Article:
Partel, Meelis, Öpik, Maarja, Tedersoo, Leho et al. (12 more authors) (2017) Historical biome distribution and recent human disturbance shape the diversity of arbuscular mycorrhizal fungi. New Phytologist. pp. 1-12. ISSN 1469-8137

https://doi.org/10.1111/nph.14695

New Phytologist

Historical biome distribution and recent human disturbance shape the diversity of arbuscular mycorrhizal fungi

Journal:	*New Phytologist*
Manuscript ID	NPH-MS-2017-23679.R2
Manuscript Type:	MS - Regular Manuscript
Date Submitted by the Author:	05-Jun-2017
Complete List of Authors:	Pärtel, Meelis; University of Tartu, Institute of Ecology and Earth Sciences Öpik, Maarja; University of Tartu, Department of Botany Moora, Mari; University of Tartu, Department of Botany, Institute of Ecology and Earth Sciences; Tedersoo, Leho; University of Tartu, Institute of Botany and Ecology; Szava-Kovats, Robert; University of Tartu, Institute of Ecology and Earth Sciences Rosendahl, Soren; University of Copenhagen, Department of Biology, Mycology Rillig, Matthias; Free University Berlin, Institut fuer Biologie Lekberg, Ylva; MPG Ranch, Soil Ecology; University of Montana, Ecosystem and Conservation Sciences Kreft, Holger; Georg-August-University of Göttingen, Biodiversity, Macroecology & Conservation Biogeography Helgason, Thorunn; University of York, Department of Biology; Eriksson, Ove; Stockholm University, Department of Ecology, Environment and Plant Sciences Davison, John; University of Tartu, Department of Botany de Bello, Francesco; University of South Bohemia, Department of Botany Caruso, Tancredi; School of Biological Sciences, Queen's University of Belfast Zobel, Martin; University of Tartu, Department of Botany
Key Words:	Biodiversity, Dark diversity, Ice Age, Mycorrhizae, Quaternary, Species pool, Tropical grassy biome, Wilderness

Historical biome distribution and recent human disturbance shape the diversity of arbuscular mycorrhizal fungi

Pärtel M[1], Öpik M[1], Moora M[1], Tedersoo L[2], Szava-Kovats R[1], Rosendahl S[3], Rillig MC[4,5], Lekberg Y[6,7], Kreft H[8], Helgason T[9], Eriksson O[10], Davison J[1], de Bello F[11], Caruso T[12], Zobel M[1]

[1]Department of Botany, Institute of Ecology and Earth Sciences, University of Tartu, Lai 40, Tartu, 51005, Estonia

[3]Department of biology, Sect. Ecology & Evolution, University of Copenhagen, Universitetsparken 15, Building 3, DK-2100 Copenhagen, Denmark

[2]Natural History Museum, University of Tartu, Vanemuise 46, Tartu, 51014, Estonia

[4]Freie Universität Berlin, Institute of Biology, Altensteinstr. 6, D-14195 Berlin, Germany

[5]Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), D-14195 Berlin, Germany

[6]MPG Ranch, 1001 S. Higgins Ave, Missoula, MT 59801 USA

[7]Department of Ecosystem and Conservation Sciences, University of Montana, Missoula, MT, 59812, USA

[8] Department of Biodiversity, Macroecology and Biogeography, Georg-August-University Göttingen, Büsgenweg 1, 37077 Göttingen, Germany.

[9]Department of Biology, University of York, Heslington, York YO10 5DD, UK

[10]Department of Ecology, Environment and Plant Sciences, Stockholm University, 10691 Stockholm, Sweden

[11]Department of Botany, Faculty of Sciences, University of South Bohemia, Na Zlate Stoce 1, CZ-370 05 České Budějovice, and Institute of Botany, Czech Academy of Sciences, Dukelská 135, CZ-379 82, Třeboň, Czech Republic.

