New Phytologist

The relative importance of ecological drivers of arbuscular mycorrhizal fungal distribution varies with taxon phylogenetic resolution

Julien Roy[1,2], Florent Mazel[3], Moisés A. Sosa-Hernández[1,2], Juan F. Dueñas[1,2], Stefan Hempel[1,2], Lucie Zinger[4] and Matthias C. Rillig[1,2]

[1]Institut für Biologie, Ökologie der Pflanzen, Freie Universität Berlin, D-14195, Berlin, Germany; [2]Berlin–Brandenburg Institute of Advanced Biodiversity Research (BBIB), D-14195, Berlin, Germany; [3]Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; [4]Ecole Normale Supérieure, CNRS, Inserm, Institut de Biologie de l'Ecole Normale Supérieure (IBENS), PSL Research University, F-75005, Paris, France

Author for correspondence:
Julien Roy
Tel: +49 30 838-53144
Email: roy.julien@live.fr

Received: *31 May 2019*
Accepted: *23 July 2019*

New Phytologist (2019) **224:** 936–948
doi: 10.1111/nph.16080

Key words: amplicon sequence variants (ASVs), arbuscular mycorrhizal fungi, biogeography, environmental filtering, exact sequence variants (ESVs), operational taxonomic units (OTUs), phylogenetic scale, taxonomic resolution.

Summary

- The phylogenetic depth at which arbuscular mycorrhizal (AM) fungi harbor a coherent ecological niche is unknown, which has consequences for operational taxonomic unit (OTU) delineation from sequence data and the study of their biogeography.
- We tested how changes in AM fungi community composition across habitats (beta diversity) vary with OTU phylogenetic resolution. We inferred exact sequence variants (ESVs) to resolve phylotypes at resolutions finer than provided by traditional sequence clustering and analyzed beta diversity profiles up to order-level sequence clusters.
- At the ESV level, we detected the environmental predictors revealed with traditional OTUs or at higher genetic distances. However, the correlation between environmental predictors and community turnover steeply increased at a genetic distance of *c*. 0.03 substitutions per site. Furthermore, we observed a turnover of either closely or distantly related taxa (respectively at or above 0.03 substitutions per site) along different environmental gradients.
- This study suggests that different axes of AM fungal ecological niche are conserved at different phylogenetic depths. Delineating AM fungal phylotypes using DNA sequences should screen different phylogenetic resolutions to better elucidate the factors that shape communities and predict the fate of AM symbioses in a changing environment.

Introduction

Arbuscular mycorrhizal (AM) fungi are a widespread group of plant symbionts (Brundrett & Tedersoo, 2018) belonging to the subphylum Glomeromycotina (Spatafora *et al.*, 2016), and play a key role in ecosystem functioning (Powell & Rillig, 2018). To understand how organisms interact with their environments, it is necessary to adequately group individuals into a biological entity that harbors a coherent ecological niche or function. Descriptions of fungal communities are now largely based on DNA sequencing data, where the sequences obtained are usually clustered into operational taxonomic units (OTUs; Lindahl *et al.*, 2013); that is, groups of similar sequences at a given percentage of sequence similarity. However, the appropriate sequence similarity level (i.e. the phylogenetic resolution) at which these OTUs should be defined remains debated, despite renewed efforts to appropriately delineate evolutionary and/or ecologically coherent units in AM fungi (Powell *et al.*, 2011; Lekberg *et al.*, 2014). This is because, in AM fungi, we have still little understanding of the ecological variation (and underlying trait variation) at different levels of phylogenetic divergence (e.g. intra- vs interspecific genetic variation; Hazard & Johnson, 2018).

There is now compelling evidence that the community composition of AM fungi is interactively influenced by different abiotic and biotic drivers (e.g. soil pH and ecosystem type; Davison *et al.*, 2015). According to classic community ecology theory (Vellend, 2016), these drivers filter organisms from a regional pool based on their functional traits (e.g. capacity to associate with a given group of plants, tolerance to acidic soils). However, trait information is often unavailable for AM fungi, for which we often only have access at best to the organisms' identities (e.g. OTUs) and their distribution. In this case, the phylogenetic depth at which these ecological traits are conserved is pivotal for elucidating aspects of the environment that drive organism distribution (Martiny *et al.*, 2015). For instance, consider two closely related phylotypes (a1 and a2) adapted to colonize grasslands and two other closely related phylotypes (b1 and b2) adapted to colonize forests. Phylotypes a1 and a2 (and b1 and b2) are found in the same broad habitat but occur in different microhabitats: phylotypes a1 and b1 are present in acidic soils, whereas phylotypes

a2 and b2 are found in alkaline soils. Because the habitat preference (grasslands/forests) is conserved at the level of broad clades, its correlation with organism distribution will be strong at this broad phylogenetic resolution but detectable to a lesser extent at the phylotype level. By contrast, lumping the two fine-scale phylotypes into the broader clade will obscure soil pH effects. In summary, studying community composition at one phylogenetic resolution (typically at the species or one OTU level) most probably gives only partial information on the relative importance of the drivers of organism distribution (Hanson *et al.*, 2012). These drivers could be identified by screening the turnover of AM fungal communities at different OTU phylogenetic resolutions (Groussin *et al.*, 2017).

The beta diversity patterns (i.e. similarity in community composition across spatial and/or environmental gradients) of AM fungi are likely to change with OTU phylogenetic resolution because their functional traits are conserved at different phylogenetic depths. Some traits are similar among closely related species and others are not. For example, spore size and amount produced, as well as investment in extra- or intraradical mycelium, appear conserved at the family level (Hart & Reader, 2002; Maherali & Klironomos, 2007; Powell *et al.*, 2009; Chagnon *et al.*, 2013). Likewise, family- or order-level community composition is well predicted by spatial variation in soil pH, soil phosphorus (P) and nitrogen (N) content, or soil depth (Camenzind *et al.*, 2014; Rodríguez-Echeverría *et al.*, 2017; Roy *et al.*, 2017; Sosa-Hernández *et al.*, 2018a; Stürmer *et al.*, 2018a; Treseder *et al.*, 2018), supporting a phylogenetic conservatism of niche and traits at a coarse phylogenetic resolution. On the other hand, AM fungi exhibit extensive genetic variation within morphologically (spore-based) defined species, which can affect compatibility with their host plant (Angelard *et al.*, 2014) and traits such as extraradical hyphal density (Munkvold *et al.*, 2004; Koch *et al.*, 2006; Mensah *et al.*, 2015): the variability in these traits and the consequences for host performance were higher within than between AM fungal species. Accordingly, though AM fungal community composition has been shown to vary consistently across land-use management or vegetation types at a regional scale regardless of OTU phylogenetic resolution, the association with a particular plant community type at the local scale was better resolved at a fine OTU phylogenetic resolution (Lekberg *et al.*, 2014; Powell and Sikes 2014). These observations suggest that screening multiple phylogenetic resolutions, ideally both at the inter- and intraspecific levels (Johnson *et al.*, 2012; Sanders & Rodriguez, 2016), would help to better understand the ecology of AM fungi.

Ecological variation at the intraspecific level is challenging to infer from DNA sequence data owing to the difficulties in differentiating genuine sequences from PCR and sequencing artifacts. These latter are often corrected by clustering sequences at a given similarity threshold (Lindahl *et al.*, 2013; Hart *et al.*, 2015), which is often a compromise between the rate of errors produced by molecular techniques and biological variation, hence resulting in the loss of intraspecific variability. New bioinformatics methods now allow better discriminating genuine sequences from artifactual ones. These genuine sequences are termed exact sequence variants (ESVs, also known as amplicon sequence variants) and may diverge from one another by differences as subtle as one nucleotide (Eren *et al.*, 2015; Callahan *et al.*, 2016; Edgar, 2016). Inferring ESVs has advantages over classical OTUs. The first advantage is because ESVs represent a fundamental unit whereas traditional OTUs are not – a genuine DNA sequence that is comparable across studies (Callahan *et al.*, 2017). The second advantage is because ESVs keep valuable genetic information that is lost in traditional OTUs (Selosse *et al.*, 2016), thereby increasing the phylogenetic resolution of observation. However, we still do not know whether the finest phylogenetic resolution (i.e. ESVs) is useful for inferring ecological processes in AM fungal communities, or if it only increases noise-to-signal ratio.

We here inferred ESVs to resolve phylotypes at resolutions finer than provided by traditional sequence clustering, and we implemented a phylogenetic decomposition of beta diversity to test whether and how beta diversity patterns change across phylogenetic resolutions from the ESV level to order-level sequence clusters. We used two datasets of AM fungal large-subunit (LSU) ribosomal DNA (rDNA) sequences generated with an Illumina MiSeq platform: a chronosequence of 52 yr of agricultural recultivation after open-cast mining (Roy *et al.*, 2017; hereafter '*Chronosequence*') and a soil depth diversity survey in an agricultural field (Sosa-Hernández *et al.*, 2018a; hereafter '*Soil depth*'), both in western Germany. Each dataset spanned variations in several aspects of the environment that likely impose selection on traits that are conserved at different phylogenetic depths. Given the variable degree of phylogenetic trait conservatism in AM fungi, we hypothesize that, depending on the environmental gradient considered, we will observe a turnover of either closely or distantly related taxa.

