Mycorrhiza (2022) 32:135–144
https://doi.org/10.1007/s00572-022-01072-7

ORIGINAL ARTICLE

Global taxonomic and phylogenetic assembly of AM fungi

Martti Vasar[1] · John Davison[1] · Siim-Kaarel Sepp[1] · Jane Oja[1] · Saleh Al-Quraishy[2] · C. Guillermo Bueno[1] · Juan José Cantero[3,4] · Ezequiel Chimbioputo Fabiano[5] · Guillaume Decocq[6] · Lauchlan Fraser[7] · Inga Hiiesalu[1] · Wael N. Hozzein[2,8] · Kadri Koorem[1] · Mari Moora[1] · Ladislav Mucina[9,10] · Vladimir Onipchenko[11] · Maarja Öpik[1] · Meelis Pärtel[1] · Cherdchai Phosri[12] · Tanel Vahter[1] · Leho Tedersoo[2,13] · Martin Zobel[2,14]

Received: 25 October 2021 / Accepted: 1 February 2022 / Published online: 9 February 2022

Abstract

Arbuscular mycorrhizal (AM) fungi are a ubiquitous group of plant symbionts, yet processes underlying their global assembly — in particular the roles of dispersal limitation and historical drivers — remain poorly understood. Because earlier studies have reported niche conservatism in AM fungi, we hypothesized that variation in taxonomic community composition (i.e., unweighted by taxon relatedness) should resemble variation in phylogenetic community composition (i.e., weighted by taxon relatedness) which reflects ancestral adaptations to historical habitat gradients. Because of the presumed strong dispersal ability of AM fungi, we also anticipated that the large-scale structure of AM fungal communities would track environmental conditions without regional discontinuity. We used recently published AM fungal sequence data (small-subunit ribosomal RNA gene) from soil samples collected worldwide to reconstruct global patterns in taxonomic and phylogenetic community variation. The taxonomic structure of AM fungal communities was primarily driven by habitat conditions, with limited regional differentiation, and there were two well-supported clusters of communities — occurring in cold and warm conditions. Phylogenetic structure was driven by the same factors, though all relationships were markedly weaker. This suggests that niche conservatism with respect to habitat associations is weakly expressed in AM fungal communities. We conclude that the composition of AM fungal communities tracks major climatic and edaphic gradients, with the effects of dispersal limitation and historic factors considerably less apparent than those of climate and soil.

Keywords Community assembly · Metabarcoding · Microbial biogeography · Soil microbes

Introduction

Ecological niches describe how individuals respond to abiotic and biotic factors, with habitat filtering based on niche attributes expected to significantly determine variation in community composition along environmental gradients (Dumbrell et al. 2010; Li et al. 2018). However, an understanding of taxon niche characteristics does not allow the composition and diversity of entire communities to be predicted with precision (Dubuis et al. 2011). Community assembly reflects more than just the sum of species able to exploit a given habitat; it is additionally shaped by various landscape-level and historical processes (Götzenberger et al. 2012). The species pool concept aims to integrate the possible mechanisms underlying community assembly. It holds that local variation in communities depends most importantly on the availability of species, which is driven by historical diversification and migration, and later dispersal events at continental and landscape scales (Zobel 2016). Contemporary patterns of species diversity thus may depend primarily on the pool of available species that are adapted to particular habitat types (Taylor et al. 1990). Yet, empirical evidence concerning the roles of these large-scale and historical drivers is limited, mostly because of methodological challenges.

Disentangling the role of dispersal with the help of descriptive data requires measurement at sufficiently large spatial scales. When such data show that community composition exhibits significant spatial correlation, this is likely to indicate dispersal limitation, i.e., an inability of taxa to reach

Martti Vasar, John Davison, Siim-Kaarel Sepp, Leho Tedersoo, and Martin Zobel contributed equally.

✉ Martti Vasar
martti.vasar@ut.ee

Extended author information available on the last page of the article

potentially suitable habitats (Myers et al. 2013). However, it does not tell us whether this is due to barriers operating over evolutionary (e.g., isolation of biota on separate land masses) or shorter ecological time scales (e.g., caused by impermeable landscape features).

The species pool concept assumes that most organisms exhibit niche conservatism, defined as the retention through evolutionary time of ancestral ecological characteristics (Wiens and Graham 2005), and it frames turnover of community composition along environmental gradients as a reflection of ancient species adaptations (Harrison and Grace 2007). In the presence of niche conservatism, variation in "taxonomic" community structure (i.e., without consideration of relatedness between taxa) and "phylogenetic" structure (i.e., weighted by the relatedness between taxa) should respond to the same ecological gradients (Lu et al. 2016). This also is true for the effects of spatial gradients — a spatial effect on phylogenetic structure indicates ancient migration limitation, while a spatial effect on taxonomic structure represents the cumulative effect of both ancient and subsequent dispersal limitation. In contrast, when niches have diverged within clades or converged between different clades, taxonomic structure still is expected to mirror ecological gradients, but phylogenetic structure is not. Thus, the degree of congruence between "taxonomic" and "phylogenetic" structure provides an indication of the extent to which ancestral relationships underpin contemporary community composition.

While it is desirable to reach broad general conclusions concerning historical and ecological processes, there are practical challenges to compiling the necessary information in parallel: the target of the study should be a well-described organism group for which phylogenetic relationships are defined and extensive field community data are available. Arbuscular mycorrhizal (AM) fungi (phylum Glomeromycota; Tedersoo et al. 2018) are a widespread group of plant root symbionts (associating with > 80% of plants in terrestrial ecosystems; Smith and Read 2008) for which studying global variation in the structure of biotic communities is achievable. An available molecular operational taxonomic unit nomenclature — phylogenetically defined sequence groups roughly corresponding to species-level taxa (virtual taxa [VT]; Öpik et al. 2014) — provides a robust and repeatable basis for describing naturally occurring AM fungal communities.

