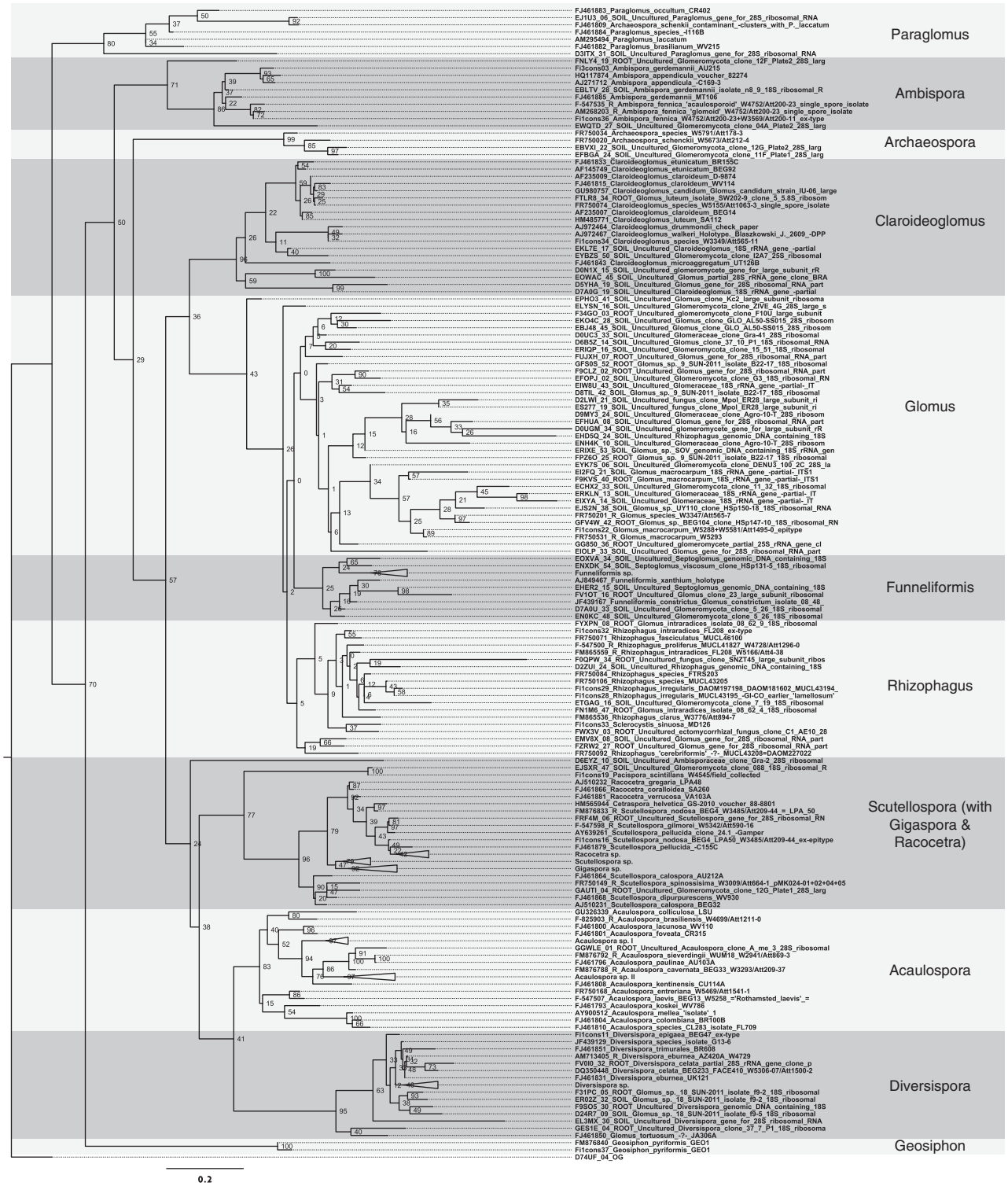

Figure 3 Maximum likelihood tree of 74 OTUs from the root and soil data set, complemented with 114 sequences from the Krüger *et al.* (2012) small sub-unit–internal transcribed spacer–large sub-unit alignment and one non-AMF outgroup ('D74UF_OG', an unidentified member of the Chytridiomycota). Tree calculation was done in RAxML. Nodes ending in triangles represent collapsed species divisions, which did not contain any of our OTUs, in order to increase visibility. Node numbers represent Bootstrap values. The node descriptions containing a 'ROOT' or 'SOIL' tag represent the OTUs defined in our study, while the other nodes represent the sequences from Krüger *et al.*

effects of the environment. Nevertheless, even if one sums that amount of variation accounted for by spatially structured and not spatially structured environmental effect, the total contribution of the environment remains small relative to the other investigated factors.

Instead, our results suggest that the AMF communities in our study area are predicted mainly by the spatial distance between samples and phylogenetic distance between OTUs, when the effect of the environment has been taken into account. As our environmental gradient was quite steep and concentrated in a small area, we have reduced confounding factors such as historical events and/or dispersal limitation, which are present in broad-scale studies. Moreover, confounding effects because of plant identity are also absent given that the observed AMF community is standardized on an organism of wide prevalence that belong to a genus (*Festuca*) that very often dominates dry grasslands worldwide. Certainly, at broader spatial scales the relative role of the various drivers of community composition may change and we stress that the local community we are here investigating must represent a local subset of the regional AMF pool. Ultimately, our local community