New Phytologist

Research

Temperature and pH define the realised niche space of arbuscular mycorrhizal fungi

John Davison[1]*, Mari Moora[1]*, Marina Semchenko[1,2]*, Sakeenah Binte Adenan[3], Talaat Ahmed[3], Asem A. Akhmetzhanova[4], Juha M. Alatalo[3], Saleh Al-Quraishy[5], Elena Andriyanova[6], Sten Anslan[1], Mohammad Bahram[7], Amgaa Batbaatar[8], Charlotte Brown[8], C. Guillermo Bueno[1], James Cahill[8], Juan José Cantero[9,10], Brenda B. Casper[11], Mikhail Cherosov[12], Saida Chideh[13], Ana P. Coelho[14], Matthew Coghill[15], Guillaume Decocq[16], Sergey Dudov[4], Ezequiel Chimbioputo Fabiano[17], Vladimir E. Fedosov[4,18], Lauchlan Fraser[15], Sydney I. Glassman[19], Aveliina Helm[1], Hugh A. L. Henry[20], Bruno Hérault[21,22,23], Indrek Hiiesalu[1], Inga Hiiesalu[1], Wael N. Hozzein[5,24], Petr Kohout[25,26], Urmas Kõljalg[1], Kadri Koorem[1], Lauri Laanisto[27], Ülo Mander[1], Ladislav Mucina[28,29], Jean-Pierre Munyampundu[30], Lena Neuenkamp[1,31], Ülo Niinemets[32], Casper Nyamukondiwa[33], Jane Oja[1], Vladimir Onipchenko[4], Meelis Pärtel[1], Cherdchai Phosri[34], Sergei Põlme[1,35], Kersti Püssa[1], Argo Ronk[11], Alessandro Saitta[36], Olivia Semboli[37], Siim-Kaarel Sepp[1], Alexey Seregin[4], Surya Sudheer[1], Clara P. Peña-Venegas[38], Claudia Paz[39], Tanel Vahter[1], Martti Vasar[1], Annelies J. Veraart[40], Leho Tedersoo[1]*, Martin Zobel[4,41]* and Maarja Öpik[1]*

[1]Institute of Ecology and Earth Sciences, University of Tartu, Tartu 51005, Estonia; [2]School of Earth and Environmental Sciences, University of Manchester, Manchester, M13 9PL, UK; [3]Environmental Science Centre, Qatar University, Doha 2713, Qatar; [4]Department of Ecology and Plant Geography, Faculty of Biology, Moscow Lomonsov State University, Moscow 119991, Russia; [5]Zoology Department, College of Science, King Saud University, Riyadh 11451, Saudi Arabia; [6]Institute of Biological Problems of the North Far East Branch of Russian Academy of Sciences, Magadan 685000, Russia; [7]Department of Ecology, Swedish University of Agricultural Sciences, Uppsala 756 51, Sweden; [8]Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada; [9]Instituto Multidisciplinario de Biología Vegetal, Universidad Nacional de Córdoba, CONICET, Córdoba X5000HUA, Argentina; [10]Departamento de Biología Agrícola, Facultad de Agronomía y Veterinaria, Universidad Nacional de Río Cuarto, Córdoba X5804BYA, Argentina; [11]Department of Biology, University of Pennsylvania, Philadelphia, PA 19104-4544, USA; [12]Institute of Biological Problems of the Cryolithozone, Siberian Branch of the Russian Academy of Sciences, Yakutsk 677000, Russia; [13]Département de Recherche en Sciences de l'Environnement, Université de Djibouti, Private bag 1904, Djibouti, Djibouti; [14]Department of Biology and CESAM, University of Aveiro, Aveiro 3810-193, Portugal; [15]Department of Natural Resource Sciences, Thompson Rivers University, Kamloops, BC V2C 0C8, Canada; [16]Ecologie et Dynamique des Systèmes Anthropisés, Jules Verne University of Picardie, Amiens F-80037, France; [17]Department of Wildlife Management and Ecotourism, University of Namibia, Private bag 1096, Katima Mulilo, Namibia; [18]Botanical Garden-Institute FEB RAS, Vladivostok 690024, Russia; [19]Department of Microbiology and Plant Pathology, University of California, Riverside, CA 92521, USA; [20]Department of Biology, University of Western Ontario, London, ON N6A 5B7, Canada; [21]CIRAD, UPR Forêts et Sociétés, Yamoussoukro, Côte d'Ivoire; [22]Forêts et Sociétés, Université de Montpellier, CIRAD, Montpellier 34000, France; [23]Institut National Polytechnique Félix Houphouët-Boigny, INP-HB, Yamoussoukro, Côte d'Ivoire; [24]Botany and Microbiology Department, Faculty of Science, Beni-Suef University, Bani Suwayf 62511, Egypt; [25]Institute of Microbiology, Czech Academy of Science, Prague 14220, Czechia; [26]Department of Experimental Plant Biology, Faculty of Science, Charles University, Prague 12843, Czechia; [27]Chair of Biodiversity and Nature Tourism, Estonian University of Life Sciences, Tartu 51006, Estonia; [28]Iluka Chair in Vegetation Science and Biogeography, Harry Butler Institute, Murdoch University, Murdoch, Perth, WA 6150, Australia; [29]Department of Geography & Environmental Studies, Stellenbosch University, Stellenbosch 7602, South Africa; [30]School of Science, College of Science and Technology, University of Rwanda, Kigali 3900, Rwanda; [31]Institute of Plant Sciences, University of Bern, Bern 3013, Switzerland; [32]Chair of Crop Science and Plant Biology, Estonian University of Life Sciences, Tartu 51006, Estonia; [33]Department of Biological Sciences and Biotechnology, Botswana International University of Science and Technology, Private bag 16, Palapye, Botswana; [34]Department of Biology, Nakhon Phanom University, Nakhon Phanom 48000, Thailand; [35]Natural History Museum, University of Tartu, Tartu 51014, Estonia; [36]Department of Agricultural, Food and Forest Sciences, University of Palermo, Palermo 90128, Italy; [37]Center of Studies and Research on Pharmacopoeia and Traditional African Medicine, University of Bangui, Bangui, Central African Republic; [38]Instituto Amazónico de Investigaciones Científicas Sinchi, Leticia, Amazonas 910001, Colombia; [39]Departamento de Biodiversidade, Universidade Estadual Paulista, Rio Claro, São Paulo 13506-900, Brazil; [40]Department of Aquatic Ecology and Environmental Biology, Institute for Water and Wetland Research, Radboud University, Nijmegen 6525AJ, the Netherlands; [41]Department of Botany, University of Tartu, Tartu 51005, Estonia

Author for correspondence:
John Davison
Email: john.davison@ut.ee

Summary

• The arbuscular mycorrhizal (AM) fungi are a globally distributed group of soil organisms that play critical roles in ecosystem function. However, the ecological niches of individual AM fungal taxa are poorly understood.

New Phytologist

Received: *8 September 2020*
Accepted: *19 January 2021*

New Phytologist (2021)
doi: 10.1111/nph.17240

Key words: arbuscular mycorrhizal fungi, ecological niche, molecular taxa, niche optimum, niche width, pH, phylogenetic correlation, temperature.

• We collected > 300 soil samples from natural ecosystems worldwide and modelled the realised niches of AM fungal virtual taxa (VT; approximately species-level phylogroups).
• We found that environmental and spatial variables jointly explained VT distribution worldwide, with temperature and pH being the most important abiotic drivers, and spatial effects generally occurring at local to regional scales. While dispersal limitation could explain some variation in VT distribution, VT relative abundance was almost exclusively driven by environmental variables. Several environmental and spatial effects on VT distribution and relative abundance were correlated with phylogeny, indicating that closely related VT exhibit similar niche optima and widths. Major clades within the Glomeraceae exhibited distinct niche optima, Acaulosporaceae generally had niche optima in low pH and low temperature conditions, and Gigasporaceae generally had niche optima in high precipitation conditions.
• Identification of the realised niche space occupied by individual and phylogenetic groups of soil microbial taxa provides a basis for building detailed hypotheses about how soil communities respond to gradients and manipulation in ecosystems worldwide.

Introduction

The concept of the ecological niche provides a framework for understanding resource partitioning by organisms and emergent patterns of coexistence and distribution (MacArthur & Levins, 1967; Tilman, 1982). Realised niches define the conditions under which organisms can survive and reproduce in the presence of biotic interactions. Niche optima (the most favourable environmental conditions for an organism), and widths (the range of environmental conditions in which the organism can persist) describe different aspects of the way organisms interact with the environment (Ozinga *et al.*, 2013; Gerz *et al.*, 2018). Estimates of such niche parameters are available for some regional floras (Ellenberg *et al.*, 1991; Ozinga *et al.*, 2013). However, we are unaware of analogous attempts to quantify the realised niche characteristics of low-level taxa (e.g. species or molecular taxa of similar resolution) throughout the diversity of any microbial organism group. For estimated niche parameters to have applicability beyond individual studies requires a consistent taxonomic framework (Llado *et al.*, 2018), which is a challenge in microbial ecology.

Arbuscular mycorrhizal (AM) fungi (phylum Glomeromycota; Tedersoo *et al.*, 2018) represent a key group of microbial organisms. They form a widespread symbiosis with plants (Smith & Read, 2010) that shapes the composition of plant communities (Klironomos *et al.*, 2011) and the functioning of ecosystems (Wurzburger *et al.*, 2017). They also represent a microbial group for which a widely used taxonomic classification is available in the form of the small subunit rRNA gene-based virtual taxonomy of the Maarj*AM* database (Öpik *et al.*, 2010). The units of this taxonomy, virtual taxa or VT, represent well supported monophyletic clades where within-clade sequence similarity exceeds a threshold of 97%. The resolution of VT corresponds approximately to that of AM fungal morphospecies (i.e. species described on the basis of morphological characteristics; Öpik & Davison, 2016), while the estimated divergence times of sister VT are on average intermediate (18 million years ago (Ma)) between those of plant species (16 Ma) and genera (19 Ma) (Davison *et al.*, 2018; although see Bruns & Taylor, 2016, Öpik *et al.*, 2016).