Previous work has disentangled the roles of climatic and edaphic factors in driving AM fungal taxon distributions and community composition, identifying pH and temperature as important environmental gradients (Davison et al. 2015, 2021). Although there is some evidence of dispersal limitation at varying scales, patterns of widespread VT (Davison et al. (2015) and within-species genotypes (Savary et al. 2018) generally have been interpreted as indicating efficient dispersal abilities among AM fungal taxa. There also is evidence of niche conservatism among AM fungi. For instance, the general architecture of hyphal growth is conserved at the family level (Hart and Reader 2002; Maherali and Klironomos 2007; Powell et al. 2009). Davison et al. (2021) additionally showed that several environmental effects on AM fungal taxon distribution and relative abundance are correlated with phylogeny, indicating that closely related taxa exhibit similar niche optima and widths. While studies on individual taxa provide clues about characteristics of the communities they inhabit, community assembly is modulated by the coincidence or discordance of taxon distributions, and incorporates processes, such as competition and dispersal, that are not predictable from the niche attributes of potentially available taxa. No previous work has explicitly considered the role of ancestral organism-environment relationships in driving contemporary patterns of AM fungal community composition.

When considering evolutionary and spatial drivers, classification of communities can be used to generate intuitive, spatial representations (De Cáceres et al. 2015). At the global scale, the community patterns of dispersal-limited organisms may be described with the help of biogeographic realms — defined as unique areas with unifying features of geography and biota (Udvardy 1975). In contrast, in the absence of dispersal limitation, large-scale variation in communities should track environmental gradients, and global community patterns may be described with the help of biomes — defined as broad cross-regional types of biotic communities and their environments (Mucina 2019). Yet, it is unclear whether any global community types or biogeographical regions are apparent within AM fungal diversity.

Here, we address global variation in the taxonomic and phylogenetic community composition of AM fungi to disentangle the potential roles of spatial and historical factors in driving AM fungal community composition. We expect AM fungal community composition to primarily track previously identified environmental gradients (including climatic and edaphic properties). However, because earlier authors have suggested that AM fungi exhibit niche conservatism (Hart and Reader 2002; Maherali and Klironomos 2007; Powell et al. 2009), we hypothesize that taxonomic and phylogenetic community structures vary in parallel and thus support a core prediction of the species pool hypothesis: turnover of community composition along environmental gradients is a reflection of ancestral adaptations. Given the global distribution of many AM fungal taxa (Davison et al. 2015), we anticipate a pattern of cross-regional clustering in community composition, following the structure of biomes more than biogeographic realms.

Material and methods

Sequencing data and bioinformatics

We used data from Davison et al. (2021) on the composition of AM fungal communities along with metadata describing climatic and soil characteristics in 327 soil samples collected worldwide. While Davison et al. (2021) described the niche attributes of individual AM fungal taxa, we compiled those taxon-level data into whole-community descriptions. A full description of data collection and laboratory methods is given in Davison et al. (2021). Briefly, at each sampling location, a site that was least disturbed by human activities was identified and about 20 g of topsoil (1–5 cm) was collected from < 40 randomly located points within an approximately 50 × 50 m sampling area. For further analysis, the samples were pooled per site, dried, and homogenized. A 2 g subsample of soil was collected from the pooled sample for molecular analysis; the remainder was used for geochemical analysis. The climatic variables mean annual temperature, mean annual precipitation, and precipitation seasonality (coefficient of variation of monthly precipitation) were taken from the CHELSA database (Karger et al. 2017). Soil pH, total N, organic C, and plant available P, K, Mg, and Ca were determined as described in Davison et al. (2021). Sampling locations were assigned to biogeographic realms following Olson et al. (2001). The historical stability of biomes at sampling locations was estimated by comparing the Olson et al. (2001) current biome classification with an analogous classification for the last glacial maximum (approximately 21kyBP; Ray and Adams 2001). Where the biome remained the same at both times, the sampling location was classified as stable; where the biome classification differed, the sampling location was classified as unstable.

DNA was extracted from 2 g of dried soil using a PowerMax® Soil DNA Isolation Kit (MoBio Laboratories, Carlsbad, USA), with some modifications (described in Davison et al. 2021). AM fungal sequences were amplified from soil DNA extracts using AM fungal specific primers for the small-subunit (SSU) ribosomal RNA gene: WANDA (Dumbrell et al. 2011) and AML2 (Lee et al. 2008). Samples were sequenced on an Illumina MiSeq platform, using a 2 × 300 bp paired-read sequencing approach. Demultiplexed paired-end reads were analyzed as described by Vasar et al. (2017) using the gDAT pipeline (Vasar et al. 2021) which includes the following steps: barcode and primer were allowed to have 1 mismatch for both reads; removal of barcode and primer sequences; average quality threshold for both reads of 30; combining paired-end reads using a 75% match threshold; chimera identification and removal using default parameters in reference mode against the MaarjAM database, and a BLAST (Camacho et al. 2009) search against VT in the MaarjAM database (using 97% identity and 95% alignment length thresholds).

Statistical analyses

Samples with low sequence counts (< 100 reads) were removed, and, to normalize remaining sequence count data, we implemented the variance stabilizing transformation (using R package DESeq2 v1.28.1; Love et al. 2014), as suggested by McMurdie and Holmes (2014). The method uses fitted dispersion-mean relationships to normalize data with respect to sample size (sequencing depth of individual samples) and variance. A set of independent predictor variables was selected based on pairwise correlations and variance inflation factors (VIF) in full dbRDA models (see below), retaining the variables that exhibited the strongest explanatory power. Thus, we removed N (correlated with organic C $r = 0.83$) and Mg (correlated with Ca $r = 0.5$) from final analyses (all variable VIF values for final models ≤ 4.04). To address phylogenetic community composition, we calculated a neighbor-joining tree containing the type sequences of all VT in the MaarjAM database (Öpik et al. 2010) and some unpublished phylogroups, using the methods described in Öpik et al. (2013).

Abiotic drivers of soil communities were identified using a combination of distance-based redundancy analysis (dbRDA; *vegan* package in R; Oksanen et al. 2008) and generalized dissimilarity modelling (GDM; Guerin et al. 2021), because the former can incorporate categorical predictor variables while the latter allows non-linear effects to be identified and a spatial distance matrix to be included. In the dbRDA analyses, variation in sample resemblance matrices (Bray–Curtis dissimilarity or phylogenetic distance) was modelled against all measured abiotic variables. Phylogenetic distance was represented by a pairwise, between-sample phylogenetic distance matrix, weighted by AM fungal VT relative abundances (constructed using the function comdist from the *picante* package; Kembel et al. 2010). The significance of effects was measured using permutation ($n = 999$). In GDM models, all continuous abiotic variables were included along with a great-circle distance matrix representing spatial distances between sampling locations. Non-linear effects were visualized by plotting variable i-splines.