is therefore also the result of broad-scale dispersal processes and environmental filtering processes that we cannot resolve in our study. For this same reason, we believe that, given the state-of-the-art, our approach offers a fair compromise between the ecologists' quest for general conclusions derived from large-scale fully randomized design (for example, no focal plant) and the need for the collection of robust patterns from field studies performed at local spatial, temporal and taxonomic scales. In other words, the locality of our study is showing fairly dominant nonrandom phylogenetic and spatial patterns in AMF communities: these patterns could have been neglected in the past given the multitude of factors that structure AMF assemblages from very local to regional scales. Indeed, in other studies stronger environmental effects have been found: Dumbrell *et al.* (2010b) studied an extremely pronounced pH gradient (<4–8), the study of Lekberg *et al.* (2007) focuses on agricultural fields at larger scales, and thus different community-structuring mechanisms

Table 1 *T*-test results ($P<0.05$ bolded) of the NRI of mean pairwise distance (SES-MPD) of the root and soil community matrices, including a division of the data set into *Glomus* only and non-*Glomus* OTUs

Data set	*t*	*Df*	P-*value*
All OTUs			
Root			
+abu	2.644	53	**0.011**
−abu	0.929	53	0.357
Soil			
+abu	2.031	47	**0.048**
−abu	1.156	47	0.254
Glomus *OTUs only*			
Root			
+abu	2.889	46	**0.006**
−abu	2.588	46	**0.013**
Soil			
+abu	2.750	44	**0.009**
−abu	3.227	44	**0.002**
All OTUs except Glomus			
Root			
+abu	2.994	42	**0.005**
−abu	1.479	42	0.147
Soil			
+abu	3.347	43	**0.002**
−abu	1.477	43	0.147

Abbreviations: NRI, net relatedness index; OTU, operational taxonomic unit; SES-MPD, standardized effect size of mean pairwise distance.
Either abundance data (+abu) or presence–absence data (−abu) was used when calculating the effect sizes and *P*-values.