Materials and Methods

Dataset description

The *Chronosequence* dataset consists of 10 fields spanning a recultivation gradient of 52 yr since open-cast mining near Jülich, Germany. Briefly, the recultivation process starts from a mix of old agricultural soil and the overburden loess layer from where lignite is extracted. It then consists of three phases: (1) 3 yr of soil restoration with alfalfa, (2) 2 yr of conversion to conventional agriculture, and (3) implementation of conventional agriculture by local farmers. In each field, five replicated soil cores were collected. The *Soil depth* dataset was obtained from an agricultural field near Bonn, Germany. It consists of soil samples at depths of 10–30 cm and 60–75 cm, and distributed in three soil compartments: the rhizosphere (soil directly influenced by roots), drilosphere (soil directly influenced by earthworms), and bulk soil, replicated three times. The same protocol (including primers, PCR conditions, and sequencing platform) was used to characterize the composition of the AM fungal communities in both datasets – see Roy *et al.* (2017) and Sosa-Hernández *et al.* (2018a) for the detailed protocols. Amplicons of the LSU ribosomal (rRNA) gene were sequenced on an Illumina MiSeq platform using the 2 × 300 paired-end chemistry. Raw reads and sample

metadata are freely available at https://doi.org/10.5061/dryad.t096s (*Chronosequence* dataset) and https://doi.org/10.6084/m9.figshare.8850119 (*Soil depth* dataset).

ESV inference

We used DADA2 (Callahan *et al.*, 2016) in R (R Core Team 2017) to obtain denoised, chimera-free, nonsingleton ESVs. Primers were removed and the forward and reverse reads were trimmed at 270 bp and 220 bp, respectively, due to the base-call quality degradation in the 3′ end of the reads. Sequence trimming at these positions permits a minimal overlap of 20 bp between the forward and reverse paired-end reads, which is necessary for their assembly, and an additional 70 bp of sequence length variation, which corresponds to that observed in the reference sequences of Krüger *et al.* (2012) for our amplified LSU region (length of reference sequences is 349–419 bp). Sequences with ambiguous bases were excluded, and the maximum expected number of errors was set to two and five for forward and reverse reads, respectively. We chose this threshold, instead of one, the default parameter of the USEARCH OTU pipeline (Roy *et al.*, 2017), because DADA2 infers genuine sequences using an error-rate model. Setting a higher threshold of expected errors allowed keeping a substantial sequencing depth equivalent to that of the USEARCH procedure. ESVs were inferred on a sample basis. Singleton ESVs were considered as artifactual and removed. ESVs for which the sequence corresponded to subsequences of two more abundant sequences were considered chimeras and removed. Taxonomic identity of ESVs was inferred using BLAST against reference Glomeromycotina sequences (Krüger *et al.*, 2012). Taxonomic assignment was done for queries having at least 90% of sequence similarity, 90% of sequence coverage, and a minimum *e*-value of 10^{-50} against their closest reference. The sequences not fulfilling these criteria were considered nonGlomeromycotina and were excluded from the analysis.

Additional filtering of putative PCR errors

Despite our conservative sequence data curation and the low proportion of error sequences retrieved with DADA2 (Callahan *et al.*, 2016), our dataset could still contain PCR and sequencing errors. Indeed, DADA2 assumes independent errors in the model, but some errors tend to be repeatable and can co-occur with the true sequences from which they originate (Coissac *et al.*, 2012). The co-occurrence of errors with their genuine sequence, besides inflating diversity estimate, can hence artificially inflate the correlation between phylogenetic beta diversity and environmental gradients. To correct for this potential bias, we used LULU in R (Frøslev *et al.*, 2017). We used the default parameters based on the co-occurrence and relative abundance relationships between closely related ESVs. We considered as a PCR error the sequences that are consistently at a lower abundance than their true sequence relative across all samples. Putative PCR error sequences were removed without summing their read counts to the inferred true sequence relative. LULU curation served two purposes: (1) to examine the robustness of the results to putative repeatable sequence errors; (2) to estimate lower and higher boundaries of phylogenetic resolution at which biological units correlate with the environment in case true diversity is removed.

Phylogenetic tree reconstruction

ESVs were aligned to the reference alignment of partial small subunit (SSU)–internal transcribed spacer (ITS)–LSU rDNA Glomeromycotina sequences (Krüger *et al.*, 2012), which we used as a phylogenetic backbone. The alignments were performed using MAFFT (Katoh *et al.*, 2002) with the *-addfragments* option to align short sequences to an alignment of long sequences. A phylogenetic tree was built with RAxML (Stamatakis, 2014) from these alignments, containing both ESVs and reference sequences, by conducting a bootstrap analysis (100 bootstraps) and searching for the best-scoring maximum-likelihood tree under a GTRGAMMA model of nucleotide substitution. The nucleotide sequence divergence expressed in mean substitution per site (subs/site) is hereafter referred to as genetic distance. The phylogenetic tree was analyzed and visualized in R using APE (Paradis *et al.*, 2004) and GGTREE (Yu *et al.*, 2017).

Decomposition of beta diversity across phylogenetic resolutions

Beta diversity through time In order to assess beta diversity patterns at different OTU phylogenetic resolutions, we used the beta diversity through time (BDTT) approach (Groussin *et al.*, 2017) and clustered ESVs into OTUs by steps of nucleotide sequence divergence (hereafter referred to as OTU_{BDTT}; ESVs are OTU_{BDTT0s}) across the phylogenetic tree. The set of samples in which an OTU_{BDTT} occurs is defined as the union of the samples where each ESV belonging to this OTU_{BDTT} is found. The abundance of the OTU_{BDTT} is the sum of the count of its respective ESVs. For each genetic distance, a new OTU contingency table was created. The function was adapted to cluster tree edges in nonultrametric trees to maintain cluster monophyly and, as such, uses directly the estimated evolutionary distances. This method will tell us at which phylogenetic resolution the OTUs best correlate with the different environmental gradients by testing whether the genetic distance used to define phylotypes is too narrow (i.e. there is noise, and it is therefore necessary to lump finer OTUs into broader OTUs to detect a signal) or has insufficient resolution to detect a correlation with the environmental (i.e. it is necessary to split broader OTUs into finer OTUs to detect a signal).

UPARSE OTU clustering We used also UPARSE OTU clustering, a common approach in AM fungal ecology (Horn *et al.*, 2014; Roy *et al.*, 2017; Sosa-Hernández *et al.*, 2018a), to test the consistency of beta diversity patterns between different methods. Sequence processing followed Roy *et al.* (2017). Briefly, paired-end reads were merged and quality filtered in USEARCH (Edgar 2010). Sequences were dereplicated and singletons removed. Sequences that did not overlap our expected region in the reference Glomeromycotina sequences (Krüger *et al.*, 2012) were

removed using MOTHUR (Schloss *et al.*, 2009; Kozich *et al.*, 2013; Schloss, 2013). Sequences were clustered into OTUs using UPARSE (hereafter referred to as OTU_{UPARSE}; Edgar, 2013) at sequence similarity ranging from 99% to 70% to test the robustness of our phylogenetic decomposition of beta diversity to OTU clustering methods. LULU curation was run, as described earlier, to discard putative PCR errors.

Mean nearest taxon distance and mean pairwise distance Both the BDTT and the UPARSE clustering rely on fixed genetic distance thresholds, and hence do not take into account the different rates of trait evolution across clades, which are nonnegligible (e.g. Öpik *et al.*, 2010; Powell *et al.*, 2011; Lekberg *et al.*, 2014). To consider this aspect, we additionally used beta diversity metrics that measure the shared evolutionary history between communities (measured in branch length) to further estimate the depth of phylogenetic turnover across different environmental gradients without an OTU clustering step. To this end, we calculated the mean nearest taxon distance (MNTD) and the mean pairwise distance (MPD) using PICANTE in R (Kembel *et al.*, 2010). MNTD and MPD are averaged measures of the phylogenetic distance between pairs of taxa (ESVs here) drawn from two communities at fine (MNTD) or coarse (MPD) phylogenetic resolution.

Taxonomic resolution To test whether taxonomic annotations reflect phylogenetic OTU clustering and phylogenetic beta diversity measures, we used the taxonomic annotations of ESVs to cluster them (and to sum their abundances) at taxonomic ranks from the species to order level.