The composition of AM fungal communities is driven by climatic and edaphic factors at a range of scales from local to global (Dumbrell *et al.*, 2010; Kivlin *et al.*, 2011; Lekberg *et al.*, 2011; Hazard *et al.*, 2013; Davison *et al.*, 2015; Větrovský *et al.*, 2019). Such community-level responses to environmental conditions, with the same taxon combinations re-occurring in analogous environments, appear symptomatic of niche-related processes (Leibold & McPeek, 2006; Kraft *et al.*, 2008). There have been earlier attempts to model the distributions of two AM fungal families (Veresoglou *et al.*, 2013) and to calculate a nitrogen-association index for AM fungal genera (Treseder *et al.*, 2018). Yet, besides a study modelling the distribution of one abundant and widespread species (*Rhizophagus irregularis*; Kivlin *et al.*, 2017), there is little detailed information about organism–environment relationships among approximately species-level AM fungal taxa.

Many AM fungal VT are geographically widespread and have been shown to occur in multiple habitat types (Davison *et al.*, 2015; Savary *et al.*, 2018; Kivlin, 2020). However, these observations are coarse summaries of occurrence data and it is unknown to what extent individual VT vary in abundance along potential niche axes. AM fungal taxa have also been classified into different ecotypes (Alzarhani *et al.*, 2019) and into generalists and specialists, based on geographic (Moora *et al.*, 2011; Bouffaud *et al.*, 2016), habitat (Sýkorová *et al.*, 2007; Oehl *et al.*, 2010; Vályi *et al.*, 2015) or host plant species (Helgason *et al.*, 2007) ranges, but these niche descriptions only apply to limited ranges of conditions covered by single studies.

Certain functional attributes are similar among related AM fungal morphospecies (Powell *et al.*, 2009; Hoeksema *et al.*, 2018). Patterns of phylogenetic conservatism in niche attributes may inform about selective pressures, while considering organism–environment relationships at different taxonomic or phylogenetic resolutions may minimise potential biases connected with molecular marker selection or the definition of taxonomic units. However, taxonomically described AM fungal morphospecies represent a small fraction of the AM fungal molecular diversity recorded in natural environments (Öpik *et al.*, 2014), and

*These authors contributed equally to this work.

attempts to describe the functional characteristics of molecular taxa have so far been limited to a very small number of variables. Several authors have suggested that VT comprising sequence data collected from spore-producing morphospecies, described as cultured VT, represent the so-called ruderal life history strategy (van der Heijden *et al.*, 2008; Ohsowski *et al.*, 2014; García de León *et al.*, 2016, 2018), while the spore diameters of component morphospecies may be correlated with dispersal ability of VT (Davison *et al.*, 2018; Chaudhary *et al.*, 2020). In contrast to this fragmentary information, realised niche data covering the full molecular diversity of AM fungi could provide ecologists with a set of tools with which to study mycorrhizal fungi and mycorrhizal interactions in far greater detail.

Here, we address the realised niches of AM fungal VT and higher-order phylogenetic clades along major climatic and edaphic axes using a set of > 300 soil samples collected worldwide. Our approach is based on the theoretical expectation that the response of organisms to environmental gradients reflects the realised ecological niche *sensu* Hutchinson (Araújo & Guisan, 2006; Zimmermann *et al.*, 2010; Wasof *et al.*, 2015). We model the distribution and relative abundance of AM fungal VT in relation to a range of abiotic environmental and spatial variables (species distribution models; SDMs), and describe the width and volume of AM fungal niches with respect to these environmental characteristics. On the basis of widespread distribution patterns among AM fungal VT (Rosendahl *et al.*, 2009; Davison *et al.*, 2015; Savary *et al.*, 2018; Kivlin, 2020), we expect weak associations with climatic variables and little large-scale spatial structure. By contrast, evidence from community-level studies (Lekberg *et al.*, 2011; Hazard *et al.*, 2013) suggests that there should be clear niche optima along soil gradients. Correspondingly, we expect narrower niche widths along axes reflecting soil variables than in relation to climatic gradients. Earlier case studies demonstrated that among the few available trait proxies for VT, cultured status may predict disturbance tolerance in VT (Ohsowski *et al.*, 2014), while spore diameter may predict dispersal ability (Chaudhary *et al.*, 2020). Here we assess how those characteristics correlate with niche optima and widths. Evidence of phylogenetic conservatism in functional attributes (e.g. Powell *et al.*, 2009) leads us to expect higher-order phylogenetic correlation in niche properties.

Materials and Methods

Sample collection

This study made use of 327 soil samples from locations worldwide, including 84 previously published samples (Tedersoo *et al.*, 2014; Gazol *et al.*, 2016; García de León *et al.*, 2018) and 243 newly collected samples (including six for which AM fungal DNA in plant roots from the same sites has been published previously; Davison *et al.*, 2015, 2018) (Supporting Information Fig. S1; Table S1). Sampling was generally conducted at sites with little disturbance from human activities, including forest ($n = 236$), grassland ($n = 35$), scrub ($n = 41$) and semidesert ($n = 15$) ecosystems (Table S1). Soil sampling largely followed the approach described in Tedersoo *et al.* (2014). Briefly, *c.* 300 g of topsoil was collected from up to 40 points within about a 50 × 50 m sampling area and then pooled. The soil samples were dried within 24 h using silica gel or at room temperature and then carefully homogenised. Preservation with silica gel appears to maintain high concentrations of good quality AM fungal DNA in environmental samples (Bainard *et al.*, 2010), although there is some evidence that it may yield relatively low estimates of fungal richness (U'Ren *et al.*, 2014). Subsamples (2 g) of soil were extracted for further molecular analysis; the remainder was stored for geochemical analysis (for variation in the protocol see Table S1).

Soil chemical analyses and climatic variables

Sieved soil samples (2 mm) were used for analysis of soil chemical properties: pH, total N, organic C, Na, P, Mg and K. Soil pH was measured in 1 M KCl solution following ISO 10390:2005, using a Seven Easy pH meter with an InLab Expert Pro electrode (Mettler Toledo, Malaysia). The content of total N in soil was determined using the Kjeldahl method with a DK-20 digestion block and a UDK-126 distillation unit (Velp Scientifica Srl, Italy). To determine the organic carbon content of soils Tjurin's method was used, with oxidation provided by boiling samples in H_2SO_4 and $K_2Cr_2O_7$ solution. For determination of P, K, Mg and Ca, the Mehlich III extraction method was used, with the content of elements determined using an MP-4200 microwave plasma atomic emission spectrometer (Agilent, USA). Chemical analyses were performed at the Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Tartu, Estonia. Estimates of mean annual temperature (MAT) and mean annual precipitation (MAP) at sample locations were taken from the CHELSA database (Karger *et al.*, 2017). Abiotic environmental measures are shown in Table S1.

Molecular methods and bioinformatics

DNA was extracted from 2 g of dried soil using the PowerMax® Soil DNA Isolation Kit. AM fungal DNA was 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) that were both equipped with unique 12-base Golay barcodes for multiplexing. The PCR mixture contained 5 µl of 5XHOT FirePol Blend Master Mix (Solis Biodyne, Tartu, Estonia); 0.5 µl of each 20 µM primer; 1 µl of template DNA; and nuclease-free water to reach a total reaction volume of 25 µl. The PCR was performed in two replicates per sample under the following thermocycling conditions: 95°C for 15 min, 30 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s followed by 72°C for 5 min. PCR products from replicate samples were pooled and the amplification success was checked on 1% agarose gel. Samples with no visible gel band were re-amplified with 35 or 40 cycles. PCR products were pooled into seven libraries at approximate ratios as determined by the gel band strength. Libraries were purified using the FavorPrep Gel/PCR Purification kit (Favorgen Biotech Corp.,

Vienna, Austria), following the manufacturer's instructions. Both negative (distilled water) and positive (synthetic double-stranded DNA with relevant priming sites) controls were included in PCR and sequencing runs (one run did not contain a positive control; see Notes S1). Each library was ligated with Illumina adaptors using the TruSeq DNA PCR-free library prep kit (Illumina Inc., San Diego, CA, USA). Libraries were sequenced on the Illumina MiSeq platform, using a 2×300 bp paired-read sequencing approach, at Asper Biogene (Tartu, Estonia).

Demultiplexed paired-end reads were analysed following the bioinformatics steps described by Vasar *et al.* (2017). Primer sequences were matched allowing one mismatch for both pairs, and primers were removed from the paired-end sequences. After removal of barcode and primer sequences, only pairs where both reads had an average quality score of > 30 were retained. Quality filtered paired-end reads were combined using FLASH (v.1.2.10, Magoč & Salzberg, 2011) with the default parameters (10–300 bp overlap with at least 75% identity). Orphan reads (paired-end reads, where one pair had low average quality) and unpaired reads (paired-end reads that did not meet the conditions for combination) were removed from the analyses. The VSEARCH (v.2.14.1, Rognes *et al.*, 2016) reference-based chimera filtering algorithm was used to remove putative chimeric reads using the default parameters with the MaarjAM database as a reference set. Retained reads were subjected to a BLAST + search (v 2.8.1, Camacho *et al.*, 2009) against the MaarjAM database using 97% identity and 95% alignment length thresholds, revealing 327 VT in 327 samples. Samples were retained if they contained > 100 reads, while global singleton VT were omitted, leaving 278 VT in 286 samples. The full data set was then normalised to the median read count (3385) using resampling with replacement (Veresoglou *et al.*, 2019), leaving a final data matrix with 268 VT and 286 samples. This normalising approach has the benefit of correcting the number of taxa recorded in samples to account for differential sampling depth while maintaining the abundance relationships between remaining taxa. Raw reads from this Targeted Locus Study have been deposited in the NCBI SRA (BioProject PRJNA659159), and representative sequences of each VT were deposited in the NCBI GenBank under the accession no. KELL00000000 (the version described in this paper is the first version, KELL01000000).

Analysis of AM fungal VT niches

Species distribution modelling To investigate VT environmental niche axes while simultaneously identifying and accounting for spatial correlates of VT occurrence and relative abundance, we used a subset of sites for which there were complete environmental data (266 out of 286 sites) and VT that were present at ≥ 20 of these sites (137 VT out of 268 VT in the final data matrix). The threshold of 20 was chosen as a minimum to allow construction of models containing multiple predictor variables (see below). We used a two-step modelling approach: (1) first modelling the occurrence (presence/absence) of VT at sites using generalised linear models (binomial error structure); then (2) modelling the logit-transformed relative abundance (the proportion of reads per sample belonging to each VT) of VT at sites where they were present using Gaussian linear models.