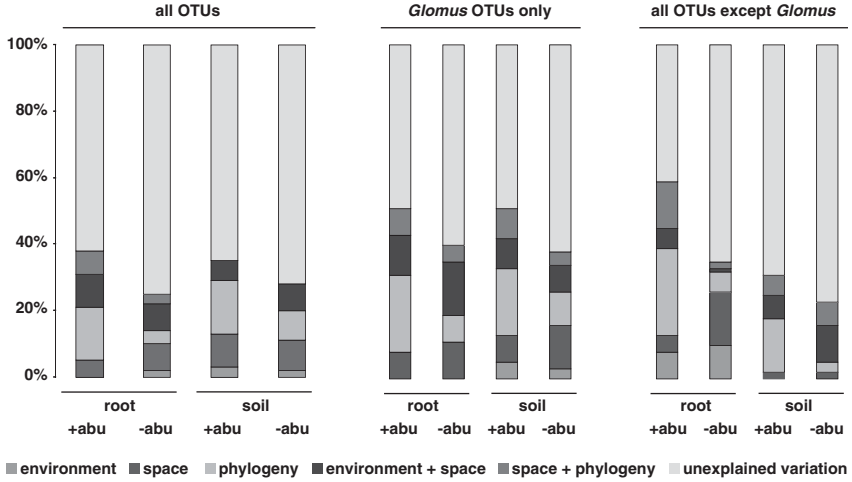

Figure 4 Percentage of variation explained by permutation tests based on redundancy analysis (RDA) and decomposition of the total variation in the community matrix into unique environmental (soil properties), spatial (geographic position) and phylogenetic (genetic distance) components. Bars of combined effects represent the shared variation between these two components. Either abundance data (+abu) or presence–absence data (−abu) was used when calculating the phylogenetic component for the variance partitioning. Values are based either on the whole OTU set or on the *Glomus* OTUs and the non-*Glomus* groups, respectively.

may operate under different ecological settings. It is also possible that significant effects of the environment on AMF may be confounded with environmental effects on the host plants (Sharma *et al.*, 2009).

The results therefore indicate that spatial and phylogenetic distance are the major representatives of the underlying processes shaping the community at small spatial scales, with soil results being similar to the roots, but more clearly separated into spatial and phylogenetic components (Figure 4; Supplementary Table S2). An explanation for the higher amount of variation explained in soil is that root communities may be strongly shaped by heavily root colonizing (that is, abundant taxa). The communities may also be more (temporally) dynamic, and thus more prone to sampling effects, that is, which plant species and when during their life cycle is sampled.

How much influence do distance-based effects and stochastic events have on AMF community structure?
We observed a large fraction of AMF community variance explained by spatial patterns after controlling for environmental factors and phylogenetic distance. Dispersal or unmeasured environmental factors as well as biotic interactions not leaving a phylogenetic signal are all possible factors behind these spatial patterns (Chang *et al.*, 2013). Given the variables we measured, it is unlikely that we missed out major environmental predictors of AMF communities. In addition to that, every measured environmental variable was spatially structured in our study area along the sampled gradient and it is therefore reasonable to assume that effects of unmeasured environmental variables are included in the variation shared by spatial eigenvectors and the measured environmental variables. On the small scale of our study, dispersal limitation is less likely but AM communities can be exceptionally patterned already at a sub-meter scale (Mummey and Rillig, 2008), so that dispersal constraints can indeed have a role at a 15-m scale. Stochastic population dynamics because of irregular, unpredictable environmental or demographic fluctuation may also contribute to these patterns. Spatial structure that is independent of environmental factors indicates that chance-events have a role in community composition although biotic interactions such as competition may also contribute to spatial patterning. Dumbrell *et al.* (2010a) suggested that chance-events could lead to a positive feedback mechanism on any taxon in the community, which could be random and self-reinforcing. This hypothesis could explain a diminished environmental signal and a strong spatial patterning. Regardless of the contribution of stochastic effects, the significant phylogenetic structure of the assemblages shows that AMF communities are also significantly shaped by deterministic processes.