Statistical analyses

Estimating the correlation between beta diversity at different OTU phylogenetic resolutions and the environmental predictors To assess how the beta diversity patterns at different OTU phylogenetic resolutions correlated to environmental gradients, we used nonparametric multivariate ANOVA (PERMANOVA; Anderson, 2001) in VEGAN (Oksanen *et al.*, 2016) based on Bray–Curtis dissimilarity. This method partitions the variance in beta diversity among the different environmental predictors. For the *Chronosequence* dataset, 24 soil variables were measured for the 50 samples considered here. We summarized these soil variables into three orthogonal ecological gradients, corresponding to the first three principal components of a principal components analysis. Principal component 1 (PC1; 44.6% of variance) corresponds to a gradient of soil extractable P and total N contents that increase from young (1- to 3-yr-old fields and 4- to 5-yr-old fields) to old fields (> 10 yr old; Supporting Information Fig. S1). Principal component 2 (PC2; 17.7%) corresponds to a gradient of inorganic N, which peaks in the 5- to 10-yr-old fields. Principal component 3 (PC3; 9.7%) differentiated fields from one another across time with a peak in magnesium in the 10-yr-old field. From now and on, PC1, PC2, and PC3 refer to the environmental predictors used for the *Chronosequence* dataset. For the *Soil depth* dataset, we used the soil layer (topsoil vs subsoil) and compartment (rhizosphere, drilosphere, bulk soil) of 18 samples from one agricultural field (Sosa-Hernández *et al.*, 2018a). The significance of the environmental predictors was assessed with 999 Monte Carlo permutations. Beta diversity patterns at different phylogenetic resolutions and for each beta diversity metric were visualized using a principal coordinates analysis (PCoA).

To ensure that our results were due to a genuine phylogenetic signal (i.e. departing from random phylogenetic relationships), we randomized the tip labels of the phylogeny while keeping constant the community composition as in Groussin *et al.* (2017) 100 times. Variance partitioning was conducted for each permutation step to generate a null distribution of the explained variance (PERMANOVA R^2) of each variable. R^2 values were considered not due to chance when departing from the 5–95% quantiles of the null distribution. A departure from the null distribution further indicates that an increase in explained variance at coarser phylogenetic resolutions is not solely due to a reduction in OTU number. We also took into account the uncertainty in phylogenetic tree topology in our analysis. Variance partitioning was repeated for each of the 100 bootstrapped RAXML trees, and predictors were deemed significant if the P values were < 0.01 for at least 90% of the trees.

Decomposition of MNTD and MPD phylogenetic beta diversity measures into phylogenetic distance classes To estimate the phylogenetic similarity between communities across the environmental gradients, we further decomposed the correlation of MNTD and MPD phylogenetic beta diversity with the environmental gradients into small (shallow phylogenetic turnover) and large (deep phylogenetic turnover) phylogenetic distance classes using Mantel correlogram analysis in ECODIST (Goslee & Urban, 2007). The sign of the correlation indicates whether communities within a phylogenetic distance class are found in similar (positive correlation) or dissimilar (negative correlation) environments compared with communities in the other distance classes. The significance was assessed with 999 Monte Carlo permutations of the phylogenetic distance matrix.

Phylogenetic signal of ESV ecological traits To complement the beta diversity analyses, we further tested whether closely related ESVs tend to share similar ecological traits (i.e. phylogenetic signal) and estimated the phylogenetic depth of this correlation. To this end, we conducted phylogenetic correlograms analyses using the PHYLOSIGNAL R package (Keck *et al.*, 2016) to test a phylogenetic signal in ESV co-occurrence (i.e. whether closely related ESVs tend to occur in the same samples) and in the optimum of each of the environmental predictors (as a measure of ecological traits). Two closely related ESVs may not co-occur yet be found in samples of similar environmental values. The co-occurrence between any pair of ESVs was measured as the Euclidean distance of ESV presence/absence in each sample. We inferred the ecological trait of each ESV by calculating their optimum for each of the environmental predictors (i.e. PC1, PC2, PC3, soil depth, and compartment). These optima corresponded to the abundance weighted mean of each of the environmental predictors for each ESV. Phylogenetic correlograms were calculated on a continuous basis by computing Moran's I index

(Moran, 1948) in the case of PCs, soil depth, and compartment class optima, or with the Mantel statistic in the case of co-occurrence. A matrix of phylogenetic weights was computed with default parameters to account for the nonuniform distribution of tips within the phylogeny (e.g. caused by different rates of diversification between lineages or phylogenetic sampling extent). Significance was tested using 100 nonparametric bootstrap resamplings of the phylogenetic tree to estimate the 95% confidence interval (CI) of the correlation. CI curves not overlapping with zero are significant.

Identification of clades that covary with the environmental predictors To further identify clades at different phylogenetic depths that covary in abundance with each of the environmental predictors, the correlation of node abundance (sum of the count of a node's respective ESVs) with each environmental predictor, for all nodes along the phylogenetic tree, was further tested with a linear model and corrected for multiple testing using the false discovery rate (Benjamini & Hochberg, 1995).

Sensitivity analysis To assess to what extent the approaches and OTU phylogenetic resolution mentioned earlier yielded similar results in terms of beta diversity, we conducted a sensitivity analysis by comparing beta diversity patterns with Mantel tests (Pearson's R correlation coefficient) using VEGAN (Oksanen *et al.*, 2016). It revealed that beta diversity patterns varied according to the phylogenetic resolution at which they were assessed and not the method used (Figs S2–S5). Therefore, in the following, we discuss the BDTT analysis using the LULU curated dataset.

Results

Number of OTUs across phylogenetic depths

In total, we obtained 574 and 233 Glomeromycotina ESVs in the *Chronosequence* and *Soil depth* datasets, respectively, of which 240 (42%) and 75 (33%), respectively, were kept after LULU curation, accounting for 84% and 77% of the reads. In both the *Chronosequence* and *Soil depth* datasets, we observed a drop in the number of OTUs from 0.01 to 0.03 subs/site (Figs 1–3a,c). There was a six-fold decrease in the number of OTU_{BDTT}s at a genetic distance of 0.03 subs/site (44 and 15 $OTU_{BDTT0.03}$s in the *Chronosequence* and *Soil depth* datasets; Figs 1–3a,c), indicating that ESV inference resolved a high diversity within a range of sequence divergence from 0 to 0.03 subs/site. The number of OTU_{BDTT}s continued to decrease more gently up to 0.1 subs/site (12 and 10 $OTU_{BDTT0.1}$s). As a matter of comparison, traditional OTU clustering at 97% with UPARSE ($OTU_{UPARSE0.03}$) yielded 52 (*Chronosequence* dataset) and 22 (*Soil depth* dataset) OTUs, which is in between $OTU_{BDTT0.02}$ and $OTU_{BDTT0.03}$.

Environmental predictors of AM fungal beta diversity across OTU phylogenetic resolutions

For the *Chronosequence* dataset, PC1 was the strongest predictor of AM fungal beta diversity and was significant at all OTU phylogenetic resolutions, but the variance explained varied strongly (Fig. 3b): PC1 explained 11% of the variance at the ESV level, but rose strongly to 22% at 0.03 subs/site and rose again to 30% at 0.1 subs/site. In addition, PC2 also explained 8% of the variance at the ESV level, rose to 12% at 0.03 subs/site, where it peaked, and then dropped. PC3 explained little but significant variance in community composition at the ESV level and was not significant at higher genetic distances. These results are in agreement with our PCoA analyses (Fig. S3): at 0.2 subs/site, samples clustered into two groups corresponding to fields recultivated for less (low PC1 value) and for more (high PC1 value) than 10 yr. At 0.03 subs/site, communities from the three first years of recultivation (low PC2 value) and the 4- and 5-yr old fields (high PC2 value) clustered apart from each other and formed two new clusters with low variability within each cluster. At the ESV level, fields from each of these clusters were further separated according to year of recultivation.

Decomposition of MNTD and MPD beta diversity into classes of low to high phylogenetic dissimilarity revealed estimates of community phylogenetic turnover across environmental gradients similar to the BDTT analysis. Communities within an MNTD of 0.03 subs/site and an MPD of 0.3–0.6 subs/site were positively correlated with PC1 (i.e. these communities tend to occur in samples of similar PC1 values), whereas communities with higher phylogenetic dissimilarity (MNTD > 0.03 subs/site and MPD > 0.03–06 subs/site) were negatively correlated with PC1 (i.e. tend to occur in samples of dissimilar PC1 values; Fig. 4a,b). The correlation with PC2 was difficult to decompose because the relationship with MNTD and MPD was nonmonotonic, indicating low phylogenetic turnover for high PC2 distances (Fig. S8); yet, the positive correlation of MPD to PC2 was observed for lower genetic distance classes than PC1; and at high distance classes, MPD negatively correlated to PC2 while it was still positively correlated to PC1 (Fig. 4b).

In the *Soil depth* dataset, soil layer was the unique predictor of beta diversity and was significant at all genetic distances (Figs 3, S5), but, again, the explained variance varied strongly. Soil layer explained 28% of the variance at the ESV level and rose to 45% to reach a plateau at 0.03 subs/site. It then increased slightly from 0.15 to 0.16 subs/site. Soil compartment effects were never significant. MNTD was strongly correlated with soil layer within the first genetic distance class of 0.03 subs/site (Fig. 4c). MPD was positively and negatively correlated at intermediate and high distance classes, respectively (Fig. 4d). Neither MNTD nor MPD correlated with soil compartment.