Statistical clustering and interpolation techniques were used to visualize global spatial patterns in the taxonomic and phylogenetic composition of AM fungal communities. First, *k*-means clustering was used to assign samples to clusters based on between-sample similarity in the form of PCoA vectors derived from Bray–Curtis or phylogenetic matrices. Clustering with values of *k* 2–16 with 100

random starting configurations per k was used. Silhouette scores, derived from the difference between within- and between-cluster distances, were used to ascertain the best-fitting cluster configurations among the range of *k*-values, and bootstrapping was used to assess the stability of clustering solutions (using Jaccard similarity; Hennig 2007). Afterwards, interpolated maps of cluster distribution were produced using weighted categorical *k*-nearest neighbor (KNN) classification (kknn() in R package kknn; Schliep et al. 2016) with the soil sample cluster identities as the training set, and a $0.5° \times 0.5°$ map grid as the test set. The weights were based on great-circle distances of the *k*-nearest training set points from the respective test set grid point. Grid cell cluster identity was taken as the class with the highest KNN predicted probability for that location. The *k*-value for KNN interpolation ($k = 18$; note that this is different from the *k* parameter used in the *k*-means clustering) was set as the rounded square root of the number of samples, based on the suggestion of Duda et al. (2012).

Results

Models of taxonomic community variation achieved reasonable explanatory power (dbRDA $R^2 = 25\%$; GDM deviance explained = 31%). On the basis of dbRDA (*F* value and significance) and GDM (variable importance and significance) results, soil pH and mean annual temperature were the most important drivers of the taxonomic community composition of AM fungi (Table 1, Fig. 1a, b). Biogeographic realm was the next most important predictor, though somewhat less so, based on the dbRDA results. The effects of temperature and pH on community taxonomic composition were non-linear, with community turnover greatest at low pH and high temperatures (Fig. 2). Other soil, climatic, spatial, and historical factors exerted weaker effects. dbRDA ordination separated samples in the Palearctic and Nearctic (i.e., Holarctic) from those in other realms with warmer climates (Fig. 1a).

Models of phylogenetic community variation achieved comparatively low explanatory power (dbRDA $R^2 = 8\%$; GDM deviance explained = 2%). Mean annual temperature

Table 1 Distance-based redundancy analysis (dbRDA) and generalized dissimilarity modelling (GDM) of community composition among AM fungal communities using taxonomic (upper section) and phylogenetic (lower section) composition-based distances. *MAT*, mean annual temperature; *MAP*, mean annual precipitation; *SeaPrec*, seasonal variation in precipitation; stability — categorical value showing biome continuity from the last glacial maximum

Variable	Summary of dbRDA					Summary of GDM	
	Residuals	Sum of squares	*F*	*p*	VIF	Variable importance	*p*
Taxonomic composition-based distance							
pH	84.227	2.81	9.108	**0.001**	2.142	**25.915**	**< 0.01**
P		0.248	0.802	0.775	1.87	0.009	0.78
K		0.411	1.331	0.108	2.26	0.353	0.22
Ca		0.572	1.854	**0.010**	2.273	0.384	0.2
OrgC		0.623	2.018	**0.006**	1.862	**2.738**	**< 0.01**
MAT		2.268	7.353	**0.001**	2.689	**29.259**	**< 0.01**
MAP		0.659	2.135	**0.004**	2.095	**3.061**	**< 0.01**
Realm		7.015	3.79	**0.001**	4.038	-	-
SeaPrec		0.629	2.038	**0.008**	1.374	0.002	0.74
Stability		0.53	1.718	**0.026**	1.246	-	-
Spatial	-	-	-	-	-	**6.585**	**< 0.01**
Phylogenetic composition-based distance							
pH	42.173	0.318	2.061	**0.001**	2.142	4.506	0.56
P		0.137	0.888	0.806	1.87	0	0.98
K		0.126	0.817	0.71	2.26	0	1
Ca		0.144	0.932	0.757	2.273	0.654	0.66
OrgC		0.157	1.019	0.342	1.862	0	0.98
MAT		0.354	2.291	**0.001**	2.689	**86.841**	**0.02**
MAP		0.167	1.079	0.227	2.095	0	0.92
Realm		0.987	1.065	0.109	4.038	-	-
SeaPrec		0.163	1.056	0.201	1.374	0.046	0.7
Stability		0.155	1.001	0.392	1.246	-	-
Spatial	-	-	-	-	-	1.653	0.54