For each modelling approach and for each VT, a spatial variable set was constructed by first using Moran's Eigenvector Mapping (MEM) to define spatial vectors from site coordinates, with spatial weighting according to a range of neighbourhood connectivity measures: sites within {10, 50, 100, 200, 500} km; and the nearest {1, 5, 10, 20, 50} sites (Dray *et al.*, 2006). For each connectivity measure an optimal set of MEM vectors was identified using stepwise selection based on AIC in a model containing VT occurrence or relative abundance as the response variable (Blanchet *et al.*, 2008). Then, the best fitting of the optimised MEM sets was selected on the basis of AIC (Table S2).

In parallel, sets of abiotic environmental predictors were defined among the measured abiotic variables: pH, P, N, K, Mg, Ca, organic C, MAT, MAP. These variables include abiotic environmental gradients known to correlate with AM fungal diversity and community composition (Davison *et al.*, 2015). For each environmental variable we used AIC to determine whether VT occurrence or relative abundance was better explained by the linear or quadratic form of the variable. Then, analogous to the procedure for selecting spatial MEMs, we applied principal components analysis followed by AIC-based stepwise selection to the full environmental variable matrix (including quadratic forms when selected) to define for each VT a set of well supported, orthogonal environmental vectors.

Finally, the occurrence or relative abundance of each VT was modelled, using linear or GLM models, in relation to: (1) the spatial MEM set, (2) the environmental variable set; (3) the combined MEM and environmental variable sets; and (4) each environmental variable in isolation (including quadratic forms when selected). The explanatory power of models was estimated using explained deviance (adjusted D^2; analogous to the coefficient of determination (R^2) used for linear models; Guisan & Zimmermann, 2000) for GLM models or adjusted R^2 for linear models. Model comparison was used to partition D^2 and R^2 between spatial and environmental sets. Model fits were assessed by visual inspection of representative residual plots. GLM model fit was assessed from Pearson residuals and Q–Q plots based on quantile residuals (Augustin *et al.*, 2012; Figs S2, S3). Leave-one-out cross-validation was used to assess model accuracy (Hastie *et al.*, 2009). Using observed and cross-validation predictions, the area under the receiver operating characteristic (ROC) curve was calculated for GLM models; and the root squared mean error (RSME) and median absolute error (MedAE) were calculated for linear models; Table S3).

Niche widths and volumes of AM fungal VT We defined the widths of VT niches in relation to individual environmental variables for those VT occurring at ≥ 5 sites ($n = 230$ out of 268 VT in the final data matrix). The threshold of five was taken as the minimum necessary for estimation of a niche width parameter (standard deviation), and assessed using a sensitivity analysis (Fig. S4). Each environmental variable (the full data set) was first scaled by one standard deviation. Then for each VT we estimated niche width as the standard deviation of the variable, weighted by

the relative abundance of the VT at the respective sites (the weighting modulates the contribution of individual observations to the calculation of the standard deviation). Total abiotic niche volume for each VT was then calculated as the product of the width estimates for different environmental variables. We used a null model approach to assess whether standardised niche widths and volumes for VT were larger or smaller than expected if occurrence of VT at sites were random. We used the 'quasiswap' algorithm as implemented in the R package VEGAN (Oksanen *et al.*, 2019) to generate randomised matrices. This algorithm preserves matrix row and column sums and matrix fill (i.e. the proportion of nonzero cells). We then calculated standardised effect sizes (SES; (observed value − mean of randomised values)/SD of randomised values). SES values above zero indicate wider niches than expected; values below zero indicate narrower niches than expected. SES values with absolute magnitude > 1.96 were taken to indicate significantly narrower or wider ranges than expected.

For display purposes and to produce estimates of VT niche optima and widths that might serve as explanatory variables in future analyses, we also estimated the unstandardised niche optimum and width for each VT and abiotic variable as the mean and standard deviation, respectively, of the variable, weighted by the relative abundance of the VT at the respective sites.

Correlation of VT niche characteristics with phylogeny and traits Parameters describing AM fungal niches resulting from the previous analyses were placed in the context of phylogenetic relationships between VT and correlated with previously described VT characteristics (cultured status, spore diameter).

Following the approach used by Davison *et al.* (2015) we constructed a phylogenetic tree of the aligned VT type sequences of all AM fungal VT (Maarj*AM* status January 2020) using BEAST (v.1.8.0; Drummond *et al.*, 2012). Phylogenetic analysis was conducted using the GTR + I + G nucleotide substitution model and a log-normal relaxed clock model with a coalescent tree model. Three separate chains of 30 000 000 iterations were constructed and combined after removal of burn-in. A Gaussian prior (mean = 505 Ma, standard deviation = 54 My) was used for the age of the root (see Davison *et al.*, 2015 for details of this tree calibration). The results were summarised on a maximum clade credibility tree (see Fig. S5).

Phylogenetic correlograms (Gittleman & Kot, 1990) were constructed to identify correlation in niche parameters in relation to the phylogenetic relatedness of VT. VT relatedness is estimated from the phylogenetic distance (tree branch length) separating pairs of VT, or equivalently, the age of the most recent ancestral node in the tree. The correlogram approach then estimates correlation at different 'lags' of relatedness (e.g. common ancestors 0–5 Ma, 5–10 Ma, 10–15 Ma etc.). A confidence envelope around the null expectation (correlation = 0) is then constructed using bootstrapping ($n = 500$), and where observed correlation exceeds the envelope, correlation was considered significant. Correlation was measured using Moran's I (for individual parameters) or Mantel r (where multiple related parameters were simultaneously investigated). As final species distribution models for different VT contained different spatial and environmental predictor variables, individual coefficients could not be uniformly compared with assess phylogenetic correlation in effects. Instead, we assessed multivariate correlation in the fitted values (for each VT an estimate of probability of occurrence (GLM) or abundance (linear models) for each site) returned by models containing only the variables of interest (i.e. only environmental variables or only spatial variables). The fitted values for each VT were first divided by the VT sum to provide a relative estimate of probability of occurrence or abundance at different sites, independent from overall VT abundance.

Correlations with cultured status and mean spore diameter for certain cultured VT (using data from Davison *et al.*, 2018) were examined for AM fungal niche parameters measured in this analysis. Mantel tests were used for multivariate correlations; Pearson's correlation was used to correlate single variables (in both cases treating cultured status as a binary (0,1) variable).

Results

Sample characteristics

Analysis of soil samples ($n = 327$) collected worldwide yielded AM fungal sequence data of sufficient quality from 284 samples (Table S1; Fig. S6). The measured abiotic environment at sample locations broadly reduced to two axes: high temperature/low C/low N on the one hand; high precipitation/low pH/low P/low K and several other soil nutrients on the other (Fig. 1). The samples yielded 268 AM fungal VT belonging to nine families (Acaulosporaceae = 22, Ambisporaceae = 3, Archaeosporaceae = 9, Claroideoglomeraceae = 10, Diversisporaceae = 15, Gigasporaceae = 7, Glomeraceae = 189, Pacisporaceae = 1, Paraglomeraceae = 11; Table S4). Taxa included in the distribution modelling (i.e. those that were present at ≥ 20 sites; $n = 137$ VT) belonged to eight families (Pacisporaceae missing), while all sampled families were included in estimates of niche dimensions (present at ≥ 5 sites; $n = 230$ VT; Table S4).

Distribution modelling

More variation in site-level occurrence was explained (i.e. D^2) by spatial compared with environmental variables for 64% of VT (88 out of 137 VT); by contrast in models of site-level relative abundance, the variance explained (i.e. R^2) by environmental variables was higher for 97% of VT (133 out of 137 VT) (Table S3). There was marginally significant positive phylogenetic correlation in the relative importance of spatial vs environmental variables in determining VT occurrence (the confidence interval around the point correlation estimate excluded zero at some phylogenetic scales; Fig. 2); but not in determining relative abundance (Fig. S7). Therefore, the distributions of closely and moderately related VT (diverging up to around 200 Ma) were similarly influenced by different limiting factors, in the forms of the abiotic environment and spatial constraints.

In models of VT occurrence, the selected neighbourhood connectivity measure used to generate spatial MEMs comprised a small number of close sites (< 10 nearest neighbours or < 10 km)

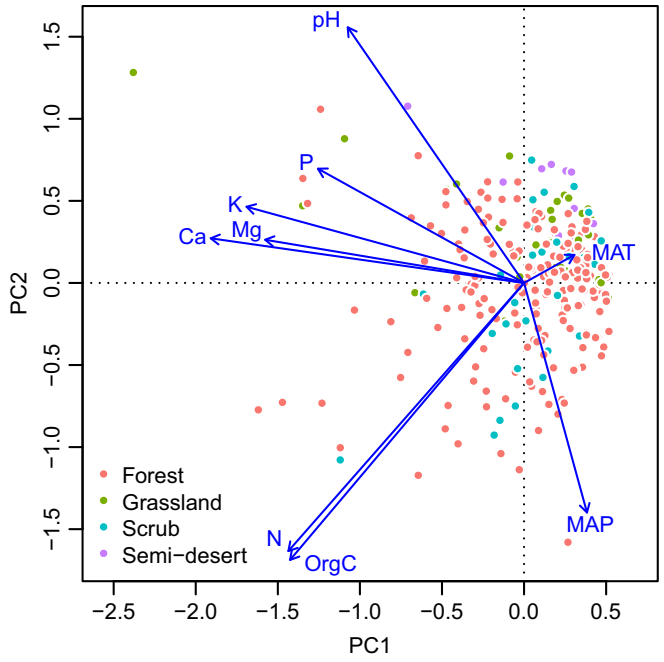

Fig. 1 Environmental characteristics at sample locations. PCA biplot of environmental variables (PC1 and PC2 respectively explain 33% and 22% of variance in the original edaphic and climatic variables). MAP, mean annual precipitation; MAT, mean annual temperature.