In summary, our results show that clustering at *c.* 0.03 subs/site has lumped ESVs into phylogenetically and ecologically coherent units that strongly correlate with the environment. They also show a turnover of distantly related AM fungal clades (> 0.03 subs/site) across PC1 and soil layer, and a turnover of relatively closely related AM fungal clades (*c.* 0.03 subs/site) across PC2. PERMANOVA P values did not change across the 100 bootstrapped trees, supporting that these results were robust to phylogenetic uncertainty. PERMANOVA R^2 were not comprised within the 5–95% distribution of null distribution of R^2 from 100 phylogenetic randomizations, further indicating that these

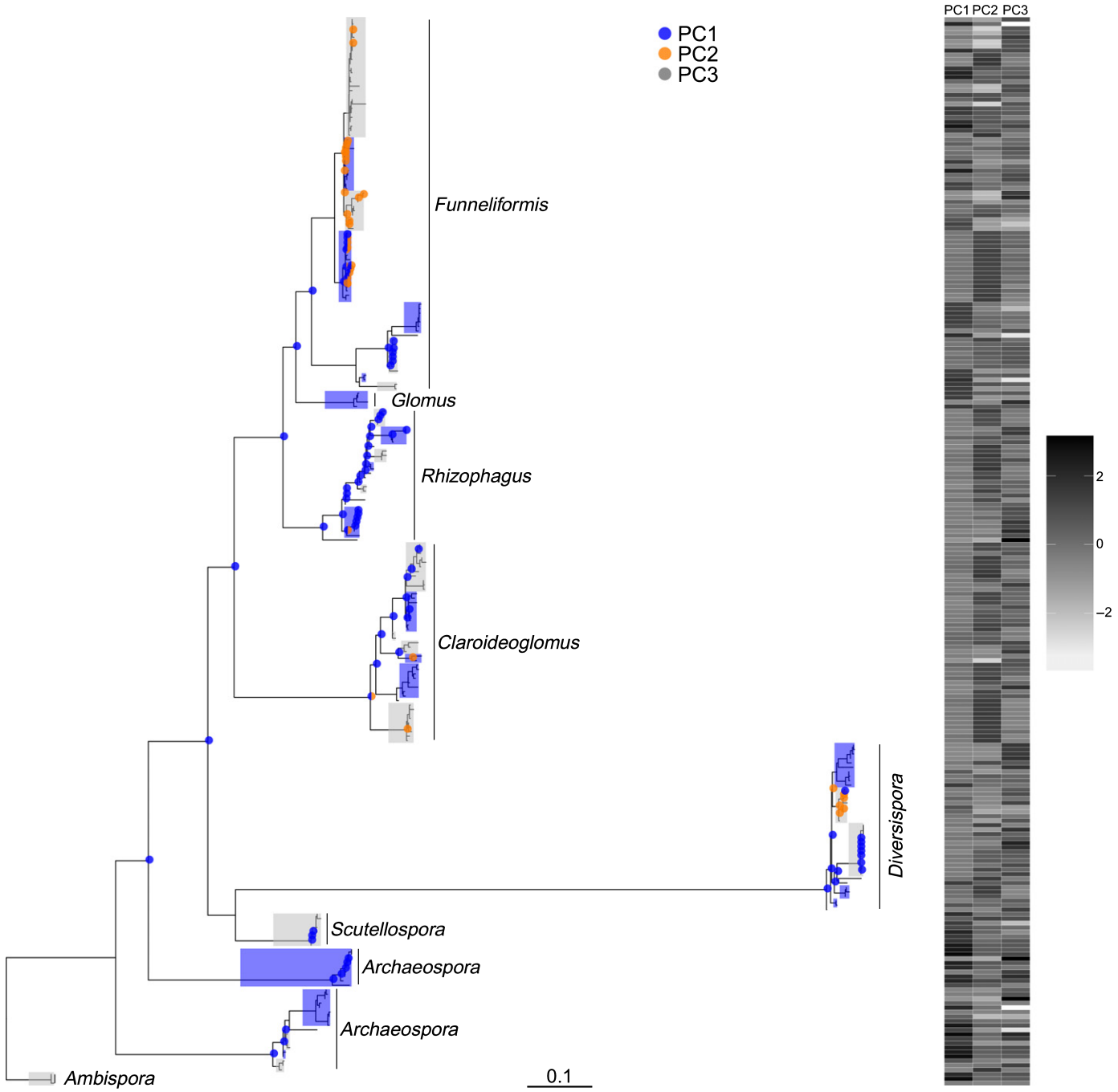

Fig. 1 Maximum-likelihood phylogenetic tree of arbuscular mycorrhizal fungal exact sequence variants (ESVs) for the *Chronosequence* dataset. The phylogenetic clustering of ESVs at a genetic distance of 0.03 mean substitutions per site is highlighted in blue and gray rectangles. Significant correlation between the abundance of clades (nodes) across the phylogeny and principal component (PC) 1 (blue), PC2 (orange), or PC3 (gray) is depicted at each node. Bar, 0.1 mean substitutions per site. The heatmap on the right of the tree shows the centered and scaled PC1, PC2 and PC3 optima.

results derived from a genuine phylogenetic signal and that none of the changes in the correlation strength across phylogenetic resolutions could be solely attributed to change in OTU number.

Phylogenetic signal of ESV ecological traits

In the *Chronosequence* dataset, we observed a positive autocorrelation (Mantel's $R = 0.05$) in ESV co-occurrence up to deep phylogenetic distances classes (Fig. 5a). However, PC1 optimum had the highest positive autocorrelation (Moran's $I = 0.3$) and the signal was significantly positive up to 0.4 subs/site patristic distance and significantly negative at largest genetic distances (Fig. 5b). PC2 optimum showed positive autocorrelation (Moran's $I = 0.09$) up to 0.2 subs/site patristic distance (Fig. 5c). We observed similar results in the nonLULU curated dataset, but the phylogenetic signal for PC2 was deeper (Fig. S9). In the *Soil depth* dataset, a positive phylogenetic signal was significant only for co-occurrences (Fig. 5d), but in the noncurated LULU dataset

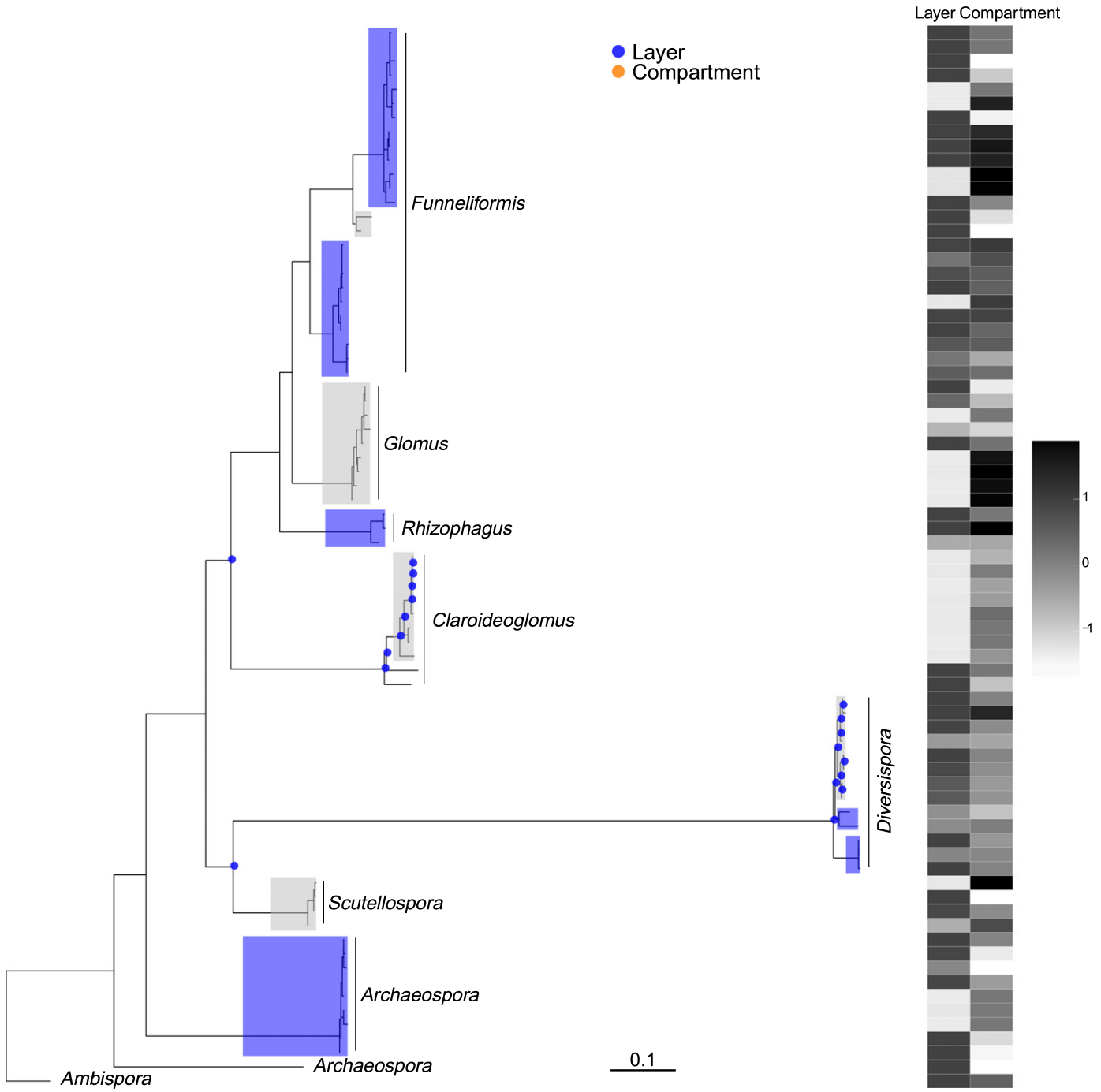

Fig. 2 Maximum-likelihood phylogenetic tree of arbuscular mycorrhizal fungal exact sequence variants (ESVs) for the *Soil depth* dataset. The phylogenetic clustering of ESVs at a genetic distance of 0.03 mean substitutions per site is highlighted in blue and gray rectangles. Significant correlation between the abundance of clades (nodes) across the phylogeny and soil layer (blue) or soil compartment (orange) is depicted at each node. Bar, 0.1 mean substitutions per site. The heatmap on the right of the tree shows the centered and scaled soil layer and soil compartment optima.

the soil layer optimum, and soil compartment to a lesser extent, was positively autocorrelated for ESVs as distant as 0.6 subs/site patristic distance (Moran's $I = 0.2$) and negatively autocorrelated after 1 sub/site patristic distance (Fig. S9).