Significant values are shown in bold

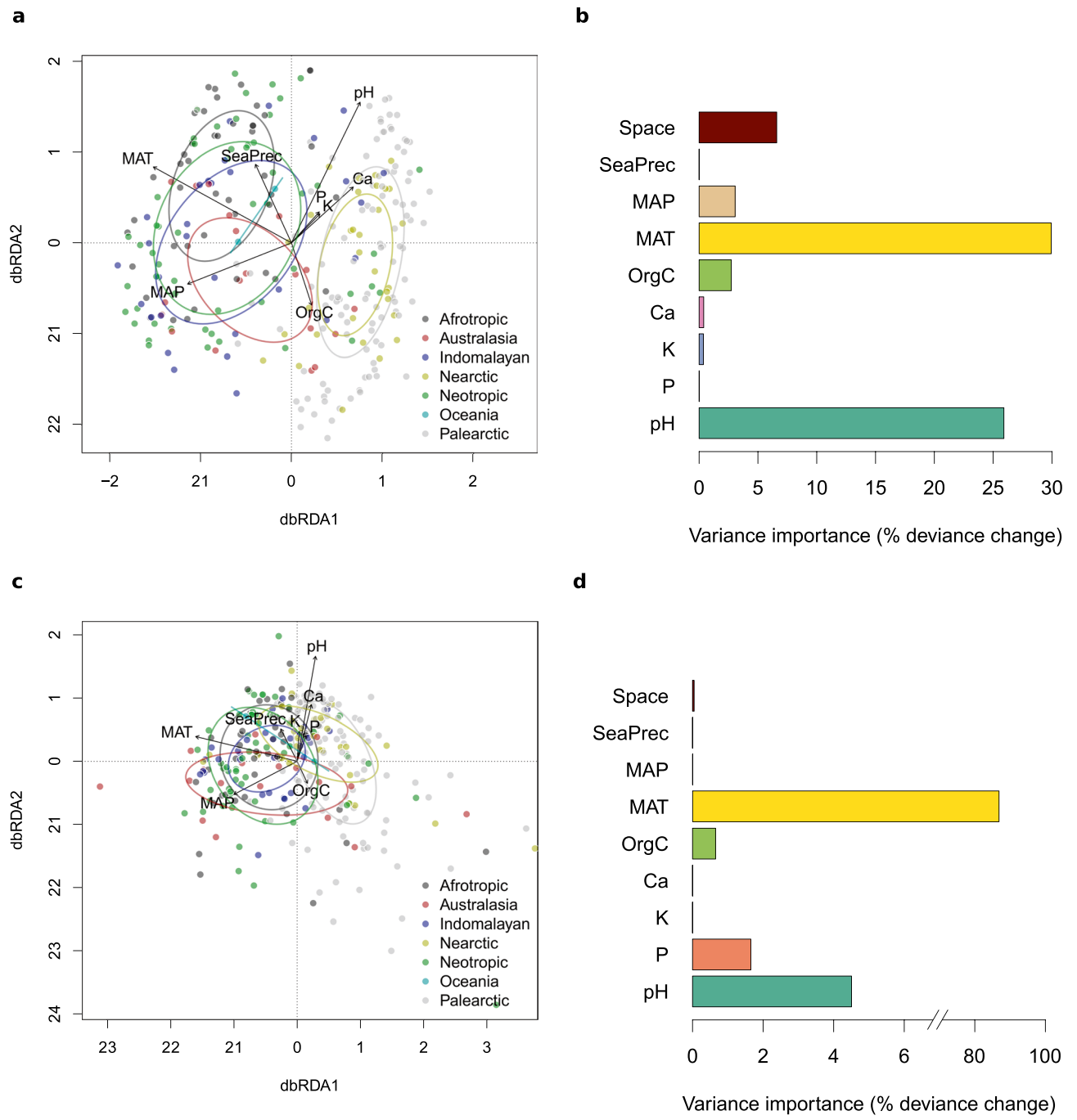

Fig. 1 The main global determinants of taxonomic **a**, **b** and phylogenetic **c**, **d** soil AM fungal community composition. **a** and **c** distance-based redundancy analysis (dbRDA) ordination biplots. Arrows indicate the direction of maximum change in environmental variables. Ellipses show 1 standard deviation around the centroid of different biogeographic realms (using the same colors used to distinguish points on the plot); **b** and **d** variable importance in generalized dissimilarity models (GDM), measured as the change in model deviance caused by variable permutation. The GDM models do not include the categorical predictors of the dbRDA models but do include a spatial distance matrix that is not in dbRDA models. OrgC, % organic carbon; MAT, mean annual temperature; MAP, mean annual precipitation; SeaPrec, seasonal variation in precipitation

(dbRDA and GDM) together with soil pH and biogeographic realm (both only apparent in the dbRDA) had significant but weak effects on the phylogenetic community composition of AM fungi (Table 1, Fig. 1c, d). The effects of other factors were nonsignificant. Samples from the Palearctic and Nearctic realms clustered together on the ordination biplot (Fig. 1c).

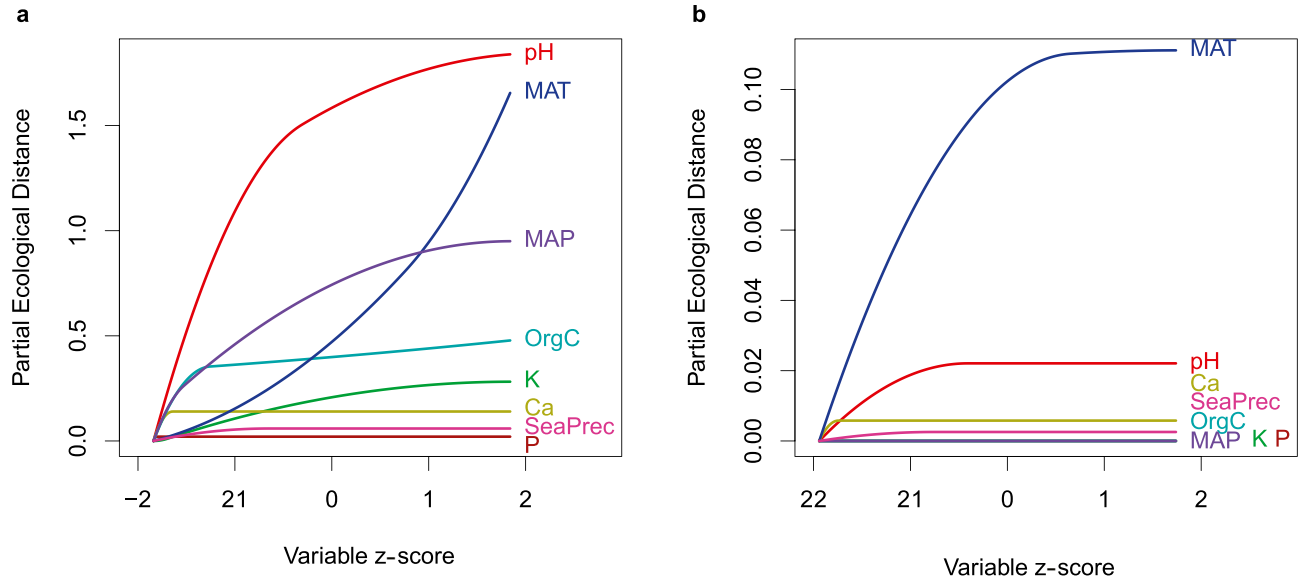

Fig. 2 I-splines (partial ecological distance) from generalized dissimilarity modelling (GDM) of AM fungal **a** taxonomic and **b** phylogenetic community composition. The slope of the i-spline indicates the rate of compositional turnover and how it changes along the range of the variable (variable *z*-score). For instance, in **a**, community turnover is rapid at low values of pH and slower at high values, but the converse is true of turnover in relation to MAT. OrgC, % organic carbon; MAT, mean annual temperature; MAP, mean annual precipitation; SeaPrec, seasonal variation in precipitation. Note the different scales of the ordinate axes