Is the AMF community phylogenetically structured?
We find that phylogenetic distance can account for a relative large and statistically significant fraction of AMF community: AMF communities consist of taxa that are more related than expected by chance. This can be an effect of at least three processes: convergence via habitat filtering because taxa that are similar in traits respond in a similar way to environmental factors; or plant–AMF interactions are such that the focal plant selects phylogenetically clustered AMF assemblages. Third, fungal interactions with the soil biotic community (for example, arthropods) could create interactions that support assemblages of conserved traits: the selected AMF are those that share traits that allow them to coexist. Whichever is the cause, the effect propagates to the soil AMF assemblage and seems even stronger in some cases in the soil than in the roots. Given that the soil abiotic environment has little effect on AMF, especially when controlling for spatial autocorrelation, our results suggest that biotic interactions are more likely to be responsible for the AMF phylogenetic community pattern, although we cannot completely rule out environmental filtering as one of the source of the observed phylogenetic signal.

In AMF, phylogenetic community patterns can inform on assembly processes (HilleRisLambers *et al.*, 2012) because AMF traits are phylogenetically conserved (Powell *et al.*, 2009). The fact that phylogenetic clustering was more intense when abundance was taken into account suggests that taxa within the most abundant group, *Glomus*, share traits that allow them to coexist. This coexistence can take place because of similar, positive interactions with the host: if the host plant selects for a particular set of conserved AMF traits from a pool that varies from one place to the other, this will result in higher clustering than expected by chance. Besides this process, the neighboring plant community of our focal species could also have a role in determining phylogenetic patterns in the AMF community: analyzing the neighboring plant community of the *F. brevipila* plants showed that significant plant–plant interactions contribute to plant community composition in close proximity of *F. brevipila* (Horn *et al.*, unpublished), and this could in turn also influence the AMF communities of the focal plant (Hausmann and Hawkes, 2009), but it is not straightforward what the effect would be in terms of expected phylogenetic pattern (clustering vs dispersion). Our results are similar to those of Roger *et al.* (2013), who found closely related AMF to be more likely to coexist, presumably because of lack of competitive exclusion. This counterintuitive agreement between studies appears to indicate a general pattern and warrants future study. It may indicate that closely related AMF are similar in traits that are favored by plants (because of spatial-temporal dynamics), and that this is not offset by competition for root or soil space because competition should reduce phylogenetic clustering if traits

involved in the competition processes are phylogenetically conserved.

Other members of the plant microbiome have been shown to exhibit similar community patterns (that is, phylogenetic clustering) as we find here for AMF, for instance, rhizobia (Sachs *et al.*, 2009). Facilitative interactions between fungi have been shown in ectomycorrhiza (Shaw *et al.*, 1995; Koide *et al.*, 2005), ericoid mycorrhiza species (Gorzelak *et al.*, 2012) and have been recently indicated for AMF as well (Thonar *et al.*, 2014). Facilitation between closely related AMF as well as antagonism between distantly related taxa would ultimately result in a phylogenetically clustered AMF community. Only more mechanistic, experimental studies will in the future tell which of the proposed mechanisms contribute to community phylogenetic clustering in AMF.

Conclusions

Here we report that in AMF communities spatial and phylogenetic patterns independent of environmental factors appear to be a major source of community variation even at the small scale of this study, which suggests that environmentally independent and even stochastic events can deeply affect AMF assemblages already at fairly small (1–10 m) scales. AMF communities are strongly structured in terms of phylogenetic relationships between fungi as evidenced from their phylogenetic clustering. Given the weak effects of the environment, we propose that this pattern is explained by direct or indirect positive interactions among fungi and their biotic environment. Phylogenetic clustering was observed both in the roots and the soil and in some cases phylogeny explained more variation in soil. In order to elucidate the mechanisms behind these patterns, the study of fungal traits offers a promising research avenue in microbial ecology.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Janis Antonovics for his insightful and inspiring comments on the manuscript, as well as the Deutsche Forschungsgemeinschaft (DFG) for their generous financial support.

References

Akaike H. (1973). Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki BF (eds) *Second International Symposium on Information Theory*. Academiai Kiado: Budapest, pp 267–281.