These results are consistent with the beta diversity analyses. They indicate that closely related ESVs tend to be found in the same samples; and if not, then in samples with similar soil conditions. PC1 and soil layer optimum had the strongest phylogenetic signal, whereas PC2 optimum was conserved at a shallow phylogenetic depth.

Identity of clades that covary with the environmental predictors

In the *Chronosequence* dataset, multiple nodes at all phylogenetic depths showed significant correlation with PC1: notably,

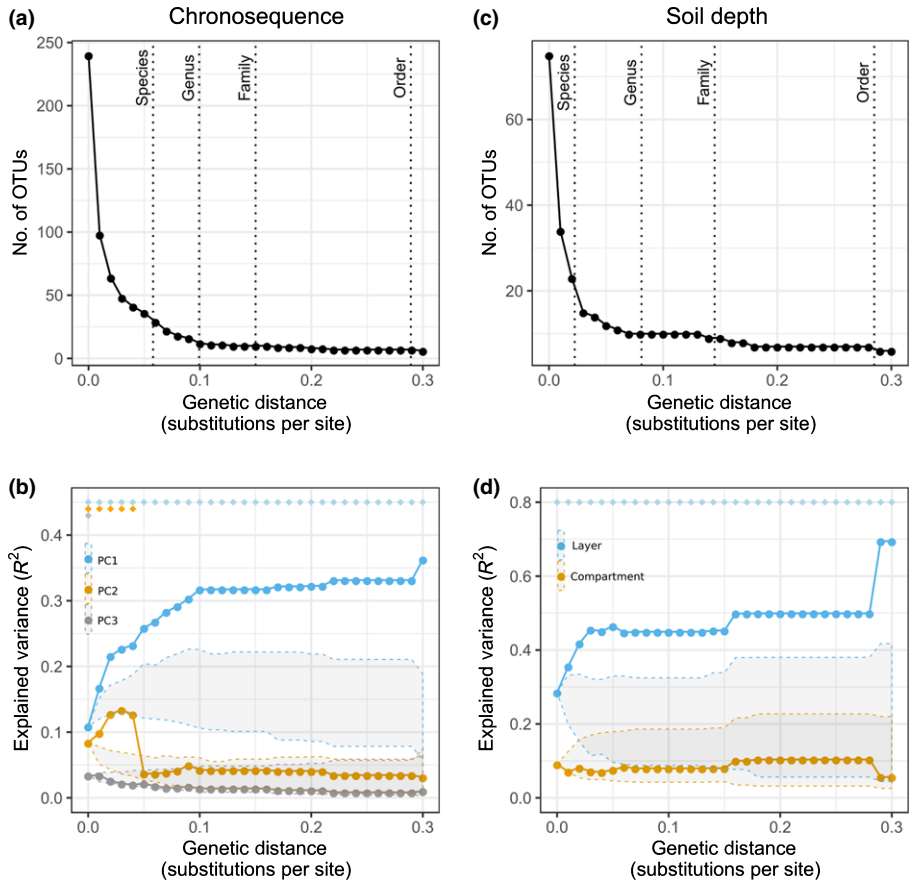

Fig. 3 Predictors of operational taxonomic units (OTUs) beta-diversity in arbuscular mycorrhizal (AM) fungal communities across phylogenetic resolutions. Number of OTUs at different phylogenetic resolutions for the (a) *Chronosequence* and (c) *Soil depth* datasets. ESVs were agglomerated from a genetic distance (*x*-axis) of 0 (the ESV resolution, i.e. the tips of the phylogenetic tree) to a distance of 0.3, by steps of 0.01 nucleotide substitution per site. The dotted vertical lines indicate the mean genetic distance among ESVs within taxa, at each taxonomic rank. (b, d) Variance partitioning of AM fungal communities at different OTU phylogenetic resolutions. The *y*-axis is the part of explained variance (R^2) in community dissimilarities (Bray–Curtis) by principal component (PC) 1 (blue), PC2 (orange), and PC3 (gray) for the *Chronosequence* dataset, and by soil layer (blue) and soil compartment (orange) for the *Soil depth* dataset. Shaded areas correspond to the 5–95% quantiles of the R^2 calculated with 100 randomizations of the phylogeny. Diamond at the top of the figure indicates significant R^2 values ($P < 0.01$) for > 90% of the 100 bootstrapped trees. The figure presents results after LULU correction (see Materials and Methods section). Results spanning the entire tree and without LULU correction are presented in Supporting Information Figs S6 and S7, respectively.

members of *Rhizophagus*, *Funneliformis*, and *Claroideoglomus*, including the Glomerales node and to a lesser extent *Diversispora*, are mostly found in young fields (low PC1), whereas *Archaeospora* members are found in old fields (high PC1; Figs 1a, S10). By contrast, only nodes mostly at or contained within 0.03 subs/site genetic distance correlated with PC2, including members of *Funneliformis*, *Diversispora*, and *Claroideoglomus*, found at either low or high PC2 values. In the *Soil depth* dataset, the *Claroideoglomus* node and its members were found in subsoil, whereas *Diversispora* and the node *Diversispora*/*Scutellospora* (Diversisporales) were found in topsoil (Figs 1b, S11). The significant increase in (relative) abundance of the Glomerales node with soil depth is probably driven by a strong increase in the abundance of Claroideoglomeraceae and does not correspond with an increase in abundance of other members of this node (i.e. *Funneliformis*, *Glomus*, and *Rhizophagus*).

Discussion

In this study, we tested whether and how beta diversity patterns change with OTU phylogenetic resolution from the ESV to order-level sequence clusters. At the ESV level, we were able to detect changes in community composition similar to traditional OTU clustering methods or at coarser phylogenetic resolutions. However, we showed a stronger correlation between the environmental predictors and beta diversity assessed at coarser phylogenetic resolutions, especially at *c.* 0.03 subs/site. Furthermore, we observed a turnover in either distantly or closely related clades along different environmental gradients, such that the variance explained by different environmental predictors varied with OTU phylogenetic resolution. This suggests that accounting for the evolutionary history of AM fungi could help better understanding of the drivers of their spatial distribution.

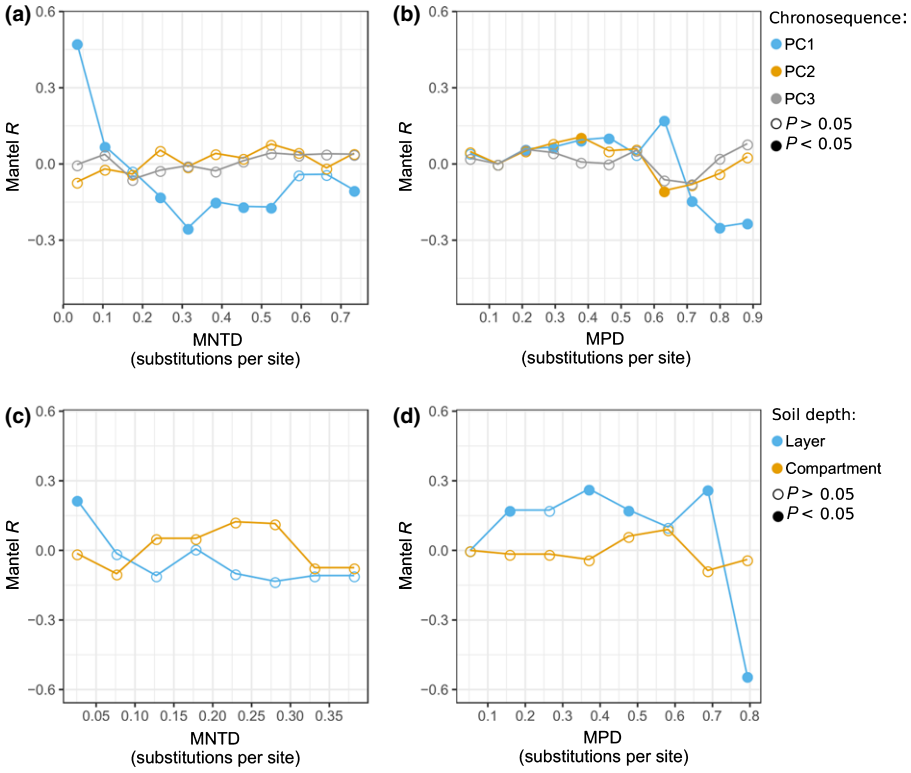

Fig. 4 Predictors of phylogenetic beta diversity in arbuscular mycorrhizal fungal communities across phylogenetic resolutions. Pearson's *R* Mantel correlograms between (a) the weighted mean nearest taxon distance (MNTD) and (b) the weighted mean pairwise distance (MPD) phylogenetic beta diversity measures and principal component (PC) 1 (blue), PC2 (orange), and PC3 (gray), for the *Chronosequence* dataset and between (c) weighted MNTD and (d) weighted MPD, and soil layer (blue) and soil compartment (orange) for the *Soil depth* dataset. Significant correlation coefficients ($P < 0.05$) are indicated by filled symbols.