[12]School of Biological Sciences, Queen's University of Belfast, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland

Corresponding author: Meelis Pärtel, meelis.partel@ut.ee

Main text: 5877 words

3 figures (all in colour)

0 tables

Supplementary Information (2 figures, 9 tables)

Summary

The availability of global microbial diversity data, collected using standardized metabarcoding techniques, makes microorganisms promising models for investigating the role of regional and local factors in driving biodiversity.

We modelled the global diversity of symbiotic arbuscular mycorrhizal (AM) fungi using currently available data on AM fungal molecular diversity (SSU-rRNA gene sequences) in field samples. To differentiate between regional and local effects, we estimated species pools (sets of potentially suitable taxa) for each site, which are expected to reflect regional processes. We then calculated community completeness, an index showing the fraction of the species pool present, which is expected to reflect local processes.

We found significant spatial variation, globally in species pool size, as well as in local and dark diversity (absent members of the species pool). Species pool size was larger close to areas containing tropical grasslands during the last glacial maximum, which are possible centres of diversification. Community completeness was larger in regions of high wilderness (remoteness from human disturbance). Local diversity was correlated with wilderness and current connectivity to mountain grasslands.

Applying the species pool concept to symbiotic fungi facilitated a better understanding of how biodiversity can be jointly shaped by large-scale historical processes and recent human disturbance.

Keywords

Biodiversity, Dark diversity, Ice Age, Mycorrhizae, Quaternary, Species pool, Tropical grassy biome, Wilderness

Introduction

Global diversity patterns have frequently been described for macroorganisms, including vascular plants and vertebrates (Gaston, 2000, Orme *et al.*, 2005, Kreft & Jetz, 2007). Yet, understanding the relative roles of different processes in shaping diversity patterns is an ongoing challenge (Pärtel *et al.*, 2016). Local diversity patterns in any group of taxa are expected to emerge as a

consequence of simultaneous, and potentially confounding, effects of regional (evolutionary changes, historical dispersal) and local processes (dispersal in contemporary landscapes, local biotic and abiotic filters, natural and anthropogenic disturbances; Huston, 1994; Ricklefs, 2004, 2007; Zobel, 2016). Distinguishing between regional and local processes requires diversity data that are comparable and replicated over large spatial scales. Molecular identification of microbial taxa from environmental samples might provide data that are much closer to meeting this requirement than traditional sampling of macroorganisms. However, macroecology of microbes is a recent field (Hanson et al., 2012; Wardle & Lindahl, 2014) and descriptions of global diversity patterns and their potential underlying drivers are largely lacking.

Identifying species pools – sets of potentially available species that are able to inhabit and reproduce under particular habitat conditions in given sites (Cornell & Harrison, 2014) – is a useful starting point for distinguishing regional and local processes acting on diversity. Species pools develop via speciation under particular habitat conditions, as well as via historical migrations between regions with similar conditions (Zobel 2016; Pärtel et al. 2016). Hence, one may expect that species pools are shaped mainly by regional factors. Species pools can be partitioned into locally present and locally absent fractions; the latter has been referred to as dark diversity (Pärtel *et al.*, 2011). From these two pieces of information, community completeness – an index characterizing the share of the species pool present at a given site (Pärtel *et al.*, 2013) – can be calculated as the log-transformed ratio of local and dark diversity. Community completeness indicates how easily potentially suitable species reach and establish in local communities, but also how well local populations persist. Hence it can be expected that community completeness is mainly driven by local factors.

There is only limited empirical support for the theoretical expectations stemming from the species pool concept (see Lessard *et al*., 2012 and Zobel, 2016 for review). Empirical species pool studies have hitherto addressed vertebrates, insects and plants, but large scale generalizations have been limited due to the multitude of methods and scales used to assess diversity and the hugely variable depth of diversity data from different parts of the globe. Consequently, local diversity estimates used in large-scale comparisons have often been derived from coarse grid-based distributions, or even from distribution range maps, and have therefore lacked information about actual diversity in local communities. A more suitable approach to

disentangling the relative roles of regional and local factors in driving large-scale patterns of biodiversity is to use local community data that are collected in a comparable manner throughout an area of interest and take proper account of species pools.