Borcard D, Legendre P, Drapeau P. (1992). Partialling out the spatial component of ecological variation. *Ecology* **73**: 1045–1055.

Borcard D, Legendre P, Avois-Jacquet C, Tuomisto H. (2004). Dissecting the spatial structure of ecological data at multiple scales. *Ecology* **85**: 1826–1832.

Caruso T, Hempel S, Powell JR, Barto EK, Rillig MC. (2012). Compositional divergence and convergence in arbuscular mycorrhizal fungal communities. *Ecology* **93**: 1115–1124.

Chang L-W, Zeleny D, Li C-F, Chiu S-T, Hsieh C-F. (2013). Better environmental data may reverse conclusions about niche- and dispersal-based processes in community assembly. *Ecology* **94**: 2145–2151.

Dray S, Legendre P, Peres-Neto PR. (2006). Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecol Model* **196**: 483–493.

Dray S. (2011). SpacemakeR: Spatial modelling.

Drummond AJ, Rambaut A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**: 214.

Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH. (2010a). Idiosyncrasy and overdominance in the structure of natural communities of arbuscular mycorrhizal fungi: is there a role for stochastic processes? *J Ecol* **98**: 419–428.

Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH. (2010b). Relative roles of niche and neutral processes in structuring a soil microbial community. *ISME J* **4**: 337–345.

Edgar RC. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.

Gai JP, Tian H, Yang FY, Christie P, Li XL, Klironomos JN. (2012). Arbuscular mycorrhizal fungal diversity along a Tibetan elevation gradient. *Pedobiologia* **55**: 145–151.

González-Cortés JC, Vega-Fraga M, Varela-Fregoso L, Martínez-Trujillo M, Carreón-Abud Y, Gavito ME. (2012). Arbuscular mycorrhizal fungal (AMF) communities and land use change: the conversion of temperate forests to avocado plantations and maize fields in central Mexico. *Fungal Ecol* **5**: 16–23.

Gorzelak MA, Hambleton S, Massicotte HB. (2012). Community structure of ericoid mycorrhizas and root-associated fungi of Vaccinium membranaceum across an elevation gradient in the Canadian Rocky Mountains. *Fungal Ecol* **5**: 36–45.

Gotelli NJ. (2000). Null model analysis of species co-occurrence patterns. *Ecology* **81**: 2606–2621.

Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R. (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Meth* **5**: 235–237.

Hao XL, Jiang R, Chen T. (2011). Clustering 16S rRNA for OTU prediction: a method of unsupervised Bayesian clustering. *Bioinformatics* **27**: 611–618.

Hausmann NT, Hawkes CV. (2009). Plant neighborhood control of arbuscular mycorrhizal community composition. *New Phytol* **183**: 1188–1200.

Hempel S, Renker C, Buscot F. (2007). Differences in the species composition of arbuscular mycorrhizal fungi in spore, root and soil communities in a grassland ecosystem. *Environ Microbiol* **9**: 1930–1938.

HilleRisLambers J, Adler PB, Harpole WS, Levine JM, Mayfield MM. (2012). Rethinking community

assembly through the lens of coexistence theory. *Annu Rev Ecol Evol Syst* **43**: 227–248.

Hofstetter V, Clémençon H, Vilgalys R, Moncalvo J-M. (2002). Phylogenetic analyses of the Lyophylleae (Agaricales, Basidiomycota) based on nuclear and mitochondrial rDNA sequences. *Mycol Res* **106**: 1043–1059.

Jansa J, Smith FA, Smith SE. (2008). Are there benefits of simultaneous root colonization by different arbuscular mycorrhizal fungi? *New Phytol* **177**: 779–789.

Karanika ED, Voulgari OK, Mamolos AP, Alifragis DA, Veresoglou DS. (2008). Arbuscular mycorrhizal fungi in northern Greece and influence of soil resources on their colonization. *Pedobiologia* **51**: 409–418.