We used unsupervised *k*-means clustering to categorize variation in global soil AM fungal communities. Clusters were defined on the basis of compositional variation between samples, with the optimal number of clusters (*k*) selected using silhouette plots (Fig. S1). Classification into two clusters ($k=2$) was best supported (Fig. S1). In this configuration, the first taxonomic cluster (cluster 1) was mostly distributed at high latitudes of the northern hemisphere, exhibiting an approximately Holarctic distribution, while the second cluster had a pantropical and southern temperate distribution (Fig. 3a). The habitat conditions of sampling sites in cluster 1 were on average colder and also slightly drier and slightly more alkaline, than those in cluster 2 (Fig. S2). Phylogenetic community composition also was classified best into two clusters (Fig. 3b). The first phylogenetic cluster was distributed at high latitudes in the northern hemisphere only; the second over the rest of the globe. The habitat conditions of sampling sites in phylogenetic cluster 1 were on average colder and also slightly more acid and drier than those of sampling sites in phylogenetic cluster 2 (Fig. S2). Taxonomic and phylogenetic community classifications both produced clusters that differed in the representation of different AM fungal families (Fig. 4). The northern high-latitude clusters (cluster 1 in both taxonomic and phylogenetic community classifications) were characterized by diminished abundance of Glomeraceae (for the phylogenetic community classification) and increased relative abundance of several other families, including Claroideoglomeraceae (both community classifications) and Acaulosporaceae (phylogenetic community classification).

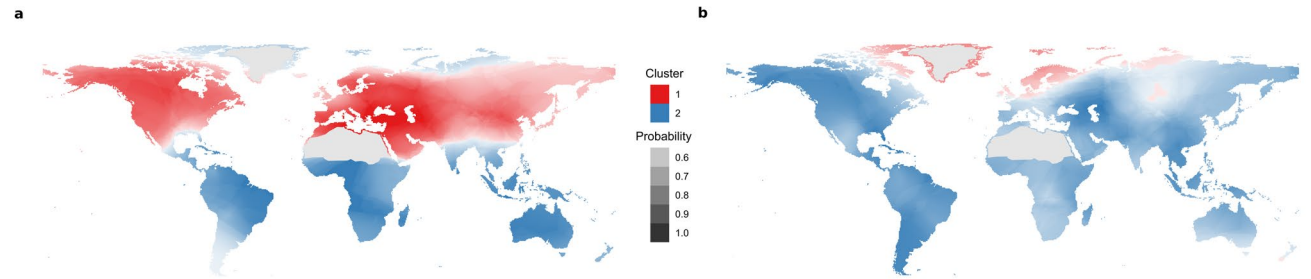

Fig. 3 Global interpolated maps of AM fungal communities using *k*-means clustering ($k=2$) of **a** taxonomic and **b** phylogenetic distances. Greenland and the Sahara region were excluded from interpolations because of insufficient sampling and highly contrasted, different abiotic conditions

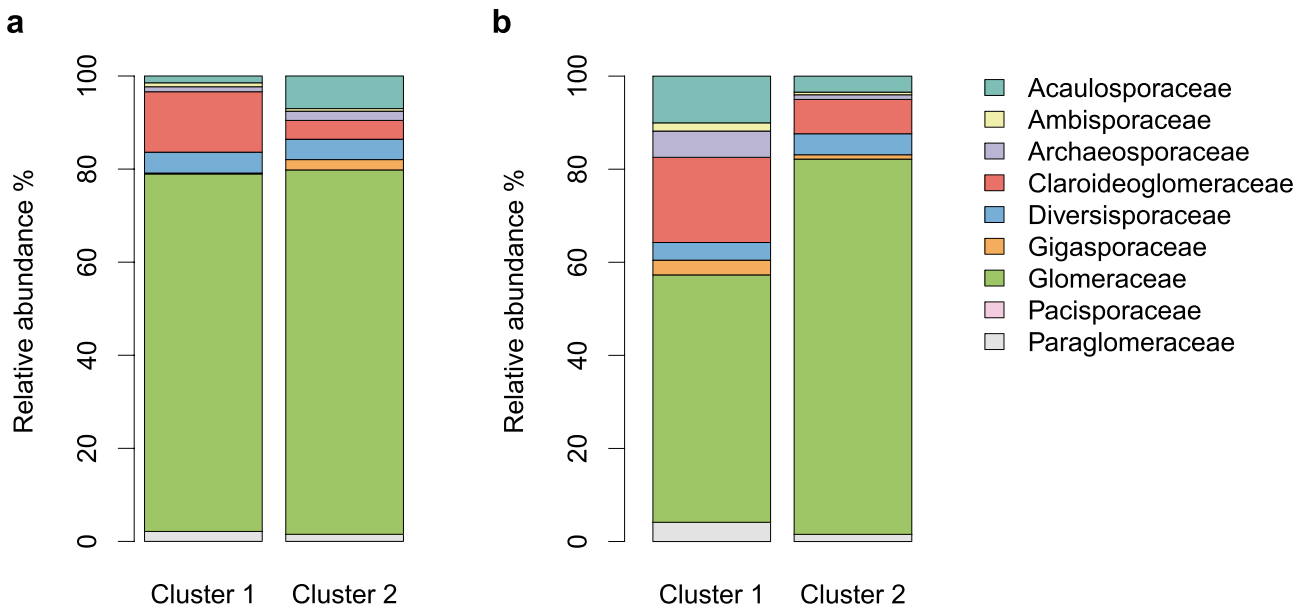

Fig. 4 AM fungal family-level composition of **a** taxonomic and **b** phylogenetic community clusters. Stacked bars indicate the relative abundance of reads corresponding to different AM fungal families. Cluster labels follow the numbering in Fig. 3. Note that there is no direct correspondence between clusters in both classifications, with each bar only comprising read counts from those samples within it