The paucity of current data also poses challenges for dark diversity estimation (Pärtel *et al.*, 2016). For well-studied organisms, expert opinion has been used to estimate dark diversity, either by linking species to habitat types or giving indicator scores along the main environmental gradients (de Bello *et al.*, 2016). Current developments in mathematical dark diversity methods based on species co-occurrences or species distribution modelling provide a promising alternative (Lewis *et al.*, 2016; Ronk *et al.*, 2016). These techniques assume that co-occurring taxa share similar ecological preferences and possibly also joint biogeographic history. Such an assumption is probably valid for stable ecosystems but should be applied with caution to successional ecosystems where many species are not in equilibrium with environmental conditions.

Perhaps surprisingly, suitable data for exploring global biodiversity patterns and processes may already be available in the form of microbial community data. Microbial diversity estimates are frequently derived using fairly standardized metabarcoding approaches and thus seem to more easily satisfy criteria of comparability than existing macro-organism data sets (Taberlet *et al*., 2012; Ficetola *et al*., 2015). Although microbes had until recently received little attention in macroecology (Wardle & Lindahl, 2014), new information is accumulating rapidly (e.g. Põlme et al. 2013; Tedersoo *et al*., 2014; Pärtel *et al*., 2017; Maestre *et al*., 2015; Louca *et al*., 2016), providing suitable data for dark diversity calculations using species co-occurrences without relying on empirical expert opinion about habitat preferences.

A potentially suitable target for studying regional and local effects on diversity are the microscopic arbuscular mycorrhizal (AM) fungi (subphylum Glomeromycotina; Spatafora *et al.*, 2016). AM fungi live in symbiosis with the roots of about 80% of terrestrial plant species (Smith & Read, 2008) and provide nutrients (mainly P and N) to their host plants in exchange for plant-assimilated carbon. AM fungi alleviate plant abiotic stress and are able to increase plant resistance to pathogens (Smith & Read, 2008; Pozo *et al*., 2015). There is accumulating information about the geographic distribution of these fungi (Öpik *et al*., 2010, 2013; Kivlin *et al*., 2011; Yang *et al*., 2012; Tedersoo *et al*., 2014). Most recently, Davison *et al.* (2015)

analysed AM fungal diversity in plant roots based on systematic sampling of 67 sites globally and found little endemism at the continental scale. At the same time, the diversity of AM fungal communities varied in relation to environmental variables (precipitation, soil organic C content and pH), and spatial distance. The species pool concept promises a more powerful approach for disentangling possible large- and small-scale factors determining AM fungal diversity, such as proximity to centres of evolutionary diversification and the effect of contemporary human influence.

AM fungi have several advantages as a model group for studying global diversity patterns and underlying processes. Standardised methodologies for delineating AM fungal taxa (Öpik *et al*., 2014; Öpik & Davison, 2016) and processing environmental samples exist and are widely used (Hart *et al*., 2015). DNA-based species delimitation is challenging due to the scarcity of sequences from morphologically described species (Öpik & Davison, 2016), so phylogenetically-delimited sequence groups (phylogroups) are often used (groupings of taxa based on 97% similarity of the target gene sequence; Öpik *et al.*, 2010, 2014). Furthermore, the global diversity of such approximately species-level phylogroups of AM fungi is fairly low (< 2000 groups globally; Öpik *et al*., 2014; Öpik & Davison, 2016).

As well as addressing theoretical challenges concerning the roles of regional and local factors in driving observed diversity patterns, the study of global AM fungal diversity can provide additional specific information about the role of historical factors in shaping the global distribution patterns of these fungi. While Beck *et al*. (2012) emphasized the significance of integrating past environmental conditions into macroecological analyses, little is known about the effect of historical factors on global microbial diversity. Davison *et al*. (2015) recorded only a minor effect of continental paleogeographic history on AM fungal community composition. The more recent past, however, might have left an important imprint. For example, during the Quaternary period, glacial periods have been more common than warmer conditions, such as the current interglacial, and biodiversity might be better described by conditions during the most recent glaciation (e.g., the Last Glacial Maximum or LGM) than by contemporary factors (Weigelt *et al*., 2016). Biomes associated with large species pools might indicate regions where AM fungi have diversified.