Kembel SW, Hubbel SP. (2006). The phylogenetic structure of a neotropical forest tree community. *Ecology* **87**: 86–99.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD *et al.* (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463–1464.

Kivlin SN, Hawkes CV, Treseder KK. (2011). Global diversity and distribution of arbuscular mycorrhizal fungi. *Soil Biol Biochem* **43**: 2294–2303.

Koide RT, Xu B, Sharda J, Lekberg Y, Ostiguy N. (2005). Evidence of species interactions within an ectomycorrhizal fungal community. *New Phytol* **165**: 305–316.

Krüger M, Stockinger H, Krüger C, Schüßler A. (2009). DNA-based species level detection of Glomeromycota: one PCR primer set for all arbuscular mycorrhizal fungi. *New Phytol* **183**: 212–223.

Krüger M, Krüger C, Walker C, Stockinger H, Schüßler A. (2012). Phylogenetic reference data for systematics and phylotaxonomy of arbuscular mycorrhizal fungi from phylum to species level. *New Phytol* **193**: 970–984.

Legendre P, Legendre L. (1998). *Numerical Ecology*. Elsevier Science: Amsterdam.

Lekberg Y, Koide RT, Rohr JR, Aldrich-Wolfe L, Morton JB. (2007). Role of niche restrictions and dispersal in the composition of arbuscular mycorrhizal fungal communities. *J Ecol* **95**: 95–105.

Lekberg Y, Gibbons SM, Rosendahl S, Ramsey PW. (2013). Severe plant invasions can increase mycorrhizal fungal abundance and diversity. *ISME J* **7**: 1424–1433.

Maherali H, Klironomos JN. (2007). Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science* **316**: 1746–1748.

Maherali H, Klironomos JN. (2012). Phylogenetic and trait-based assembly of arbuscular mycorrhizal fungal communities. *PLoS One* **7**: e36695.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA *et al.* (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376–380.

McGonigle TP, Miller MH, Evans DG, Fairchild GL, Swan JA. (1990). A new method which gives an objective measure of colonization of roots by vesicular-arbuscular mycorrhizal fungi. *New Phytol* **115**: 8.

Minin VN, Bloomquist EW, Suchard MA. (2008). Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol Biol Evol* **25**: 1459–1471.

Mummey DL, Rillig MC. (2008). Spatial characterization of arbuscular mycorrhizal fungal molecular diversity at the submetre scale in a temperate grassland. *FEMS Microbiol Ecol* **64**: 260–270.

Oksanen J, Guillaume Blanchet F, Kindt R, Legendre P, Minchin PR, O'Hara RB *et al.* (2012). Vegan: community ecology package.

Öpik M, Moora M, Liira J, Zobel M. (2006). Composition of root-colonizing arbuscular mycorrhizal fungal communities in different ecosystems around the globe. *J Ecol* **94**: 778–790.

Öpik M, Metsis M, Daniell TJ, Zobel M, Moora M. (2009). Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreonemoral forest. *New Phytol* **184**: 424–437.

Öpik M, Vanatoa A, Vanatoa E, Moora M, Davison J, Kalwij JM *et al.* (2010). The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytol* **188**: 223–241.

Powell JR, Parrent JL, Hart MM, Klironomos JN, Rillig MC, Maherali H. (2009). Phylogenetic trait conservatism and the evolution of functional trade-offs in arbuscular mycorrhizal fungi. *Proc R Soc Lond B Biol Sci* **276**: 4237–4245.

Quince C, Lanzen A, Curtis TP, Davenport RJ, Hall N, Head IM *et al.* (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat Meth* **6**: 639–U627.

Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ. (2011). Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* **12**: 38.

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna.

Rambaut A. (2012). FigTree Tree Figure Drawing Tool, 1.4.0 edn.

Ristow M, Rohner M-S, Heinken T. (2011). Die Oderhänge bei Mallnow und Lebus. *Tuexenia*.