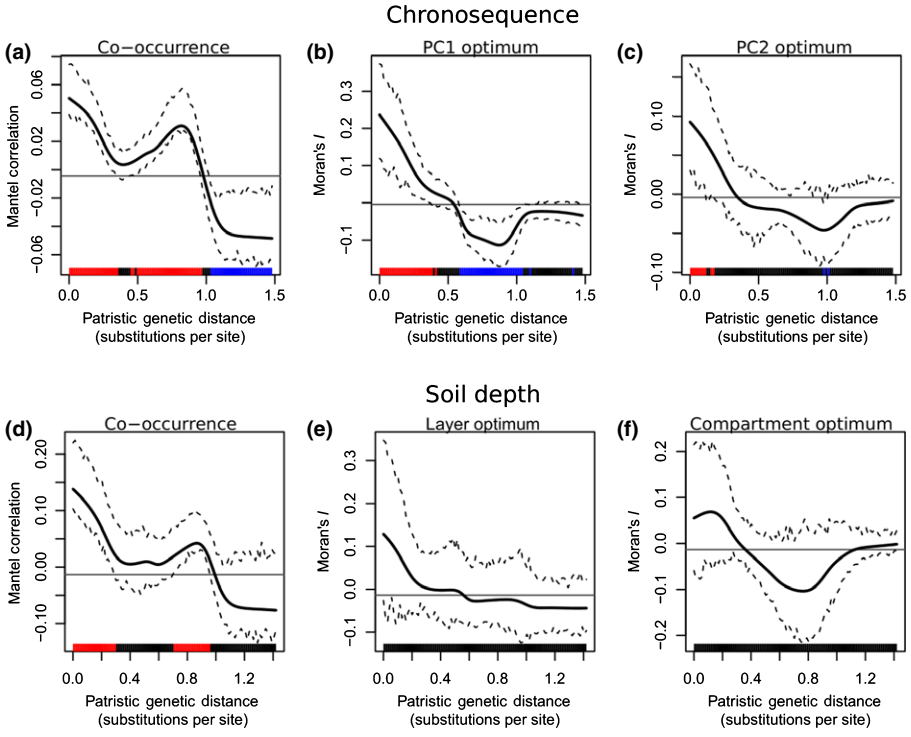

Fig. 5 Phylogenetic signal in the ecological traits inferred for exact sequence variants (ESVs). Phylogenetic correlograms between pairwise phylogenetic (patristic) distances of ESVs and (a) their co-occurrence, (b) their principal component (PC) 1 optimum, and (c) their PC2 optimum, across genetic distances, for the *Chronosequence* dataset, and (d) their co-occurrence, (e) their soil layer, and (f) their soil compartment for the *Soil depth* dataset. The solid bold black line represents (a, d) the Mantel correlation and (b, c, e, f) Moran's *I* index of autocorrelation. The dashed black lines represent the lower and upper bounds of the 95% confidence interval. The horizontal black line indicates the expected value of Mantel Pearson's *R* or Moran's *I* under the null hypothesis of no phylogenetic autocorrelation. Colored bar along the *x*-axis indicates no correlation (black) or significant ($P < 0.05$) positive (red) or negative (blue) correlation, respectively. See Supporting Information Fig. S9 for the results without LULU curation.

At the ESV level, we statistically detected the environmental predictors revealed with traditional OTUs or at higher genetic distances, consistent with previous results for bacteria (Thompson *et al.*, 2017) and fungi using the ITS marker (Glassman & Martiny, 2018). This was because closely related ESVs tend to occur in the same samples. However, we showed a strong increase in the correlation between most environmental predictors and beta diversity at a genetic distance of *c.* 0.03 subs/site, and even

higher for certain environmental predictors. This shows that clustering ESVs into OTUs at 0.02–0.04 subs/site (and 0.03–0.05 subs/site for the nonLULU-curated dataset), or higher, reduced the noise in the biological signal by lumping closely related ESVs that occur in otherwise similar environments but in different communities. These results support the idea that fungal community assembly, and in particular AM fungal communities, are governed by environmental selection of organisms exhibiting strong phylogenetic niche conservatism (Lekberg *et al.*, 2014; Powell & Sikes, 2014; Botnen *et al.*, 2018). In this case, given that AM fungi exhibit a significant degree of phylogenetic niche conservatism, clustering ESVs into OTUs at *c.* 0.03 subs/site or higher may serve a better purpose for niche modeling by increasing the power to detect correlations with environmental predictors (Powell & Sikes, 2014).

Interestingly, for both the *Chronosequence* and *Soil depth* datasets, the beta diversity approaches based on a phylogenetic tree (BDTT, MNTD, and MPD) converged towards *c.* 0.03 subs/site as a level of evolutionary divergence carrying a strong ecological specialization signal. This threshold to delineate AM fungal OTUs with LSU will certainly be lower for SSU and higher for ITS (Thiéry *et al.*, 2016); this may change when studying other organisms and, as shown here, with other environmental gradients. Interestingly, the number of OTUs plateaued after 0.03 subs/site (both after and before LULU curation). These characteristics indicate that OTUs at 0.03 subs/site can be defined as ecologically coherent biological units, which could be species (Bruns *et al.*, 2018). More broadly, biological units that are simultaneously evolutionarily and ecologically coherent can be functional groups amenable to quantitative PCR monitoring.

Our results suggest an additional ecological structure in beta diversity at the ESV level: in the *Chronosequence* dataset, we observed a weak but significant correlation between beta diversity and PC3, and a split among each fields recultivated from 1 to 5 yr, which was not the case at coarser phylogenetic resolutions, including when considering OTU_{UPARSE}. This result is supported in our sensitivity analysis by the slight clustering of beta diversity patterns at the ESV level apart from the other beta diversity patterns inferred at relatively fine phylogenetic resolutions. What biological unit an ESV captures remains unknown, however, and it would ultimately depend on the rate of evolution of the targeted gene and clades. The observed diversity of LSU variants within a range of sequence divergence from 0 to 0.03 subs/site is consistent with estimates of intraspecific diversity obtained from single isolates and spores within *Rhizophagus irregularis* and *Gigaspora margarita* (Thiéry *et al.*, 2016), suggesting that successional population processes were captured. Through ESV inference, we could separate the process of sequence denoising from the biological unit delineation, with the potential to infer the ecological variation at intra- vs interspecific levels.

Our results show that the degree of phylogenetic turnover among AM fungal communities differs along different environmental gradients. We observed a turnover of distantly related clades across PC1 and soil depth (clades diverging at least by 0.2 subs/site), whereas clades at or within a genetic distance of *c.* 0.03 subs/site varied along PC2. These results suggest that ecological specializations to different aspects of the environment (e.g. soil P and N content, soil depth) are conserved at different phylogenetic depths in AM fungi. This could reflect the primary factors imposing pressure on the AM symbiosis. For instance, fungi reported to have a ruderal strategy, such as *Funneliformis*, *Rhizophagus*, and *Claroideoglomus* within the Glomeraceae (Hart & Reader, 2002; Chagnon *et al.*, 2013), were abundant in young fields, whereas *Archaeospora* only occurred in soils recultivated for > 10 yr. PC1 represented a strong temporal gradient, with steep variations in plant-available P content. The occurrence and abundance of AM fungi across this gradient are probably based on their ability to disperse, colonize, and persist in soil and roots, and maintain an association with plants in soils of different P content. The strong mutual exclusion of *Diversispora*/*Scutellospora* found in topsoil and of *Claroideoglomus* found in subsoil (*Soil depth* dataset) and in young fields where the subsoil : topsoil mix has been recently deposited (*Chronosequence* dataset) indicates that some *Claroideoglomus* fungi cannot persist in topsoil and/or have specialized to subsoil (Sosa-Hernández *et al.*, 2018b), a habitat with low plant-carbon supply probably selecting for stress-tolerant fungi. Our results are consistent with colonization–persistence or competitors–stress tolerators–ruderals life-history strategies probably being conserved at the family level (Hart & Reader, 2002; Chagnon *et al.*, 2013). They are also consistent with P uptake being the primary function of AM fungi (Smith & Read, 2008). PC2 represented a strong increase in inorganic N availability but also in N : P in fields recently converted to agriculture. The turnover of closely related AM fungi along PC2 contradicts the idea that AM fungal association with N availability (plant-available N in our case) is conserved at family or genus level (Treseder *et al.*, 2018). This discrepancy could be due to the difference in the study ecosystems (natural or agricultural ecosystems), to fungal physiological requirement (e.g. stoichiometry) but also to the difficulty in disentangling confounding abiotic and/or biotic factors. Along with investigations on isolates, further research conducted at multiple phylogenetic resolutions will certainly help test whether soil P, more than soil N content, could have driven the early diversification of AM fungi.