Here, we use the framework of the species pool concept to study the effects of regional and local drivers on the diversity of AM fungal communities. We used the Maarj*AM* database (Öpik *et al.*, 2010) to compile data from all available studies addressing AM fungal molecular (SSU rRNA gene sequence) diversity in environmental samples. The specific objectives of the study were: (1) to quantify and map global patterns in the species pools, local diversity, dark diversity and community completeness of AM fungi; and (2) to link these AM fungal diversity measures to various regional and local drivers, including latitude, current and past (LGM) biome distribution, current and past climate, wilderness index (remoteness from human influence) and local vegetation type. Our results show that species pools, local diversity and dark diversity exhibited significant spatial structure at the global scale. Species pool and dark diversity were related to regional factors (LGM biome configuration and climate), community completeness to local factors (wilderness), and local diversity was jointly associated with regional and local factors (wilderness and current biome configuration).

Materials and Methods

We used the Maarj*AM* database (cf. Öpik *et al*., 2010; updated in November 2016) as a source of AM fungal distribution data. Maarj*AM* is a curated repository containing AM fungal sequence-based records from published studies, each including information about Virtual Taxa (VT) in a specific geographical location. VT are SSU rRNA gene sequence-based approximately species-level phylogroups of AM fungi, which are phylogenetically delimited on the basis of sequence similarity and clade support (Öpik *et al*., 2010, 2014). A record in the Maarj*AM* database represents the presence of a VT in a plant species at a site in the case of individual plant root-based records, or the presence of a VT at a site in the case of soil samples or mixed-root samples. The database includes records from both Sanger and 454 sequencing platforms and incorporates 2-3 representative sequences per VT per site or per plant species per site from each study (see Öpik *et al.,* 2010 for details). The Maarj*AM* database currently contains c. 24 000 SSU rRNA gene sequence records associated with c. 400 VT. We associated all records of VT to unique geographical coordinates (sites). We also used information about vegetation type recorded for

each site: woodland vegetation (forest, woodland, shrubland) or grassland (both natural and semi-natural). Records from disturbed successional habitats were excluded.

For further analysis, we selected only sites that were associated with at least 20 records, since very low numbers of records might not allow precise extrapolations of local diversity. This resulted in a total of 128 sites and 361 VT (Fig. 1a, Table S1).

We calculated four related diversity measures: i) species pool size, ii) local diversity, iii) dark diversity (the locally absent fraction of the species pool), and iv) community completeness (the ratio of local and dark diversity). Natural logarithm transformation was used for all these measures to express relative differences. On a log scale, differences indicate how many times diversity values differ, e.g. on a log scale the difference between 5 and 10 VT is equivalent to the difference between 50 and 100 VT rather than the difference between 50 and 55 VT. It should be noted that several of these diversity measures are inherently related (e.g. local and dark diversity are additive components of the species pool), and patterns from these measures are expected to covary. At the same time, the pairs local - dark diversity, and species pool size - community completeness are mathematically independent (Pärtel et al. 2013).

In order to estimate species pool size (we use this term for the number of AM fungal VT in the pool for simplicity), it is necessary to sum local diversity and dark diversity. Local diversity was determined from observations at individual sites. The number of records per site ranged from 20 to 815 (mean 125). To account for differences in sampling intensity between sites, we used the Shannon index-based effective number of species and extrapolation to an asymptote implemented in the iNEXT software (Hsieh et al., 2016). The asymptotic diversity equates to expected local diversity at full sample coverage *sensu* Hsieh *et al.* (2016). This technique made it possible to maximise use of the information in the original data, which would have been lost with rarefying approaches whereby many observations are removed (Chao et al., 2016). Supporting Information Figure S1 shows rarefaction and extrapolation curves for each site. On average, extrapolated local diversity was 1.3 times larger than observed local diversity. The ratio of extrapolated / observed local diversity was not related to sequencing platform and was not strongly spatially clustered (Fig S1b).