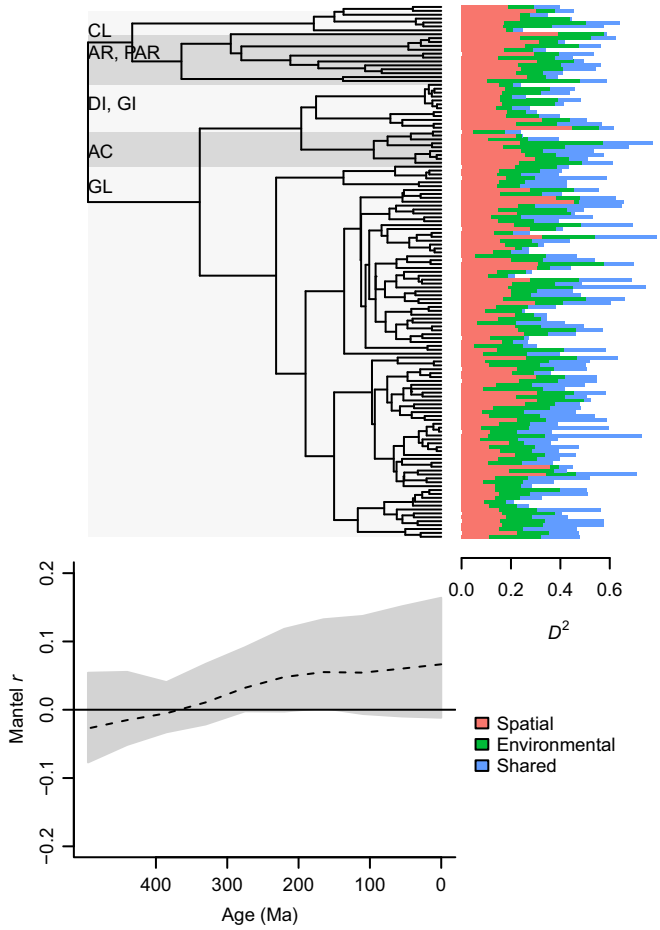

Fig. 2 Phylogenetic correlation in the explanatory power of spatial and environmental variables in generalised linear models of virtual taxon (VT) presence–absence in soils worldwide. The barplot at the tips of the small subunit nuclear ribosomal DNA phylogenetic tree shows the deviance explained (D^2; a measure of explained variation in the dependent variable, analogous to the coefficient of determination (R^2) used for linear models) by spatial and environmental variable sets, and their intersection, for each VT (each tip of the tree). A phylogenetic correlogram (below) shows how the correlation in the estimates of D^2 changes as a function of VT relatedness (i.e. evolutionary age; My). The dashed line shows the Mantel r estimate, while the grey envelope shows the 95% confidence interval derived from bootstrapping. AC, Acaulosporaceae; AR, Archaeosporaceae; CL, Claroideoglomeraceae; DI, Diversisporaceae; GI, Gigasporaceae; GL, Glomeraceae; PAR, Paraglomeraceae. Ma, million years ago.

for 75% of VT (104 out of 137 VT) and as such reflected local-scale rather than large-scale spatial correlation (Table S2); in models of VT relative abundance the corresponding figure was somewhat lower at 52% (71 out of 137 VT).

MAT and soil pH were the most influential environmental variables explaining VT occurrence (MAT highest D^2 for 47% (64 out of 137 VT); pH highest D^2 for 36% (49 out of 137 VT)) (Fig. 3; Table S3). In models of relative abundance, these two variables remained the most important, but variable importance was more evenly spread (Fig. S8; Table S3). There were some correlations between VT cultured status and spore diameter on the one hand and the variation in VT relative abundance explained by different variables on the other hand, such that MAP explained more variation in relative abundance among cultured VT, while pH and Ca explained more variation in relative abundance among large-spored VT (Table S5).

There was significant positive phylogenetic correlation in the influence (D^2) of P, pH and MAP on VT occurrence (Table S5). In all cases correlation reflected within-family clades exhibiting similar variable explanatory power. In general, Acaulosporaceae was characterised by niche optima in low pH conditions (the median niche optimum was 4), while a clade of Glomeraceae containing no recorded morphologically described species was characterised by the highest pH niche optima (median niche optimum of 6; clade 4 in Figs S5, 4; estimated niche optima for each VT are shown in Table S6). Acaulosporaceae niche optima were also characterised by low temperature (median niche optimum just below 10°C), while Gigasporaceae was characterised by high temperature niche optima (median niche optimum *c.* 21°C) and there was considerable variation within Glomeraceae: in a clade containing *Glomus hoi*, *G. indicum*, *G. macrocarpum* and *G. perpusillum* the median temperature niche optimum was also *c.* 21°C (clade 2 in Fig. S5), while in the clade containing no morphospecies (clade 4 in Fig. S5), the median niche optimum was around 12°C (Fig. 4). Gigasporaceae were also marked by notably high niche optima in relation to precipitation (the median value was over 2000 mm). There was significant positive phylogenetic correlation in the fitted values produced by environmental and spatial models of VT occurrence but not those of relative abundance (Fig. S9; Table S5). Positive correlation indicated generally similar responses to environmental and spatial variables among taxa that diverged up to *c.* 150 Myr BP (approximately the family-level or among within-family clades). The

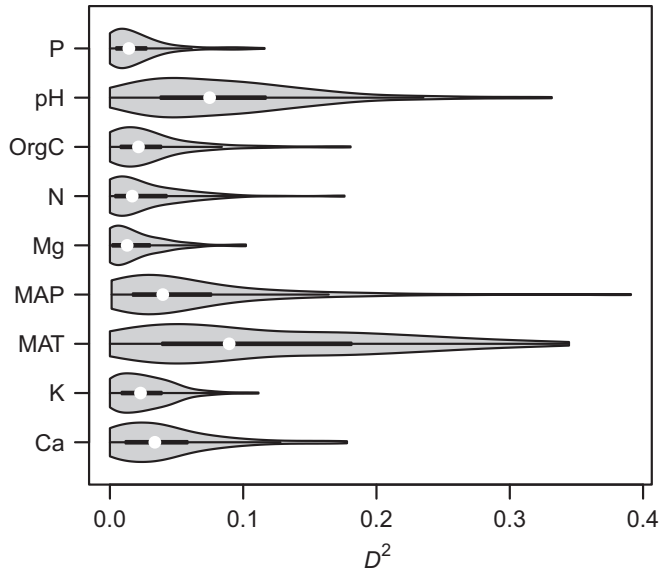

Fig. 3 Explanatory power of different environmental variables in generalised linear models of virtual taxon (VT) presence–absence in soils worldwide. Explained deviance (D^2) is analogous to the coefficient of determination (R^2) used for linear models and provides a measure of how much variation in the dependent variable was explained by the model. The curves show the probability density of the data at different values of D^2. White dots indicate median values; thick black lines indicate interquartile ranges. MAP, mean annual precipitation; MAT, mean annual temperature.

fitted values from environmental and spatial models of VT relative abundance were significantly correlated with estimated VT spore diameter (Table S5), such that VT of similar spore size responded similarly to environmental conditions and spatial constraints.

AM fungal niche width

VT niches were often narrower than expected based on comparison with a null model, but the fractions of VT exhibiting narrow or significantly narrow niches varied between measures of VT niche width and volume (Fig. 5). pH and MAT niches systematically exhibited the lowest SES values, with *c.* 50% of taxa exhibiting significantly narrow niches, indicating that VT often associated with a particularly narrow range of pH and temperature conditions. Furthermore, P niche width was positively correlated with VT spore diameter (Table S5). Estimated niche volume exhibited phylogenetic correlation, such that some within-family clades had similar niche volumes (clades 3 and 4 within Glomeraceae (cf. Fig. S5) appeared to exhibit systematically large niche volume; Fig. 6), as did the width of niches along the P, K and Mg axes (Table S5; Fig. S10; estimated niche widths for each VT are shown in Table S6).

Discussion

Global sampling detected around 70% of recorded virtual taxa and incorporated all AM fungal families recorded in the MaarjAM database except Geosiphonaceae and Pervetustaceae; as such the study provides a first detailed characterisation of realised niches throughout the known diversity of AM fungi. We found that the characteristics of AM fungal VT niches varied between niche axes. In general, the narrowest niches and most pronounced optima were apparent with respect to temperature and soil pH. There was also significant phylogenetic clustering in relation to spatial constraints and multiple niche properties, including niche optima and dimensions. The results provide insight into the processes shaping AM fungal diversity and a resource, in the form of VT-level niche parameters, to underpin future analysis of AM fungal communities.

Spatial vs environmental constraints on AM fungal occurrence and abundance

It has been suggested that the distribution of AM fungi bears the footprint of dispersal over geological time (Morton *et al.*, 1995; see also Fitter, 2005). However, our findings are largely consistent with recent large-scale studies indicating near-global distributions among many morphospecies (Stürmer *et al.*, 2018), VT (Davison *et al.*, 2015) and even genotypes within an AM fungal morphospecies (Savary *et al.*, 2018). The best fitting spatial variable sets in VT distribution models predominantly reflected spatial relationships between small numbers of neighbouring sites, rather than regional-scale or larger-scale spatial patterns. As such, they may have reflected small-scale dispersal limitation (Grünfeld *et al.*, 2020) or, perhaps, local environmental differences that were not captured in the environmental variable set (Rasmussen *et al.*, 2018). It was also apparent that while spatial variables explained some variation in VT occurrence, VT relative abundance at sites showed little spatial pattern, however it was driven by environmental variables. Previous work has shown dispersal limitation of the occurrence of AM fungal taxa in landscapes (Lekberg *et al.*, 2011; Bouffaud *et al.*, 2016; Torrez *et al.*, 2016; Grünfeld *et al.*, 2020), but our results clearly show the role of the environment in shaping local relative abundance.