Discussion

The global-scale taxonomic structure of AM fungal communities is driven primarily by habitat filtering, notably by temperature and soil pH, although some biogeographic and finer-scale spatial structure is apparent. The same factors significantly drive phylogenetic community structure, but the relationships are considerably weaker. Two widely distributed types of AM fungal communities were apparent, integrating the effects of environmental conditions and biogeographic history. The types broadly are associated with cooler versus warmer climates and differ with respect to their biogeographic history — approximately corresponding to the ancient landmasses of Gondwana and Laurasia, and with the first type also including recently glaciated areas. The hypothesis that taxonomic and phylogenetic community structures vary in parallel thus was supported only partially because habitat conditions had considerably stronger effects on taxonomic than on phylogenetic structure. These results indicate only weak niche conservatism with respect to the major environmental drivers such as temperature and soil pH. Also, there was little indication of dispersal limitation in either the taxonomic or phylogenetic data sets with global community types representing cross-regional clustering in community composition, following the structure of biomes more than realms. Thus, historical processes such as diversification and migration, which constitute the essence of the species pool hypothesis, seem to play only moderate roles in determining the assembly of contemporary AM fungal communities.

There have been no previous attempts to directly compare taxonomic and phylogenetic community structure of AM fungi at the global scale. The present analysis indicates broadly similar gradients structuring both taxonomic and phylogenetic community composition. The former largely tracks temperature and soil pH gradients and is consistent with previous studies at the species (Antunes et al. 2011; Davison et al. 2021) and community levels (Lekberg et al. 2011; Dumbrell et al. 2010; Davison et al. 2015). Overall, the effects of environmental factors on phylogenetic community structure were much weaker than those on taxonomic structure, providing little support for strong niche conservatism in AM fungi. This result contrasts with earlier findings (Hart and Reader 2002; Maherali and Klironomos 2007; Powell et al. 2009). The discrepancy may be because of variation in the traits that are conserved among closely related AM fungal species (Roy et al. 2019). Prior results indicating strong niche conservatism mainly have concerned the general architecture of hyphal growth, which is conserved at the family level (Hart and Reader 2002; Maherali and Klironomos 2007; Powell et al. 2009). Such traits may drive fungal competition (Maherali and Klironomos 2007), host specificity (Hart and Reader 2002; Powell et al. 2009), or response to disturbances (Chagnon et al. 2013; Weber et al. 2019), but perhaps do not determine responses to basic climatic and soil factors. It is notable that the VT-level niche conservatism in relation to abiotic gradients identified by Davison et al. (2021) did not translate into clear phylogenetic structure at the community level. It is difficult, however, to compare these results directly. Davison et al. (2021) noted

conservatism only at particular phylogenetic lags (generally among relatively closely related VT); and it is unclear how that would be expected to translate into a community measure that incorporates the full range of relatedness between taxa. Nevertheless, it also is possible that community assembly processes to some extent obscure the cumulative effect of niche conservatism among component taxa. Instead, our results are consistent with those of Davison et al. (2015) who concluded that the wide distribution of AM fungal VT in conditions in which VT taxogenesis generally occurred after major continental reconfigurations indicates efficient recent dispersal rather than ancient vicariance. We supplement this conclusion by arguing that AM fungal communities do not exhibit a clear imprint of abiotic niche conservatism. Thus, we cannot support the core view of the species pool concept that the turnover of community composition along environmental gradients primarily reflects ancient species adaptations to habitat conditions in the regions where diversification occurred (Harrison and Grace 2007; Zobel et al. 2011). At the same time, we fully agree with Spasojevic et al. (2018) that species pools differ substantially in functional-trait composition, which can strongly influence current community assembly as well as biodiversity responses to environmental change.

Two broad types of AM fungal community were distinguished by the best-supported clustering solution, primarily representing climatic zones and, in the case of taxonomic community composition, also large-scale differences in soil pH. Thus, one broad AM fungal community type is present in cold regions and a second community type in warm regions, the latter being somewhat skewed towards the southern hemisphere. The warm-climate community type also occurs in areas that are characterized by a large AM fungal taxon pool and a high density and diversity of host plants, especially in areas containing tropical grasslands during the last glacial maximum (Pärtel et al. 2017). After the last glacial maximum, the proportion of AM plants in the vegetation of the northern hemisphere decreased (Zobel et al. 2018). At present, the cool regions of the northern hemisphere are predominantly covered by forests dominated by ectomycorrhizal trees, the abundance of which is negatively correlated with the diversity of AM fungi (Toussaint et al. 2020). In our phylogeny-based clustering, one community type was distributed only at very high latitudes of the northern hemisphere (including southern Scandinavia and the British Isles). The Glomeraceae family, of which occurrence often is correlated with low root colonization by pathogens but limited improved nutrition (Maherali and Klironomos 2007), was most abundantly represented in warm climates. Claroideoglomeraceae and Acaulosporaceae — cold tolerant families (Davison et al. 2021) with the latter also commonly reported in low pH environments — were notably abundant in the northern clusters.

We conclude that habitat filtering in relation to major climatic and edaphic gradients largely explains the taxonomic composition of AM fungal communities, while the phylogenetic composition of the AM fungal communities did not correlate strongly with either environmental or spatial gradients. Evidence of niche conservatism and spatial structure thus was quite limited, indicating that the effects on community composition of dispersal limitation and historic factors were weak. Cluster analysis distinguished pan-regional AM fungal community types that tracked climatic, and to a lesser degree soil characteristics, resembling biomes more than biogeographic realms. These observations provide insight into the ecological correlates and evolutionary context of AM fungal community composition, and additionally offer a practical template for studying the large-scale diversity patterns of organisms.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s00572-022-01072-7.

Funding This study was supported by the European Regional Development Fund (Centre of Excellence EcolChange), Estonian Research Council (PRG609, PRG632, PRG1065, PRG1170, MOBTP105), the Russian Ministry of Science and Higher Education (# 075–15-2021–1396), and the University of Tartu (PLTOM20903).

Declarations

Conflict of interest The authors declare no competing interests.