Dark diversity was estimated using species co-occurrence patterns (Lewis et al., 2016). This approach defines taxa as belonging to dark diversity when they are absent from a site but

otherwise frequently co-occur with those species present at the site. Thus, species that are locally present are used as indicators for absent species: if there are frequent co-occurrences, it is assumed that the species share similar ecological requirements. A co-occurrence index, also known as Beals index, was calculated for each VT in each site. Threshold values for assigning VT to the dark diversity were determined on a VT-by-VT basis since the co-occurrence index depends on species frequency (De Cáceres & Legendre, 2008). For each VT, we examined co-occurrence index values for all sites where it was present and recorded the minimum. Then, if the VT was absent from a site, but its co-occurrence index exceeded the minimum observed in sites where it was present, the VT was considered part of the dark diversity. See Lewis *et al.* (2016) for methodological details and working examples. Community completeness was calculated as the log-ratio of local and dark diversity (Pärtel *et al.*, 2013). Species pool size and community completeness were calculated on the assumption that local and dark diversity estimates represent distinct sets of taxa, i.e. without many overlapping taxa.

Geographical distribution

We predicted the global distribution of the four different diversity measures using Generalized Additive Models (GAMs) and the spline-over-the-sphere algorithm in R package mgcv, with the method 'sos.smooth' and the default arguments except k=30 (Wood, 2003). This model can predict smooth variation in diversity values over the globe without producing edges. For each model, we recorded its estimated degrees of freedom (*edf*), *F* and *P* values, and amount of variation described. We measured the predictive power of the model using cross-validation by dividing locations into random 20% bins and estimating values for bins using the rest of the data (Franklin, 2010). We then calculated the correlation between observed and predicted values. We present only prediction maps when predicted values were significantly correlated with observed values. As a measure of uncertainty in our predictions, we mapped the standard deviation of 100 global predictions using random subsets of 80% of sites.

AM fungal diversity drivers

In order to relate diversity values to possible drivers, we obtained measures of the following parameters for each site: (1) latitude, (2) current connectivity to biomes, (3) connectivity to biomes during the LGM, (4) major bioclimatic variables describing current conditions and (5) those during the LGM, (6) wilderness index (remoteness from human influence), and (7) local vegetation type.

We measured latitude as distance from the equator (km). Although latitude is not a biogeographic gradient *per se* and climate and biomes are expected to be more directly related to biodiversity, latitude has been often used in previous studies and we included it to permit comparison.

We used the current biome vector map from Olson et al. (2001) and the LGM (ca 21,000 yrs before present) biome vector map from Ray & Adams (2001). The current biome map defines 14 biomes, while the original LGM biome map defines 24 biomes. Therefore, we regrouped LGM biomes to match the current classifications (Supporting Information Table S2; Fig. 1b,c). To calculate connectivity to biomes, we constructed a grid of points equally distributed across the globe by using centroids of the ISEA3H geodesic discrete global grid system (Sahr et al., 2003). We used R package 'dggridR' to obtain 65,612 points. We determined biome identity for each point and applied Hanski's connectivity index (Hanski, 1994; Moilanen & Nieminen, 2002): Connectivity = $\sum \exp(-d/a)$; where d is the distance from the site to all terrestrial points of a biome. The parameter a defines the influence of distance in the exponential distribution and can be seen as the average influence distance. We used a values 500, 1000 and 2000 km. To improve its distribution, connectivity was ln-transformed for modelling.