We expected clearer niche optima in relation to local soil variables than climatic variables, where gradients occur at larger spatial scales. However, the best predictors of VT occurrence and relative abundance in species distribution models were one climatic and one soil chemical variable: temperature and pH. An earlier attempt to model the distribution of *Rhizophagus irregularis* (member of Glomeraceae clade 3 in this study) found that climatic variables were important predictors (Kivlin *et al.*, 2017), and the importance of temperature has been demonstrated by warming experiments that elicited changes in the composition of root colonising AM fungal communities (Cotton, 2018; Rasmussen *et al.*, 2018). Indeed, previous authors have noted that temperature strongly regulates the physiological development of certain AM fungi (Tibbett & Cairney, 2007), with variation in temperature responses apparent in comparisons between small numbers of taxa (Helgason & Fitter, 2009). Moreover, Antunes *et al.* (2011) reported different symbiotic function in isolates of the same AM fungal taxon originating from contrasting climates. Here, we demonstrate the general importance of temperature in defining the niches of c. species-level AM

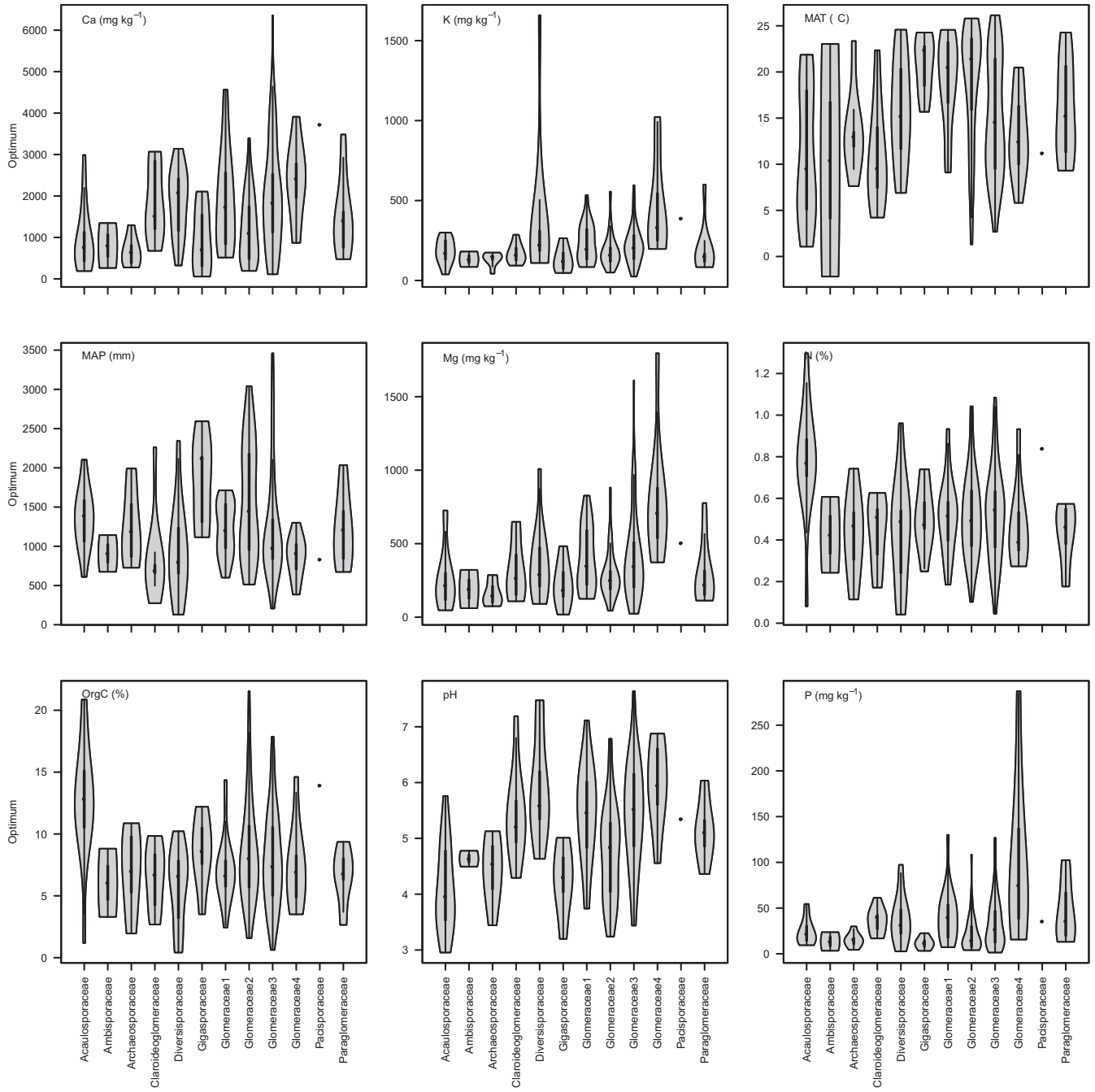

Fig. 4 Niche optimum estimates along different environmental axes for virtual taxa (VT) in different families and clades (see Supporting Information Fig. S2). Niche optima were calculated as the mean environmental conditions at sites where VT occurred, weighted by VT relative abundance. Points indicate median values; thick black lines indicate interquartile ranges. MAP, mean annual precipitation; MAT, mean annual temperature.

fungal taxa. Several case studies have also shown that soil pH importantly influences the community composition of AM fungi (Dumbrell *et al.*, 2010; Lekberg *et al.*, 2011; Hazard *et al.*, 2013; Davison *et al.*, 2015; van Geel *et al.*, 2018), and of fungal communities in general (Glassman *et al.*, 2017). Experiments showing that liming can strongly modulate AM fungal spore number and root colonisation suggest that pH may have an important direct influence on AM fungal growth and performance (Siqueira *et al.*, 1984; Wang *et al.*, 1993; Coughlan *et al.*, 2000). Our results indicate an important influence of soil pH on individual taxa throughout the phylum. Given the limited evidence of dispersal limitation and larger spatial resolution of climatic variables, it seems probable that soil pH is the most important abiotic determinant of relative abundance at the local scale.

Our results demonstrating the important role of temperature and pH are consistent with those previously reported for bacteria and other fungi (Bahram *et al.*, 2018) and for protists (Oliverio *et al.*, 2020), although both studies identified precipitation as an

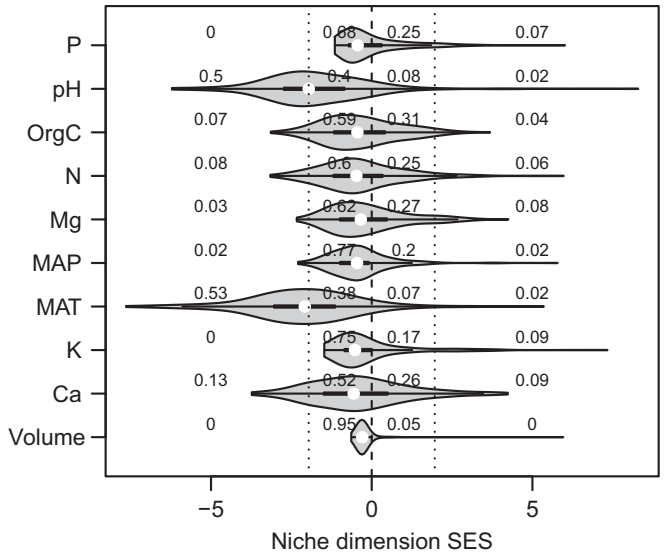

Fig. 5 Virtual taxon (VT) niche width estimates along different environmental axes. Niche widths are presented as standardised effect sizes (SES) in relation to a null model. Negative values indicate narrower niches than expected; positive values wider niches than expected. White dots indicate median values; thick black lines indicate interquartile ranges. SES values with an absolute value > 1.96 are taken to indicate significantly wider or narrower niches than expected. SES = 0 is marked with a long dash line; SES = −1.96 and SES = 1.96 are marked with short dashed lines. The proportion of VT falling into the categories: significantly narrower than expected (SES < −1.96), narrower than expected (SES = −1.96–0), wider than expected (SES = 0–1.96) and significantly wider than expected (> 1.96) are shown for each environmental axis. MAP, mean annual precipitation; MAT, mean annual temperature.

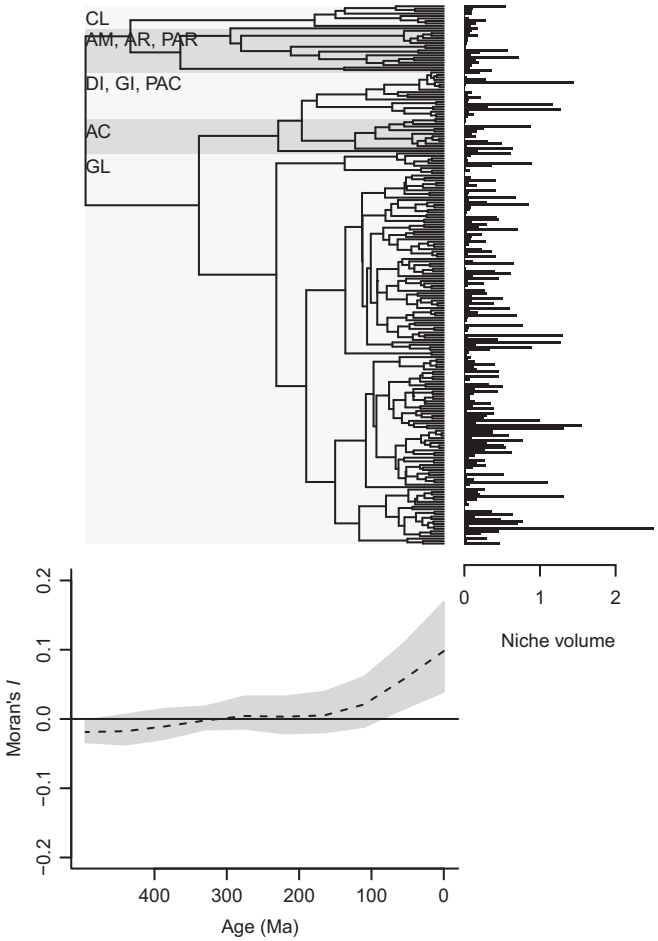

Fig. 6 Phylogenetic correlation in virtual taxon (VT) niche volume. The barplot at the tips of the small subunit nuclear ribosomal DNA phylogenetic tree shows the niche volume estimate (square root) for each VT (each tip of the tree). A phylogenetic correlogram (below) shows how correlation in the estimates of niche volume changes as a function of VT relatedness (i.e. evolutionary age; Ma). The dashed line shows the Moran's *I* estimate, while the grey envelope shows the 95% confidence interval derived from bootstrapping. AC, Acaulosporaceae; AM, Ambisporaceae; AR, Archaeosporaceae; CL, Claroideoglomeraceae; DI, Diversisporaceae; GI, Gigasporaceae; GL, Glomeraceae; PAC, Pacisporaceae; PAR, Paraglomeraceae.

important additional driver. At the same time, it is clear that there remain unmeasured soil parameters in this and former studies that may be important drivers of AM fungal performance or may underlie associations with other variables, such as those identified here (Větrovský *et al.*, 2019). For example, soil bulk density, a potential determinant of AM fungal distribution, is strongly correlated with temperature (Zhao *et al.* 2019). Aspects of the methodology used for extensive global sampling (including single-time soil analysis, pooling of soil samples within plots, the detection of dead or dormant organisms) may also have influenced the precision and generality of the findings. Furthermore, there is a theoretical possibility that some records of VT occurrence reflect only nonviable propagules arriving at a location. We anticipate that such records are generally uncommon, with bias potentially highest where such records arise in otherwise sparsely recorded taxa, however the least frequently recorded taxa were not included in niche analyses.