References

Antunes PM, Koch AM, Morton JB, Rillig MC, Klironomos JN (2011) Evidence for functional divergence in arbuscular mycorrhizal fungi from contrasting climatic origins. New Phytol 189:507–514. https://doi.org/10.1111/j.1469-8137.2010.03480.x

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:1–9. https://doi.org/10.1186/1471-2105-10-421

Chagnon PL, Bradley RL, Maherali H, Klironomos JN (2013) A trait-based framework to understand life history of mycorrhizal fungi. Trends Plant Sci 18:484–491. https://doi.org/10.1016/j.tplants.2013.05.001

Davison J, Moora M, Öpik M et al (2015) Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. Science 349:970–973. https://doi.org/10.1126/science.aab1161

Davison J, Moora M, Semchenko M, Adenan SB, Ahmed T, Akhmetzhanova AA, Alatalo JM, Al-Quraishy S, Andriyanova E, Anslan S, Bahram M (2021) Temperature and pH define the realised niche space of arbuscular mycorrhizal fungi. New Phytol 231:763–776. https://doi.org/10.1111/nph.17240

De Cáceres M, Chytrý M, Agrillo E et al (2015) A comparative framework for broad-scale plot-based vegetation classification. Appl Veg Sci 18:543–560. https://doi.org/10.1111/avsc.12179

Dubuis A, Pottier J, Rion V, Pellissier L, Theurillat JP, Guisan A (2011) Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. Divers Distrib 17:1122–1131. https://doi.org/10.1111/j.1472-4642.2011.00792.x

Duda P, Hayashi Y, Jaworski M (2012) On the strong convergence of the orthogonal series-type kernel regression neural networks in a non-stationary environment. In International conference on artificial intelligence and soft computing (pp. 47–54). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-29347-4_6

Dumbrell AJ, Ashton PD, Aziz N et al (2011) Distinct seasonal assemblages of arbuscular mycorrhizal fungi revealed by massively parallel pyrosequencing. New Phytol 190:794–804. https://doi.org/10.1111/j.1469-8137.2010.03636.x

Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH (2010) Relative roles of niche and neutral processes in structuring a soil microbial community. ISME J 4:337–345. https://doi.org/10.1038/ismej.2009.122

Götzenberger L, de Bello F, Bråthen KA, Davison J, Dubuis A, Guisan A, Lepš J, Lindborg R, Moora M, Pärtel M, Pellissier L (2012) Ecological assembly rules in plant communities—approaches, patterns and prospects. Biol Rev 87:111–127. https://doi.org/10.1111/j.1469-185X.2011.00187.x

Guerin GR, Williams KJ, Leitch E, Lowe AJ, Sparrow B (2021) Using generalised dissimilarity modelling and targeted field surveys to gap-fill an ecosystem surveillance network. J Appl Ecol 58:766–776. https://doi.org/10.1111/1365-2664.13814

Harrison S, Grace JB (2007) Biogeographic affinity helps explain productivity-richness relationships at regional and local scales. Am Nat 170:S5–S15. https://doi.org/10.1086/519010

Hart MM, Reader RJ (2002) Host plant benefit from association with arbuscular mycorrhizal fungi: variation due to differences in size of mycelium. Biol Fertil Soils 36:357–366. https://doi.org/10.1007/s00374-002-0539-4

Hennig C (2007) Cluster-wise assessment of cluster stability. Comput Stat Data Anal 52:258–271. https://doi.org/10.1016/j.csda.2006.11.025

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M (2017) Climatologies at high resolution for the earth's land surface areas. Scientific Data 4:1–20. https://doi.org/10.1038/sdata.2017.122

Kembel S (2010) An introduction to the picante package. R package version 1.3.1. https://CRAN.R-project.org/package=picante

Lee J, Lee S, Young JPW (2008) Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. FEMS Microbiol Ecol 65:339–349. https://doi.org/10.1111/j.1574-6941.2008.00531.x

Lekberg Y, Meadow J, Rohr JR, Redecker D, Zabinski CA (2011) Importance of dispersal and thermal environment for mycorrhizal communities: lessons from Yellowstone National Park. Ecology 92:1292–1302. https://doi.org/10.1890/10-1516.1

Li Y, Shipley B, Price JN et al (2018) Habitat filtering determines the functional niche occupancy of plant communities worldwide. J Ecol 106:1001–1009. https://doi.org/10.1111/1365-2745.12802

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:1–21. https://doi.org/10.1186/s13059-014-0550-8

Lu HP, Yeh YC, Sastri AR, Shiah FK, Gong GC, Hsieh CH (2016) Evaluating community–environment relationships along fine to broad taxonomic resolutions reveals evolutionary forces underlying community assembly. ISME J 10:2867–2878. https://doi.org/10.1038/ismej.2016.78

Maherali H, Klironomos JN (2007) Influence of phylogeny on fungal community assembly and ecosystem functioning. Science 316:1746–1748. https://doi.org/10.1126/science.1143082

McMurdie PJ, Holmes S (2014) Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput Biol 10:e1003531. https://doi.org/10.1371/journal.pcbi.1003531

Mucina L (2019) Biome: evolution of a crucial ecological and biogeographical concept. New Phytol 222:97–114. https://doi.org/10.1111/nph.15609

Myers JA, Chase JM, Jiménez I, Jørgensen PM, Araujo-Murakami A, Paniagua-Zambrana N, Seidel R (2013) Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. Ecol Lett 16:151–157. https://doi.org/10.1111/ele.12021

Oksanen J, Blanchet FG, Kindt R, Legendre P, O'hara RB, Simpson GL, Solymos P, Stevens MH, Wagner H (2008) vegan: community ecology package. R package version 1.15–1. https://CRAN.R-project.org/package=vegan

Olson DM, Dinerstein E, Wikramanayake ED et al (2001) Terrestrial ecoregions of the world: a new map of life on Earth: a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. Bioscience 51:933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Öpik M, Davison J, Moora M, Zobel M (2014) DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences. Botany 92:135–147. https://doi.org/10.1139/cjb-2013-0110

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 Phytol 188:223–241. https://doi.org/10.1111/j.1469-8137.2010.03334.x

Öpik M, Zobel M, Cantero JJ et al (2013) Global sampling of plant roots expands the described molecular diversity of arbuscular mycorrhizal fungi. Mycorrhiza 23:411–430. https://doi.org/10.1007/s00572-013-0482-2