Roger A, Colard A, Angelard C, Sanders IR. (2013). Relatedness among arbuscular mycorrhizal fungi drives plant growth and intraspecific fungal coexistence. *ISME J* **7**: 2137–2146.

Ross HA, Murugan S, Li WLS. (2008). Testing the reliability of genetic methods of species identification via simulation. *Syst Biol* **57**: 216–230.

Sachs JL, Kembel SW, Lau AH, Simms EL. (2009). *In situ* phylogenetic structure and diversity of wild bradyrhizobium communities. *Appl Environ Microbiol* **75**: 4727–4735.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Schüßler A, Walker C. (2010). *The Glomeromycota: A Species List with New Families and New Genera*. Createspace: Gloucester.

Shantz HL. (1954). The place of grasslands in the Earth's cover. *Ecology* **35**: 3.

Sharma D, Kapoor R, Bhatnagar AK. (2009). Differential growth response of Curculigo orchioides to native arbuscular mycorrhizal fungal (AMF) communities varying in number and fungal components. *Eur J Soil Biol* **45**: 328–333.

Shaw TM, Dighton J, Sanders FE. (1995). Interactions between ectomycorrhizal and saprotrophic fungi on agar and in association with seedlings of Lodgepole Pine (Pinus-Contorta). *Mycol Res* **99**: 159–165.

Smith SE, Read DJ. (2008). *Mycorrhizal Symbiosis*, 3rd edn. Academic Press: London.

Stamatakis A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.

Stover HJ, Thorn RG, Bowles JM, Bernards MA, Jacobs CR. (2012). Arbuscular mycorrhizal fungi and vascular plant species abundance and community structure in tallgrass prairies with varying agricultural disturbance histories. *Appl Soil Ecol* **60**: 61–70.

Sun YJ, Cai YP, Liu L, Yu FH, Farrell ML, McKendree W *et al.* (2009). ESPRIT: estimating species richness using large collections of 16S rRNA pyrosequences. *Nucleic Acids Res* **37**: e76.

Taylor A, Walker C, Bending GD. (2013). Dimorphic spore production in the genus Acaulospora. *Mycoscience* **55**: 1–4.

Thonar C, Frossard E, Smilauer P, Jansa J. (2014). Competition and facilitation in synthetic communities of arbuscular mycorrhizal fungi. *Mol Ecol* **23**: 733–746.

van der Heijden MGA, Boller T, Wiemken A, Sanders IR. (1998). Different arbuscular mycorrhizal fungal species are potential determinants of plant community structure. *Ecology* **79**: 2082–2091.

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

Verbruggen E, Roling WFM, Gamper HA, Kowalchuk GA, Verhoef HA, van der Heijden MGA. (2010). Positive effects of organic farming on below-ground mutualists: large-scale comparison of mycorrhizal fungal communities in agricultural soils. *New Phytol* **186**: 968–979.

Verbruggen E, van der Heijden MGA, Weedon JT, Kowalchuk GA, Roling WFM. (2012). Community assembly, species richness and nestedness of arbuscular mycorrhizal fungi in agricultural soils. *Mol Ecol* **21**: 2341–2353.

Veresoglou DS, Caruso T, Rillig MC. (2013). Modelling the environmental and soil factors that shape the niches of two common arbuscular mycorrhizal fungal families. *Plant Soil* **368**: 507–518.

Veresoglou SD, Chen BD, Rillig MC. (2012). Arbuscular mycorrhiza and soil nitrogen cycling. *Soil Biol Biochem* **46**: 53–62.

Vilgalys R, Hester M. (1990). Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several Cryptococcus species. *J Bacteriol* **172**: 4238–4246.

Zubek S, Stefanowicz AM, Błaszkowski J, Niklińska M, Seidler-Łożykowska K. (2012). Arbuscular mycorrhizal fungi and soil microbial communities under contrasting fertilization of three medicinal plants. *Appl Soil Ecol* **59**: 106–115.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)