AM fungal niche dimensions

There is virtually no previous information about the niche dimensions of AM fungi, in addition to a handful of attempts to classify AM fungal taxa into categories of generalists and specialists based on the observed geographic (Moora *et al.*, 2011; Bouffaud *et al.*, 2016) or habitat (Sýkorová *et al.*, 2007; Oehl *et al.*, 2010; Vályi *et al.*, 2015) range used by taxa within single studies. Our results confirmed that, on average, AM fungal VT are more specialist than expected along all measured axes, on the basis of null model analysis. Estimated niche widths were narrowest, in relation to the null model, for the two most important gradients emerging from the distribution modelling: temperature and pH.

Phylogenetically conserved AM fungal niche properties

Certain functional attributes, as well as the extent of geographic distribution, of AM fungal taxa are similar among related taxa within broad phylogenetic groupings (Powell *et al.*, 2009; Bouffaud *et al.*, 2016). Similarly, a recent analysis showed that different ecogroups of fungi, defined on the basis of their responses to abiotic and biotic environmental gradients, predominantly comprise different functional groups (e.g. symbiotic, pathogenic;

Alzarhani *et al.*, 2019). We identified phylogenetic clustering in AM fungal niche properties at the level of higher taxonomic ranks, such as the approximate family level. Such niche conservatism is encountered in the functional traits of many organism groups, and is relevant to a range of ecological questions, including community assembly, responses to environmental change and species richness patterns (Wiens *et al.*, 2010). For example, among AM fungi, niche conservatism is likely to drive observed patterns of phylogenetic clustering among local communities (Davison *et al.*, 2016) and relationships between species pool size and historical habitat availability (Pärtel *et al.*, 2017).

We found that VT within some approximately family-level or within-family clades had similar niche volumes and responses to spatial and environmental constraints. In the case of spatial constraints, this may indicate conserved responses to unmeasured environmental gradients or conserved dispersal traits, such that a given landscape configuration imposes a similar degree of dispersal limitation on closely related taxa. Some but not all clades within Glomeraceae appeared to exhibit systematically large niche volume. Therefore, despite being characterised as possessing low capacity for dispersal (Hart & Reader, 2002), and allocating the majority of biomass into hyphae growing inside plant roots (Maherali & Klironomos, 2007), this large family comprises multiple ecological strategies. Specifically, in relation to the most important niche axes, pH and temperature, it was also clear that major clades within Glomeraceae exhibited varying niche properties. For example, clade 3 (containing widespread species *Glomus irregularis*) had a median temperature optimum *c.* 5 °C lower than that of clades 1 and 2 (which include *Glomus mosseae* and *Glomus hoi* among other species); while clade 4 (DNA-based only) had a median temperature optimum nearly 10°C lower, and a pH value *c.* one unit higher, than clade 2 (it also had markedly higher P and Ca optima). It was also notable that Acaulosporaceae generally had niche optima in low pH and low temperature conditions, while Gigasporaceae generally had niche optima in high precipitation conditions. These latter patterns closely mirror the findings of Veresoglou *et al.* (2013), who identified the same environmental drivers for the occurrence of these families in database records. Finally, it should be acknowledged that analysis of relationships at the VT level and above may neglect meaningful within-VT or genotype-level niche characteristics. While local adaptation in mycorrhiza is an established phenomenon (Johnson *et al.*, 2010), the challenge of experimentally manipulating individual AM fungal taxa from natural environments (most taxa are not present in culture collections; Ohsowski *et al.*, 2014) means that little information is known about adaptation among the fungal partners.

Realised niches in the context of functional characteristics and the biotic environment

We expected niche optima and widths to vary between cultured VT, presumed to represent the so-called ruderal life history strategy (van der Heijden *et al.*, 2008; Ohsowski *et al.*, 2014) and uncultured VT. Indeed, temperature was a relatively more important driver of the distribution of cultured VT compared with uncultured VT. Also, within the group of cultured VT, large spore size was more characteristic of VT preferring high pH and Ca soils. Generally, little information is known about the ecological significance of AM fungal spore size (Aguilar-Trigueros *et al.*, 2019), although aerially distributing spores tend be smaller than average among AM fungi (Chaudhary *et al.*, 2020). However, in general, differences between cultured and uncultured taxa, and in relation to spore size, were limited, suggesting that these characteristics may capture different aspects of the organism–environment relationship than the abiotic niche axes described here.

We studied realised niches that reflect the distribution of organisms relative to abiotic gradients in situations where they are exposed to various biotic interactions (Araújo & Guisan, 2006; Wasof *et al.*, 2015). At the same time, we did not specifically incorporate any of the biotic drivers that potentially influence AM fungal distribution and abundance. Most obviously, AM fungi are obligately symbiotic organisms, obtaining carbon from a host plant, and distinct AM fungal assemblages associate with different plant functional groups (Davison *et al.*, 2020). Furthermore, AM fungi coexist with complex communities of interacting soil organisms. AM fungal communities are known to be sensitive to interactions with bacteria, other mycorrhizal fungi (Frey-Klett *et al.*, 2007; Tedersoo *et al.*, 2020) and other soil organisms, including nematodes (Rodríguez-Echeverría *et al.*, 2009) and earthworms (Paudel *et al.*, 2016), while AM fungal competition may also limit the occurrence of taxa (Engelmoer *et al.*, 2014; Knegt *et al.*, 2016). Therefore, observed environmental and spatial effects on AM fungal distribution and abundance are mediated by biotic interactions, such as the environmental preferences of a preferred host plant. As such, the relationships identified here serve as practical, empirical measures of AM fungal realised niches, but do not provide information about fundamental abiotic niches (e.g. how certain abiotic conditions directly constrain AM fungal distribution and abundance).

Conclusion

This analysis provides a detailed and comprehensive overview of realised niche attributes among a group of essential plant-symbiotic fungi. These generally widespread organisms have distinct requirements along realised niche axes reflecting large-scale and local-scale environmental conditions. There was clear phylogenetic signal in the variation of niche parameters, with closely related VT exhibiting similar abiotic niche optima and widths. The general importance of temperature and pH among niche axes of AM fungal VT may signal the sensitivity of these organisms to changing climate and soil management. We also provide quantitative measures of various niche parameters that may serve as explanatory variables for future investigations of AM fungal community ecology (Table S6).

Acknowledgements

This study was supported by the European Regional Development Fund (Centre of Excellence EcolChange), the University of Tartu (PLTOM20903), the Estonian Research Council

(MOBTP 105, MOBERC20, PRG352, PRG609, PRG 1065 and PRG 1170 and PRG 1409), Moscow State University (AAAA-A16-116021660039-1), a Natural Sciences and Engineering Research Council of Canada Discovery Grant, the Russian Science Foundation (19-14-00038), the São Paulo Research Foundation (2016/25197-0), the Swedish Research Council (Vetenskapsrådet, 2017-05019) and Qatar Petroleum (QUEX-CAS-QP-RD-18/19). We thank the Instituto da Biodiversidade e das Áreas Protegidas, for arranging permission to sample in Guinea Bissau, and Boyd Deep Canyon Reserve, California, USA (https://doi.org/10.21973/N3V66D). We also thank Ruben Heleno and François-Xavier Naramabuye for organising field sampling and identifying contacts; and Nina Farish, Teele Jairus, Rein Kalamees, Darja Koltysheva, Mart Meriste, Dagmar Mucina, Tanaka Muradzikwa, Yves Uwiragiye, Ülle Reier and Anna Zobel for assisting with fieldwork. We are grateful to Jaan Liira for his helpful comments that improved the manuscript.

Author contributions

JD, LT, MM, MÖ, MS and MZ designed the research; all authors contributed to collection or lab analysis of soil samples; JD and MV conducted the bioinformatics and statistical analysis; JD and MZ wrote the first draft of the manuscript, all authors contributed to the final version. JD, MM, MS, LT, MZ and MÖ contributed equally to this work.

ORCID

Talaat Ahmed https://orcid.org/0000-0001-8022-1855
Juha M. Alatalo https://orcid.org/0000-0001-5084-850X
Elena Andriyanova https://orcid.org/0000-0003-1441-8283
Sten Anslan https://orcid.org/0000-0002-2299-454X
Mohammad Bahram https://orcid.org/0000-0002-9539-3307
Amgaa Batbaatar https://orcid.org/0000-0002-6687-0274
Sakeenah Binte Adenan https://orcid.org/0000-0002-8931-7041
Charlotte Brown https://orcid.org/0000-0002-3989-6401
James Cahill https://orcid.org/0000-0002-4110-1516
Saida Chideh https://orcid.org/0000-0003-2942-3640
Ana P. Coelho https://orcid.org/0000-0002-3337-0310
Matthew Coghill https://orcid.org/0000-0003-1476-1567
John Davison https://orcid.org/0000-0002-0161-6195
Guillaume Decocq https://orcid.org/0000-0001-9262-5873
Sergey Dudov https://orcid.org/0000-0003-1512-0956
Vladimir E. Fedosov https://orcid.org/0000-0002-5331-6346
Lauchlan Fraser https://orcid.org/0000-0003-3998-5540
Sydney I. Glassman https://orcid.org/0000-0001-9115-3026
C. Guillermo Bueno https://orcid.org/0000-0002-7288-2271
Aveliina Helm https://orcid.org/0000-0003-2338-4564
Hugh A. L. Henry https://orcid.org/0000-0001-8397-6292
Bruno Hérault https://orcid.org/0000-0002-6950-7286
Inga Hiiesalu https://orcid.org/0000-0002-5457-2376
Wael N. Hozzein https://orcid.org/0000-0003-2467-9719
Petr Kohout https://orcid.org/0000-0002-3985-2310
Urmas Kõljalg https://orcid.org/0000-0002-5171-1668
Kadri Koorem https://orcid.org/0000-0002-3376-9372
Lauri Laanisto https://orcid.org/0000-0003-2215-7298
Ülo Mander https://orcid.org/0000-0003-2340-6989
Ladislav Mucina https://orcid.org/0000-0003-0317-8886
Jean-Pierre Munyampundu https://orcid.org/0000-0003-3002-8672
Lena Neuenkamp https://orcid.org/0000-0001-6108-5720
Ülo Niinemets https://orcid.org/0000-0002-3078-2192
Casper Nyamukondiwa https://orcid.org/0000-0002-0395-4980
Vladimir Onipchenko https://orcid.org/0000-0002-4930-3112
Maarja Öpik https://orcid.org/0000-0001-8025-7460
Meelis Pärtel https://orcid.org/0000-0002-5874-0138
Claudia Paz https://orcid.org/0000-0001-9754-4087
Clara P. Peña-Venegas https://orcid.org/0000-0001-6317-3767
Cherdchai Phosri https://orcid.org/0000-0003-3963-752X
Sergei Põlme https://orcid.org/0000-0002-9658-1166
Argo Ronk https://orcid.org/0000-0002-8716-976X
Alessandro Saitta https://orcid.org/0000-0002-5670-7780
Olivia Semboli https://orcid.org/0000-0003-2906-4609
Alexey Seregin https://orcid.org/0000-0002-1824-7453
Surya Sudheer https://orcid.org/0000-0002-4445-0500
Leho Tedersoo https://orcid.org/0000-0002-1635-1249
Tanel Vahter https://orcid.org/0000-0002-4343-5288
Martti Vasar https://orcid.org/0000-0002-4674-932X
Annelies J. Veraart https://orcid.org/0000-0001-6286-7484
Martin Zobel https://orcid.org/0000-0001-7957-6704