For each site, we compiled 19 bioclimatic variables (Supporting Information Table S3) (Hijmans *et al*., 2005) to describe both current conditions and the conditions predicted for the LGM according to the Community Climate System Model (Braconnot *et al*., 2007). The current climate map had resolution of 5′ and the LGM climate map had resolution of 10′. Precipitation measures were ln-transformed. We collapsed the 19 variables to 4 principal components using correlation matrices. The four principal components described >90% of total variation. The first axis was strongly correlated with annual mean and winter temperature (r>0.9), the second axis with precipitation during the dry period (r>0.9). The third axis was more related to precipitation during the warm period (r>0.6), and the fourth axis to modern maximum temperature (r=0.5), or

diurnal temperature range during the LGM (r>0.6). See Supporting Information Table S3 for the full correlation table.

Wilderness can be defined as a continuous index quantifying remoteness and the level of disturbance by modern technological society (Carver & Fritz, 2016). This synthetic variable was first elaborated for Australia (Lesslie & Taylor, 1985), but later applied globally by UNEP-WCMC (http://www.unep-wcmc.org/resources-and-data/global-wilderness). Available data have a resolution of ca 1.4′, and for each site we calculated the mean index value for radiuses of 5, 10 and 20 km. It should be noted that we had already excluded disturbed sites, so high wilderness index values were indicative of low human impact in the vicinity of sample sites.

We obtained information from original publications about local vegetation type for each site from the Maarj*AM* database and classified each site broadly as grassland (both natural and semi-natural) or woodland (forest and shrublands). Unfortunately, information about other potential local drivers (e.g. geological and soil characteristics, host plants) was not available for all studied sites.

We used an information theoretical approach and compared models using Akaike Information Criterion corrected for sample size (AICc, Burnham & Anderson, 2002). We first standardized all our variables to have equal inputs of mean ±1 standard deviation using the R package 'arm' (Gelman 2008). This allows direct comparisons between model coefficients of both continuous and binary variables. Then we modelled each of the driver types separately. If there were several variables available for a driver type (e.g. connectivity to different biomes, wilderness within different radiuses, Supporting Information Tables S4, S5) we selected the variable for which the model resulted in the lowest AICc values. For latitude, principal components of climate and wilderness, we investigated both linear and quadratic relationships, since unimodal patterns are theoretically possible, and selected the model with the lower AICc value. For connectivity to biomes, we only considered linear models where diversity was positively related to connectivity.

In a second step, we examined 29 models: (1) the full model with seven variables, (2) seven univariate models, addressing each driver type in isolation, (3) and all pairwise variable combinations to examine pairs of regional and local drivers in combination. Model assumptions were verified by plotting residuals versus fitted values and each independent variable. We calculated the importance of each driver as the sum of Akaike weights from models where the

driver was included. Then we took the top-ranked models (ΔAICc <4) and used full model averaging to identify the most important variables (Grueber *et al.*, 2011). Several of the independent variables were correlated (e.g. latitude with climate and biomes, or past and current climate; see Supporting Information Table S6 for a correlation matrix). Model averaging, however, is relatively insensitive to such correlations (Freckleton, 2011). Details of the top-ranked model are given in Supporting Information Table S7, of model averaging in Table S8, and a summary of all initial models can be found in Table S9. The R package 'MuMIn' was used for multi-model inference (Bartón, 2016).

Results

AM fungal local diversity, species pool size, community completeness and dark diversity

Average richness was estimated to 60 VT per site (Shannon effective number of taxa), with values ranging between 6 and 216. Species pool size per site was on average 132 VT (range: 46 to 285) and dark diversity was on average 71 VT (range: 29 to 145). Relationships between local or dark diversity and species pool size are shown in Fig. 2. As expected, AM fungal local diversity co-varied with AM fungal species pool size but variation in dark diversity introduced considerable variation into this relationship. Local and dark diversity were negatively correlated, although not tightly (Fig. 2c). Average community completeness was slightly negative (-0.37), showing that dark diversity estimates often exceeded local diversity at sites. Variation in community completeness was, however, large (range: -2.7 to 1.3).