Pärtel M, Öpik M, Moora M et al (2017) Historical biome distribution and recent human disturbance shape the diversity of arbuscular mycorrhizal fungi. New Phytol 216:227–238. https://doi.org/10.1111/nph.14695

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 Royal Soci B: Biol Sci 276:4237–4245. https://doi.org/10.1098/rspb.2009.1015

Ray N, Adams J (2001) A GIS-based vegetation map of the world at the last glacial maximum (25,000–15,000 BP). Internet Archaeology 11

Roy J, Mazel F, Sosa-Hernández MA, Dueñas JF, Hempel S, Zinger L, Rillig MC (2019) The relative importance of ecological drivers of arbuscular mycorrhizal fungal distribution varies with taxon phylogenetic resolution. New Phytol 224:936–948. https://doi.org/10.1111/nph.16080

Savary R, Masclaux FG, Wyss T, Droh G, Corella JC, 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 J 12:17–30. https://doi.org/10.1038/ismej.2017.153

Schliep K, Klaus Hechenbichler K (2016) kknn: weighted k-nearest neighbors. R package version 1.3.1. https://CRAN.R-project.org/package=kknn

Smith SE, Read DJ (2008) Mycorrhizal symbiosis. Academic Press, Amsterdam

Spasojevic MJ, Catano CP, LaManna JA, Myers JA (2018) Integrating species traits into species pools. Ecology 99(6):1265–1276. https://doi.org/10.1002/ecy.2220

Taylor DR, Aarssen LW, Loehle C (1990) On the relationship between r/K selection and environmental carrying capacity: a new habitat templet for plant life history strategies. Oikos 58:239–250. https://doi.org/10.2307/3545432

Tedersoo L, Sánchez-Ramírez S, 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. https://doi.org/10.1007/s13225-018-0401-0

Toussaint A, Bueno G, Davison J et al (2020) Asymmetric patterns of global diversity among plants and mycorrhizal fungi. J Veg Sci 31:355–366. https://doi.org/10.1111/jvs.12837

Udvardy MD (1975) A classification of the biogeographical provinces of the world. Morges: International Union for Conservation of Nature and Natural Resources.

Vasar M, Andreson R, Davison J, Jairus T, Moora M, Remm M, Young JP, Zobel M, Öpik M (2017) Increased sequencing depth does not increase captured diversity of arbuscular mycorrhizal fungi. Mycorrhiza 27:761–773. https://doi.org/10.1007/s00572-017-0791-y

Vasar M, Davison J, Neuenkamp L, Sepp SK, Young JP, Moora M, Öpik M (2021) User-friendly bioinformatics pipeline gDAT (graphical downstream analysis tool) for analysing rDNA sequences. Mol Ecol Resour 21:1380–1392. https://doi.org/10.1111/1755-0998.13340

Weber SE, Diez JM, Andrews LV, Goulden ML, Aronson EL, Allen MF (2019) Responses of arbuscular mycorrhizal fungi to multiple coinciding global change drivers. Fungal Ecol 40:62–71. https://doi.org/10.1016/j.funeco.2018.11.008

Wiens JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. Annu Rev Ecol Evol Syst 36:519–539. https://doi.org/10.1146/annurev.ecolsys.36.102803.095431

Zobel M (2016) The species pool concept as a framework for studying patterns of plant diversity. J Veg Sci 27:8–18. https://doi.org/10.1111/jvs.12333

Zobel M, Otto R, Laanisto L, Naranjo-Cigala A, Pärtel M, Maria Fernandez-Palacios J (2011) The formation of species pools: historical habitat abundance affects current local diversity. Glob Ecol Biogeogr 20:251–259. https://doi.org/10.1111/j.1466-8238.2010.00593.x

Zobel M, Davison J, Edwards ME, Brochmann C, Coissac E, Taberlet P, Willerslev E, Moora M (2018) Ancient environmental DNA reveals shifts in dominant mutualisms during the late Quaternary. Nat Commun 9:1–9. https://doi.org/10.1038/s41467-017-02421-3

Authors and Affiliations

Martti Vasar[1] · John Davison[1] · Siim-Kaarel Sepp[1] · Jane Oja[1] · Saleh Al-Quraishy[2] · C. Guillermo Bueno[1] · Juan José Cantero[3,4] · Ezequiel Chimbioputo Fabiano[5] · Guillaume Decocq[6] · Lauchlan Fraser[7] · Inga Hiiesalu[1] · Wael N. Hozzein[2,8] · Kadri Koorem[1] · Mari Moora[1] · Ladislav Mucina[9,10] · Vladimir Onipchenko[11] · Maarja Öpik[1] · Meelis Pärtel[1] · Cherdchai Phosri[12] · Tanel Vahter[1] · Leho Tedersoo[2,13] · Martin Zobel[2,14]

[1] Department of Botany, Institute of Ecology and Earth Sciences, University of Tartu, 2 Liivi St, 50409 Tartu, Estonia

[2] Zoology Department, College of Science, King Saud University, Riyadh, Saudi Arabia

[3] CONICET, Instituto Multidisciplinario de Biología Vegetal, Universidad Nacional de Córdoba, Córdoba, Argentina

[4] Departamento de Biología Agrícola, Facultad de Agronomía Y Veterinaria, Universidad Nacional de Río Cuarto, Córdoba, Argentina

[5] Department of Wildlife Management and Ecotourism, University of Namibia, Katima Mulilo, Namibia

[6] Ecologie et Dynamique des Systèmes Anthropisés, Jules Verne University of Picardie, Amiens, France

[7] Department of Natural Resource Sciences, Thompson Rivers University, Kamloops, BC, Canada

[8] Botany and Microbiology Department, Faculty of Science, Beni-Suef University, Beni-Suef, Egypt

[9] Iluka Chair in Vegetation Science and Biogeography, Harry Butler Institute, Murdoch University, Murdoch, Perth, Australia

[10] Department of Geography and Environmental Studies, Stellenbosch University, Stellenbosch, South Africa

[11] Department of Ecology and Plant Geography, Faculty of Biology, Moscow Lomonosov State University, Moscow, Russia

[12] Department of Biology, Nakhon Phanom University, Nakhon Phanom, Thailand

[13] Mycology and Microbiology Center, University of Tartu, Tartu, Estonia

[14] Department of Botany, University of Tartu, Tartu, Estonia