References

Aguilar-Trigueros CA, Hempel S, Powell JR, Cornwell WK, Rillig MC. 2019. Bridging reproductive and microbial ecology: a case study in arbuscular mycorrhizal fungi. *ISME Journal* **13**: 873–884.

Alzarhani AK, Clark DR, Underwood GJ, Ford H, Cotton TA, Dumbrell AJ. 2019. Are drivers of root-associated fungal community structure context specific? *ISME Journal* **13**: 1330–1344.

Antunes PM, Koch AM, Morton JB, Rillig MC, Klironomos JN. 2011. Evidence for functional divergence in arbuscular mycorrhizal fungi from contrasting climatic origins. *New Phytologist* **189**: 507–514.

Araújo MB, Guisan A. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography* **33**: 1677–1688.

Augustin NH, Sauleau EA, Wood SN. 2012. On quantile quantile plots for generalized linear models. *Computational Statistics and Data Analysis* **56**: 2404–2409.

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

Bainard LD, Klironomos JN, Hart MM. 2010. Differential effect of sample preservation methods on plant and arbuscular mycorrhizal fungal DNA. *Journal of Microbiological Methods* **82**: 124–130.

Blanchet FG, Legendre P, Borcard D. 2008. Forward selection of explanatory variables. *Ecology* **89**: 2623–2632.

Bouffaud ML, Creamer RE, Stone D, Plassart P, van Tuinen D, Lemanceau P, Wipf D, Redecker D. 2016. Indicator species and co-occurrence in communities of arbuscular mycorrhizal fungi at the European scale. *Soil Biology and Biochemistry* **103**: 464–470.

Bruns TD, Taylor JW. 2016. Comment on "Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism". *Science* **351**: 826.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Chaudhary VB, Nolimal S, Sosa-Hernández MA, Egan C, Kastens J. 2020. Trait-based aerial dispersal of arbuscular mycorrhizal fungi. *New Phytologist* **228**: 238–252.

Cotton TA. 2018. Arbuscular mycorrhizal fungal communities and global change: an uncertain future. *FEMS Microbiology Ecology* **94**: p.fiy179.

Coughlan AP, Dalpé Y, Lapointe L, Piché Y. 2000. Soil pH-induced changes in root colonization, diversity, and reproduction of symbiotic arbuscular mycorrhizal fungi from healthy and declining maple forests. *Canadian Journal of Forest Research* **30**: 1543–1554.

Davison J, García de León D, Zobel M, Moora M, Bueno CG, Barceló M, Gerz M, León D, Meng Y, Pillar VD *et al.* 2020. Plant functional groups associate with distinct arbuscular mycorrhizal fungal communities. *New Phytologist* **226**: 1117–1128.

Davison J, Moora M, Jairus T, Vasar M, Öpik M, Zobel, M. 2016. Hierarchical assembly rules in arbuscular mycorrhizal (AM) fungal communities. *Soil Biology and Biochemistry* **97**: 63–70.

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* **349**: 970–973.

Davison J, Moora M, Öpik M, Ainsaar L, Ducousso M, Hiiesalu I, Jairus T, Johnson N, Jourand P, Kalamees R *et al.* 2018. Microbial island biogeography: isolation shapes the life history characteristics but not diversity of root-symbiotic fungal communities. *ISME Journal* **12**: 2211–2224.

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

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**: 1969–1973.

Dumbrell AJ, Ashton PD, Aziz N, Feng G, Nelson M, Dytham C, Fitter AH, Helgason T. 2011. Distinct seasonal assemblages of arbuscular mycorrhizal fungi revealed by massively parallel pyrosequencing. *New Phytologist* **190**: 794–804.

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 Journal* **4**: 337–345.

Ellenberg H, Weber HE, Düll R, Wirth V, Werner W, Paulissen D. 1991. Zeigerwerte von Pfanzen in Mitteleuropa. *Scripta Geobotanika* **18**: 1–248.

Engelmoer DJ, Behm JE, Toby KE. 2014. Intense competition between arbuscular mycorrhizal mutualists in an *in vitro* root microbiome negatively affects total fungal abundance. *Molecular Ecology* **23**: 1584–1593.

Fitter AH. 2005. Darkness visible: reflections on underground ecology. *Journal of Ecology* **93**: 231–243.

Frey-Klett P, Garbaye JA, Tarkka M. 2007. The mycorrhiza helper bacteria revisited. *New Phytologist* **176**: 22–36.

García de León D, Davison J, Moora M, Öpik M, Feng H, Hiiesalu I, Jairus T, Koorem K, Liu Y, Phosri C *et al.* 2018. Anthropogenic disturbance equalizes diversity levels in arbuscular mycorrhizal fungal communities. *Global Change Biology* **24**: 2649–2659.

García de León D, Moora M, Öpik M, Neuenkamp L, Gerz M, Jairus T, Vasar M, Bueno CG, Davison J, Zobel M. 2016. Symbiont dynamics during ecosystem succession: co-occurring plant and arbuscular mycorrhizal fungal communities. *FEMS Microbiolgy. Ecology* **92**: p.fiw097.

Gazol A, Zobel M, Cantero JJ, Davison J, Esler KJ, Jairus T, Öpik M, Vasar M, Moora M. 2016. Impact of alien pines on local arbuscular mycorrhizal fungal communities—evidence from two continents. *FEMS Microbiology Ecology* **92**: fiw073.

Gerz M, Bueno CG, Ozinga WA, Zobel M, Moora M. 2018. Niche differentiation and expansion of plant species are associated with mycorrhizal symbiosis. *Journal of Ecology* **106**: 254–264.

Gittleman JL, Kot M. 1990. Adaptation: statistics and a null model for estimating phylogenetic effects. *Systematic Biology* **39**: 227–241.

Glassman SI, Wang IJ, Bruns TD. 2017. Environmental filtering by pH and soil nutrients drives community assembly in fungi at fine spatial scales. *Molecular Ecology* **26**: 6960–6973.

Grünfeld L, Wulf M, Rillig MC, Manntschke A, Veresoglou SD. 2020. Neighbours of arbuscular-mycorrhiza associating trees are colonized more extensively by arbuscular mycorrhizal fungi than their conspecifics in ectomycorrhiza dominated stands. *New Phytologist* **227**: 10–13.

Guisan A, Zimmermann NE. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**: 147–186.

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

Hastie T, Tibshirani R, Friedman J. 2009. *The elements of statistical learning: data mining, inference, and prediction, 2nd edn* New York, NY, USA: Springer.

Hazard C, Gosling P, Van Der Gast CJ, Mitchell DT, Doohan FM, Bending GD. 2013. The role of local environment and geographical distance in determining community composition of arbuscular mycorrhizal fungi at the landscape scale. *ISME Journal* **7**: 498–508.

van der Heijden MG, Bardgett RD, van Straalen NM. 2008. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters* **11**: 296–310.

Helgason T, Fitter AH. 2009. Natural selection and the evolutionary ecology of the arbuscular mycorrhizal fungi (Phylum Glomeromycota). *Journal of Experimental Botany* **60**: 2465–2480.

Helgason T, Merryweather JW, Young JPW, Fitter AH. 2007. Specificity and resilience in the arbuscular mycorrhizal fungi of a natural woodland community. *Journal of Ecology* **95**: 623–630.

Hoeksema JD, Bever JD, Chakraborty S, Chaudhary VB, Gardes M, Gehring CA, Hart MM, Housworth EA, Kaonongbua W, Klironomos JN *et al.* 2018. Evolutionary history of plant hosts and fungal symbionts predicts the strength of mycorrhizal mutualism. *Communications Biology* **1**: 116.

Johnson NC, Wilson GW, Bowker MA, Wilson JA, Miller RM. 2010. Resource limitation is a driver of local adaptation in mycorrhizal symbioses. *Proceedings of the National Academy of Sciences, USA* **107**: 2093–2098.

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**: 170122.

Kivlin SN. 2020. Global mycorrhizal fungal range sizes vary within and among mycorrhizal guilds but are not correlated with dispersal traits. *Journal of Biogeography* **47**: 1994–2001.

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

Kivlin SN, Muscarella R, Hawkes CV, Treseder KK. 2017. The predictive power of ecological niche modeling for global arbuscular mycorrhizal fungal biogeography. In: Tedersoo L, ed. *Biogeography of mycorrhizal symbiosis*. Cham, Switzerland: Springer, 143–158.

Klironomos J, Zobel M, Tibbett M, Stock WD, Rillig MC, Parrent JL, Moora M, Koch AM, Facelli JM, Facelli E *et al.* 2011. Forces that structure plant communities: quantifying the importance of the mycorrhizal symbiosis. *New Phytologist* **189**: 366–370.

Knegt B, Jansa J, Franken O, Engelmoer DJ, Werner GD, Bücking H, Kiers ET. 2016. Host plant quality mediates competition between arbuscular mycorrhizal fungi. *Fungal Ecology* **20**: 233–240.