Inferring life-history strategies and functions in fungi remains difficult (Powell & Rillig, 2018; Treseder *et al.*, 2018). We show that one could infer such traits from distribution data based on DNA sequencing when coupled with contextual parameters, which is a promising result given that DNA-based assessments of biodiversity are now accumulating. Increasing spatial extent and decreasing spatial grain will increase environmental heterogeneity and capture environmental gradients of varying steepness and with additional importance for the AM symbiosis (e.g. vegetation characteristics, climate) in addition to an increased breadth of AM phylogenetic diversity. It will further reveal the ecological variation at different levels of phylogenetic divergence in AM fungi and narrow the estimated depth of phylogenetic conservatism of these ecological traits. For example, traits variation at very fine phylogenetic resolution could be relevant for subtle environmental variation (Powell & Sikes, 2014) and might be resolved with ESV inference. Such variation could exist, for

instance, between extra- or intraradical mycelium (if different portions of the mycelium express functionally different ribosomes; Gilbert, 2011; Filipovska & Rackham, 2013), when considering the selection of different nucleotypes at the arbuscule scale (Limpens & Geurts, 2014), or when other deterministic processes occur (Verbruggen and Kiers 2010). At the other extreme of phylogenetic depth, differences between biomes or ecoregions could be clear at broad phylogenetic resolutions (Stürmer *et al.*, 2018b). Alternatively, a biogeographic structure might be revealed at the ESV level in the case of dispersal limitation of cryptic taxa and stochastic population-level processes (e.g. genetic drift; Bruns & Taylor, 2015; Savary *et al.*, 2018).

A phylogenetic decomposition of beta diversity may prove difficult using ITS, a genetic marker used by most molecular fungal ecologists (Schoch *et al.*, 2012), due to limitations in phylogeny reconstruction. Clustering ESVs based on (morphology-based) taxonomy (Cline *et al.*, 2017) or using traditional clustering methods (Lekberg *et al.*, 2018) can alternatively be used to reveal significant changes of community composition across environmental gradients. Our sensitivity analysis showed both methods can be used to assess change in community composition at different phylogenetic resolutions. Finally, phylogenetic placement methods, together with the recent release of improved phylogenies for the entire fungal kingdom (e.g. Tedersoo *et al.*, 2018) will certainly anchor fungal molecular ecology using ITS into a phylogenetic framework.

Conclusion

The establishment of an organism in a community is determined by a suite of processes, including ecological drift, dispersal, and abiotic and biotic selection. The interaction of these processes imprints on the phylogenetic structure of communities at multiple phylogenetic resolutions. We used this information to identify at which phylogenetic resolution AM fungal communities covary with different environmental factors. We showed that beta diversity patterns of AM fungi change with OTU phylogenetic resolution, so that one phylogenetic resolution gives only partial information on the factors shaping their spatial distribution. We suggest that assessing beta diversity patterns at varying OTU phylogenetic resolutions offers a deeper understanding of the ecology and evolution of fungi. Identifying ecological variation among taxa of AM fungi at different phylogenetic depths will ultimately help define functional groups and serve as a guide for the engineering of AM fungal communities (Rillig *et al.*, 2016).

Acknowledgements

We acknowledge funding from the Federal Ministry of Education and Research (BMBF) for the initiative 'BonaRes – Soil as a sustainable resource for the bioeconomy' to the projects INPLAMINT and Soil[3]. FM was supported by a Banting postdoctoral fellowship. We thank François Keck for advice in using PHYLOSIGNAL. We thank Maarja Öpik and four anonymous reviewers for providing comments that improved the manuscript.

Author contributions

JR elaborated the study. JR analyzed the data with the help of FM for using BDTT. JR wrote the manuscript. LZ, FM, MASH, MCR, SH and JFD contributed to ideas and helped revising the manuscript. All authors approved the final version of the manuscript.

ORCID

Florent Mazel https://orcid.org/0000-0003-0572-9901
Matthias C. Rillig https://orcid.org/0000-0003-3541-7853
Julien Roy https://orcid.org/0000-0003-2964-1314
Moisés A. Sosa-Hernández https://orcid.org/0000-0002-1943-0102
Lucie Zinger https://orcid.org/0000-0002-3400-5825

References

Anderson MJ. 2001. A new method for non parametric multivariate analysis of variance. *Austral Ecology* **26**: 32–46.

Angelard C, Tanner CJ, Fontanillas P, Niculita-Hirzel H, Masclaux F, Sanders IR. 2014. Rapid genotypic change and plasticity in arbuscular mycorrhizal fungi is caused by a host shift and enhanced by segregation. *ISME Journal* **8**: 284–294.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B* **57**: 289–300.

Botnen SS, Davey ML, Halvorsen R, Kauserud H. 2018. Sequence clustering threshold has little effect on the recovery of microbial community structure. *Molecular Ecology Resources* **18**: 1064–1076.

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

Bruns TD, Corradi N, Redecker D, Taylor JW, Öpik M. 2018. Glomeromycotina: what is a species and why should we care ? *New Phytologist* **220**: 963–967.

Bruns TD, Taylor JW. 2015. Comment on 'Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism'. *Science* **351**: 826.

Callahan BJ, McMurdie PJ, Holmes SP. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal* **11**: 2639–2643.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**: 581.

Camenzind T, Hempel S, Homeier J, Horn S, Velescu A, Wilcke W, Rillig MC. 2014. Nitrogen and phosphorus additions impact arbuscular mycorrhizal abundance and molecular diversity in a tropical montane forest. *Global Change Biology* **20**: 3646–3659.

Chagnon PL, Bradley RL, Maherali H, Klironomos JN. 2013. A trait-based framework to understand life history of mycorrhizal fungi. *Trends in Plant Science* **18**: 484–491.

Cline LC, Song Z, Al-Ghalith GA, Knights D, Kennedy PG. 2017. Moving beyond *de novo* clustering in fungal community ecology. *New Phytologist* **216**: 629–634.

Coissac E, Riaz T, Puillandre N. 2012. Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular Ecology* **21**: 1834–1847.

Davison J, Moora M, Öpik M, Adholeya A, Ainsaar L, Bâ A, Burla S, Diedhiou AG, Hiiesalu I, Jairus T *et al.* 2015. Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science* **127**: 970–973.

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

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

Edgar RC. 2016. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv*: 81257.

Eren AM, Morrison HG, Lescault PJ, Reveillaud J, Vineis JH, Sogin ML. 2015. Minimum entropy decomposition: unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *ISME Journal* **9**: 968–979.

Filipovska A, Rackham O. 2013. Specialization from synthesis: how ribosome diversity can customize protein function. *FEBS Letters* **587**: 1189–1197.

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ. 2017. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications* **8**: e1188.

Gilbert WV. 2011. Functional specialization of ribosomes? *Trends in Biochemical Sciences* **36**: 127–132.

Glassman SI, Martiny BH. 2018. Broadscale ecological patterns are robust to use of exact sequence variants versus operational taxonomic units. *mSphere* **3**: e00148-18.

Goslee SC, Urban DL. 2007. The ECODIST package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software* **22**: 19. 10.18637/jss.v022.i07

Groussin M, Mazel F, Sanders JG, Smillie CS, Lavergne S, Thuiller W, Alm EJ. 2017. Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. *Nature Communications* **8**: e14319.

Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH. 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology* **10**: 497–506.

Hart MM, Aleklett K, Chagnon P, Egan C, Ghignone S, Helgason T, Lekberg Y, Öpik M, Pickles BJ, Waller L. 2015. Navigating the labyrinth: a guide to sequence-based, community ecology of arbuscular mycorrhizal fungi. *New Phytologist* **207**: 235–247.

Hart MM, Reader RJ. 2002. Taxonomic basis for variation in the colonization strategy of arbuscular mycorrhizal fungi. *New Phytologist* **153**: 335–344.

Hazard C, Johnson D. 2018. Does genotypic and species diversity of mycorrhizal plants and fungi affect ecosystem. *New Phytologist* **220**: 1122–1128.

Horn S, Caruso T, Verbruggen E, Rillig MC, Hempel S. 2014. Arbuscular mycorrhizal fungal communities are phylogenetically clustered at small scales. *ISME Journal* **8**: 2231–2242.

Johnson D, Martin F, Cairney JWG, Anderson IC. 2012. The importance of individuals: intraspecific diversity of mycorrhizal plants and fungi in ecosystems. *New Phytologist* **194**: 614–628.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**: 3059–3066.