Kraft NJ, Valencia R, Ackerly DD. 2008. Functional traits and niche-based tree community assembly in an Amazonian forest. *Science* **322**: 580–582.

Lee J, Lee S, Young JP. 2008. Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiology and Ecology* **65**: 339–349.

Leibold MA, McPeek MA. 2006. Coexistence of the niche and neutral perspectives in community ecology. *Ecology* **87**: 1399–1410.

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.

Llado S, Lopez-Mondejar R, Baldrian P. 2018. Drivers of microbial community structure in forest soils. *Applied Microbiology and Biotechnology* **102**: 4331–4338.

MacArthur R, Levins R. 1967. The limiting similarity, convergence, and divergence of coexisting species. *American Naturalist* **101**: 377–385.

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

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

Moora M, Berger S, Davison J, Öpik M, Bommarco R, Bruelheide H, Kühn I, Kunin WE, Metsis M, Rortais A *et al.* 2011. Alien plants associate with widespread generalist arbuscular mycorrhizal fungal taxa: evidence from a continental-scale study using massively parallel 454 sequencing. *Journal of Biogeography* **38**: 1305–1317.

Morton JB, Bentivenga SP, Bever JD. 1995. Discovery, measurement, and interpretation of diversity in arbuscular endomycorrhizal fungi (Glomales, Zygomycetes). *Canadian Journal of Botany* **73**: 25–32.

Oehl F, Laczko E, Bogenrieder A, Stahr K, Bösch R, van der Heijden M, Sieverding E. 2010. Soil type and land use intensity determine the composition of arbuscular mycorrhizal fungal communities. *Soil Biology and Biochemistry* **42**: 724–738.

Ohsowski BM, Zaitsoff PD, Öpik M, Hart MM. 2014. Where the wild things are: looking for uncultured Glomeromycota. *New Phytologist* **204**: 171–179.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P *et al.* 2019. *vegan: Community Ecology Package*. R package v.2.5-6. [WWW document] URL https://CRAN.R-project.org/package=vegan.

Oliverio AM, Geisen S, Delgado-Baquerizo M, Maestre FT, Turner BL, Fierer N. 2020. The global-scale distributions of soil protists and their contributions to belowground systems. *Science Advances* **6**: eaax8787.

Öpik M, Davison J. 2016. Uniting species-and community-oriented approaches to understand arbuscular mycorrhizal fungal diversity. *Fungal Ecology* **24**: 106–113.

Öpik M, Davison J, Moora M, Pärtel M, Zobel M. 2016. Response to Comment on "Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism". *Science* **351**: 826.

Öpik M, Davison J, Moora M, Zobel M. 2014. DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences. *Botany-Botanique* **92**: 135–147.

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.

Ozinga WA, Colles A, Bartish IV, Hennion F, Hennekens SM, Pavoine S, Poschlod P, Hermant M, Schaminée JH, Prinzing A. 2013. Specialists leave fewer descendants within a region than generalists. *Global Ecology and Biogeography* **22**: 213–222.

Pärtel M, Öpik M, Moora M, Tedersoo L, Szava-Kovats R, Rosendahl S, Rillig MC, Lekberg Y, Kreft H, Helgason T *et al.* 2017. Historical biome distribution and recent human disturbance shape the diversity of arbuscular mycorrhizal fungi. *New Phytologist* **216**: 227–238.

Paudel S, Longcore T, MacDonald B, McCormick MK, Szlavecz K, Wilson GW, Loss SR. 2016. Belowground interactions with aboveground consequences: invasive earthworms and arbuscular mycorrhizal fungi. *Ecology* **97**: 605–614.

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 London B* **276**: 4237–4245.

Rasmussen PU, Hugerth LW, Blanchet FG, Andersson AF, Lindahl BD, Tack AJ. 2018. Multiscale patterns and drivers of arbuscular mycorrhizal fungal communities in the roots and root-associated soil of a wild perennial herb. *New Phytologist* **220**: 1248–1261.

Rodríguez-Echeverría S, de la Pena E, Moens M, Freitas H, van der Putten WH. 2009. Can root-feeders alter the composition of AMF communities? Experimental evidence from the dune grass *Ammophila arenaria*. *Basic and Applied Ecology* **10**: 131–140.

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**: e2584.

Rosendahl S, Mcgee P, Morton JB. 2009. Lack of global population genetic differentiation in the arbuscular mycorrhizal fungus *Glomus mosseae* suggests a recent range expansion which may have coincided with the spread of agriculture. *Molecular Ecology* **18**: 4316–4329.

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 Journal* **12**: 17–30.

Siqueira JO, Hubbell DH, Mahmud AW. 1984. Effect of liming on spore germination, germ tube growth and root colonization by vesicular-arbuscular mycorrhizal fungi. *Plant and Soil* **76**: 115–124.

Smith SE, Read DJ. 2010. *Mycorrhizal symbiosis*. Cambridge, UK: Academic Press, 2010.

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

Sýkorová Z, Ineichen K, Wiemken A, Redecker D. 2007. The cultivation bias: different communities of arbuscular mycorrhizal fungi detected in roots from the field, from bait plants transplanted to the field, and from a greenhouse trap experiment. *Mycorrhiza* **18**: 1–14.

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

Tedersoo L, Bahram M, Zobel M. 2020. How mycorrhizal associations drive plant population and community biology. *Science* **367**: 6480.

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.

Tibbett M, Cairney JW. 2007. The cooler side of mycorrhizas: their occurrence and functioning at low temperatures. *Botany-Botanique* **85**: 51–62.

Tilman D. 1982. *Resource competition and community structure*. Princeton, NJ, USA: Princeton University Press.

Torrez V, Ceulemans T, Mergeay J, De Meester L, Honnay O. 2016. Effects of adding an arbuscular mycorrhizal fungi inoculum and of distance to donor sites on plant species recolonization following topsoil removal. *Applied Vegetation Science* **19**: 7–19.

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

U'Ren JM, Riddle JM, Monacell JT, Carbone I, Miadlikowska J. 2014. Tissue storage and primer selection influence pyrosequencing-based inferences of diversity and community composition of endolichenic and endophytic fungi. *Molecular Ecology Resources* **14**: 1032–1048.

Vályi K, Rillig MC, Hempel S. 2015. Land-use intensity and host plant identity interactively shape communities of arbuscular mycorrhizal fungi in roots of grassland plants. *New Phytologist* **205**: 1577–1586.

Van Geel M, Jacquemyn H, Plue J, Saar L, Kasari L, Peeters G, van Acker K, Honnay O, Ceulemans T. 2018. Abiotic rather than biotic filtering shapes the arbuscular mycorrhizal fungal communities of European seminatural grasslands. *New Phytologist* **220**: 1262–1272.

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.

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

Veresoglou SD, Liu L, Xu T, Rillig MC, Wang M, Wang J, Chen Y, Hu Y, Hao Z, Chen B. 2019. Biogeographical constraints in Glomeromycotinan distribution across forest habitats in China. *Journal of Ecology* **107**: 684–695.

Větrovský T, Kohout P, Kopecký M, Machac A, Man M, Bahnmann BD, Brabcová V, Choi J, Meszárošová L, Human ZR. 2019. A meta-analysis of global fungal distribution reveals climate-driven patterns. *Nature Communications* **10**: 5142.

Wang GM, Stribley DP, Tinker PB, Walker C. 1993. Effects of pH on arbuscular mycorrhiza I. Field observations on the long-term liming experiments at Rothamsted and Woburn. *New Phytologist* **124**: 465–472.

Wasof S, Lenoir J, Aarrestad PA, Alsos IG, Armbruster WS, Austrheim G, Bakkestuen V, Birks HJB, Bråthen KA, Broennimann O. 2015. Disjunct

populations of European vascular plant species keep the same climatic niches. *Global Ecology and Biogeography* **24**: 1401–1412.

Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Davies J, Grytnes JA, Harrison SP *et al.* 2010. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters* **13**: 1310–1324.

Wurzburger N, Brookshire EN, McCormack ML, Lankau RA. 2017. Mycorrhizal fungi as drivers and modulators of terrestrial ecosystem processes. *New Phytologist* **213**: 996–999.

Zhao X, Yang Y, Shen H, Geng X, Fang, J. 2019. Global soil–climate–biome diagram: linking surface soil properties to climate and biota. *Biogeosciences* **16**: 2857–2871.

Zimmermann NE, Edwards TC Jr, Graham CH, Pearman PB, Svenning J-C. 2010. New trends in species distribution modelling. *Ecography* **33**: 985–989.

Supporting Information

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

Fig. S1 Map showing sampled locations and habitat types.

Fig. S2 Q–Q plots based on quantile residuals from full GLM models including spatial and environmental predictor variables.

Fig. S3 Q–Q plots from full linear models including spatial and environmental predictor variables.

Fig. S4 Sampling effects on estimated niche characteristics.

Fig. S5 Phylogenetic tree (small subunit nuclear ribosomal DNA) of AM fungal VT showing family groupings and, major clades within the most abundant family, Glomeraceae.

Fig. S6 Within-sample VT rarefaction curves.

Fig. S7 Phylogenetic correlation in the explanatory power of spatial and environmental variables in models of virtual taxon relative abundance in soils worldwide.

Fig. S8 Explanatory power (R^2) of different environmental variables in models of virtual taxon (VT) relative abundance in soils worldwide.

Fig. S9 Phylogenetic correlation (small subunit nuclear ribosomal DNA phylogenetic tree) in the fitted values of VT presence–absence in models containing environmental and spatial predictor variables.

Fig. S10 Niche width estimates along different environmental axes for VT in different families and clades.

Notes S1 Analysis of positive and negative controls.

Table S1 Sample details including location, environmental variables and whether the sample returned AM fungal sequencing data of sufficient quality.

Table S2 Selected neighbourhood matrices and MEMs for models virtual taxon (VT) presence–absence and relative abundance.

Table S3 Fitted models of virtual taxon (VT) presence–absence and relative abundance.

Table S4 Sequencing data table and virtual taxon (VT) taxonomic placement and characteristics.

Table S5 Phylogenetic correlation in characteristics of VT realised niches and correlations with AM fungal VT cultured status and spore diameter (for a subset of cultured VT for which the measure is available).

Table S6 Estimated niche optima and widths for AM fungal VT along different environmental axes.

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.