Keck F, Rimet F, Bouchez A, Franc A. 2016. PHYLOSIGNAL: an R package to measure, test, and explore the phylogenetic signal. *Ecology and Evolution* **6**: 2774–2780.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. PICANTE: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463–1464.

Koch AM, Croll D, Sanders IR. 2006. Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. *Ecology Letters* **9**: 103–110.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq. *Applied and Environmental Microbiology* **79**: 5112–5120.

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 Phytologist* **193**: 970–984.

Lekberg Y, Gibbons SM, Rosendahl S. 2014. Will different OTU delineation methods change interpretation of arbuscular mycorrhizal fungal community patterns? *New Phytologist* **202**: 1101–1104.

Lekberg Y, Vasar M, Bullington LS, Sepp S-K, Antunes PM, Bunn R, Larkin BG, Öpik M. 2018. More bang for the buck? Can arbuscular mycorrhizal fungal communities be characterized adequately alongside other fungi using general fungal primers? *New Phytologist* **220**: 971–976.

Limpens E, Geurts R. 2014. Plant-driven genome selection of arbuscular mycorrhizal fungi. *Molecular Plant Pathology* **15**: 531–534.

Lindahl BD, Nilsson RH, Tedersoo L, Abarenkov K, Carlsen T, Kjøller R, Kõljalg U, Pennanen T, Rosendahl S, Stenlid J *et al.* 2013. Fungal community analysis by high-throughput sequencing of amplified markers – a user's guide. *New Phytologist* **199**: 288–299.

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

Martiny JBH, Jones SE, Lennon JT, Martiny AC. 2015. Microbiomes in light of traits: a phylogenetic perspective. *Science* **350**: aac9323.

Mensah JA, Koch AM, Antunes PM, Kiers ET, Hart M, Bücking H. 2015. High functional diversity within species of arbuscular mycorrhizal fungi is associated with differences in phosphate and nitrogen uptake and fungal phosphate metabolism. *Mycorrhiza* **25**: 533–546.

Moran P. 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society* **10**: 243–251.

Munkvold L, Kjøller R, Vestberg M, Rosendahl S, Jakobsen I. 2004. High functional diversity within species of arbuscular mycorrhizal fungi. *New Phytologist* **164**: 357–364.

Oksanen JF, Blanchet G, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P *et al.* 2016. *VEGAN: community ecology package*. [WWW document] URL https://cran.r-project.org/package=vegan

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

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

Powell JR, Monaghan MT, Öpik M, Rillig MC. 2011. Evolutionary criteria outperform operational approaches in producing ecologically relevant fungal species inventories. *Molecular Ecology* **20**: 655–666.

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. *Proceedings of the Royal Society of Biological Sciences* **276**: 4237–4245.

Powell JR, Rillig MC. 2018. Biodiversity of arbuscular mycorrhizal fungi and ecosystem function. *New Phytologist* **220**: 1059–1075.

Powell JR, Sikes BA. 2014. Method or madness: does OTU delineation bias our perceptions of fungal ecology? *New Phytologist* **202**: 1095–1097.

R Core Team. 2017. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. [WWW document] URL https://www.R-project.org/.

Rillig MC, Sosa-Hernández MA, Roy J, Aguilar-Trigueros CA, Valyi K, Lehmann A. 2016. Towards an integrated mycorrhizal technology: harnessing mycorrhiza for sustainable intensification in agriculture. *Frontiers in Plant Science* **7**: 1625.

Rodríguez-Echeverría S, Teixeira H, Correia M, Timóteo S, Heleno R, Öpik M, Moora M. 2017. Arbuscular mycorrhizal fungi communities from tropical Africa reveal strong ecological structure. *New Phytologist* **213**: 380–390.

Roy J, Reichel R, Brüggemann N, Hempel S, Rillig MC. 2017. Succession of arbuscular mycorrhizal fungi along a 52-year agricultural recultivation chronosequence. *FEMS Microbiology Ecology* **93**: fix102–13.

Sanders IR, Rodriguez A. 2016. Aligning molecular studies of mycorrhizal fungal diversity with ecologically important levels of diversity in ecosystems. *ISME Journal* **10**: 2780–2786.

Savary R, Masclaux FG, Wyss T, Droh G, Cruz Corella J, Machado AP, Morton JB, Sanders IR. 2018. A population genomics approach shows widespread geographical distribution of cryptic genomic forms of the symbiotic fungus *Rhizophagus irregularis*. *ISME Journal* **12**: 17–30.

Schloss PD. 2013. Secondary structure improves OTU assignments of 16S rRNA gene sequences. *ISME Journal* **7**: 457–460.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ *et al.* 2009. Introducing MOTHUR: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**: 7537–7541.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal

DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences, USA* **109**: 6241–6246.

Selosse MA, Vincenot L, Öpik M. 2016. Data processing can mask biology: towards better reporting of fungal barcoding data? *New Phytologist* **210**: 1159–1164.

Smith SE, Read D. 2008. *Mycorrhizal symbiosis, 3rd edn.* San Diego, CA: Academic Press.

Sosa-Hernández MA, Roy J, Hempel S, Kautz T, Köpke U, Uksa M, Schloter M, Caruso T, Rillig MC. 2018a. Subsoil arbuscular mycorrhizal fungal communities in arable soil differ from those in topsoil. *Soil Biology and Biochemistry* **117**: 83–86.

Sosa-Hernández MA, Roy J, Hempel S, Rillig MC. 2018b. Evidence for subsoil specialization in arbuscular mycorrhizal fungi. *Frontiers in Ecology and Evolution* **6**: e67.

Spatafora JW, Chang Y, Benny GL, Lazarus K, Smith ME, Berbee ML, Bonito G, Corradi N, Grigoriev I, Gryganskyi A, *et al.* 2016. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* **108**: 1028–1046.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Stürmer SL, Oliveira LZ, Morton JB. 2018a. Gigasporaceae versus Glomeraceae (phylum Glomeromycota): a biogeographic tale of dominance in maritime sand dunes. *Fungal Ecology* **32**: 49–56.

Stürmer SL, Bever JD, Morton JB. 2018b. Biogeography of arbuscular mycorrhizal fungi (Glomeromycota): a phylogenetic perspective on species distribution patterns. *Mycorrhiza* **28**: 587–603.

Tedersoo L, Santiago S-R, Kõljalg U, Bahram M, Döring M, Schigel D, May T, Ryberg M, Abarenkov K. 2018. High-level classification of the Fungi and a tool for evolutionary ecological analyses. *Fungal Diversity* **90**: 135–159.

Thiéry O, Vasar M, Jairus T, Davison J, Roux C, Kivistik PA, Metspalu A, Milani L, Saks Ü, Moora M *et al.* 2016. Sequence variation in nuclear ribosomal small subunit, internal transcribed spacer and large subunit regions of *Rhizophagus irregularis* and *Gigaspora margarita* is high and isolate-dependent. *Molecular Ecology* **25**: 2816–2832.

Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G *et al.* 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**: 457–463.

Treseder KK, Hart MM, Egerton-warburton EBALM, Klironomos JN, Maherali H, Tedersoo L. 2018. Arbuscular mycorrhizal fungi as mediators of ecosystem responses to nitrogen deposition: a trait-based predictive framework. *Journal of Ecology* **106**: 480–489.

Vellend M. 2016. *The theory of ecological communities (MPB-57).* Princeton, NJ, USA: Princeton University Press.

Verbruggen E, Kiers ET. 2010. Evolutionary ecology of mycorrhizal functional diversity in agricultural systems. *Evolutionary Applications* **3**: 547–560.

Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. 2017. GGTREE: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**: 28–36.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Principal Component Analysis of 24 soil variables and time since recultivation of the *Chronosequence* dataset.

Fig. S2 Heatmap of pairwise Mantel correlation (Pearson's *R*) between beta diversity patterns assessed with different methods and at different OTU phylogenetic resolutions, for the *Chronosequence* data set.

Fig. S3 Principal Coordinates Analysis (PCoA) of community dissimilarities calculated at different OTU phylogenetic resolutions and with different methods, for the *Chronosequence* dataset.

Fig. S4 Heatmap of pairwise Mantel correlation (Pearson's *R*) between beta diversity patterns assessed with different methods and at different OTU phylogenetic resolutions, for the *Soil depth* dataset.

Fig. S5 Principal Coordinates Analysis (PCoA) of community dissimilarities calculated at different OTU phylogenetic resolutions and with different methods, for the *Soil depth* dataset.

Fig. S6 Predictors of beta-diversity in arbuscular mycorrhizal (AM) fungal communities across OTU phylogenetic resolutions after LULU curation spanning the entire tree.

Fig. S7 Predictors of beta-diversity in arbuscular mycorrhizal (AM) fungal communities across OTU phylogenetic resolutions before LULU curation spanning the entire tree.

Fig. S8 Relationship between phylogenetic beta-diversity (MNTD and MPD) and PC1 and PC2 Euclidean distance.

Fig. S9 Phylogenetic signal in the ecological traits inferred for exact sequence variants (ESVs) before LULU correction.

Fig. S10 Relative abundance of exact sequence variants (ESVs) across samples in the *Chronosequence* dataset.

Fig. S11 Relative abundance of exact sequence variants (ESVs) across samples in the *Soil depth* dataset.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.