Global distribution of AM fungal diversity measures

AM fungal species pool size and local and dark diversity were non-randomly distributed across the globe. Spatial GAM models accounted for 34% of the variation in AM fungal species pool size (Fig. 1e; *edf*=14.1, *F*=1.6, *P*<0.0001), 12% of the variation in AM fungal local diversity (Fig. 1f; *edf*=4.8, *F*=0.4, *P*=0.016), and 45% of the variation in AM fungal dark diversity (Fig. 1g; *edf*=20.8, *F*=2.5, *P*<0.001). Large AM fungal species pools were found in southeastern

329 Africa and eastern South America. Small species pools occurred at higher latitudes of the
330 Northern Hemisphere, especially in North America. Higher local AM fungal diversity values
331 were found in southern South America and southern Africa. North America was characterized by
332 low values. Higher AM fungal dark diversity was found close to the equator, in eastern North
333 America, eastern Australia and New Zealand. Low dark diversity was found in northeastern
334 Asia, western North America and southern South America. Cross-validation revealed moderate
335 correlation between actual and predicted values for the species pool size (r=0.41, P<0.001) and
336 dark diversity (r=0.39, P<0.001), while the correlation between actual and predicted local
337 diversity was indicative of lower predictive power (r=0.20, P=0.025). All predictions for North
338 America (and for New Zealand's dark diversity) were associated with high uncertainty
339 (Supporting Information Fig. S2).

340 The spatial GAM for AM fungal community completeness was non-significant (*edf*=5.5, *F*=0.4,
341 *P*=0.052) and cross-validation showed that actual and predicted values of AM fungal community
342 completeness were not significantly related (r=0.08, P=0.367). Thus, community completeness
343 had no identifiable geographical pattern and is more likely linked to local factors. Therefore, we
344 cannot present a prediction map and present instead a map showing observed values for AM
345 fungal community completeness (Fig. 1h); sites with low and high completeness are frequently
346 found in close proximity.

347

348 *Relationships with tested regional and local drivers*

349 According to driver importance and model averaging, AM fungal species pool size was best
350 described by connectivity to Last Glacial Maximum (LGM) tropical grasslands and savannas
351 (Fig. 3a,b). No other driver had comparable importance or significance (Table S8). For AM
352 fungal local diversity, wilderness around the sample site and current connectivity to mountain
353 grasslands had higher importance (Fig. 3c). Wilderness was significant in model averaging (Fig.
354 3d, Table S8), but current connectivity to mountain grasslands was not (P=0.184, but still
355 significant in the univariate model, Table S8, coef.= 0.23, P=0.009). No clearly important driver
356 of AM fungal dark diversity emerged (Fig. 3e). In the averaged model, AM dark diversity was
357 significantly related to current temperature (PC1, Fig 3f, Table S8). Sites with higher annual or
358 winter temperatures exhibited significantly higher dark diversity estimates.

The degree of wilderness in the surrounding area was important in describing AM fungal community completeness (Fig. 3g) and in the averaged model the relationship was close to significant (P=0.08, Table S8). Wilderness significantly explained community completeness in the model where it was the sole explanatory variable (Fig 3h, Table S9). In bivariate plots, local diversity and community completeness formed triangular-shaped relationships with wilderness (Fig 3e,h): both high and low values of diversity or community completeness were recorded at low wilderness, while only high values were recorded at high wilderness.

Discussion

Here we show that application of the species pool concept to AM fungi can reveal previously undescribed global biodiversity patterns and disentangle the effects of potential underlying drivers. Our results support theoretical expectations that the species pool size is linked to regional (and historical) factors, community completeness is linked to local (and contemporary) factors, and local diversity is a result of both. Using a global data set, we found that the species pool, local diversity and dark diversity of AM fungi showed nonrandom global patterns, with distinct regions of high and low magnitude. By contrast, community completeness did not show significant global structure. AM fungal species pool size was larger in regions that were well connected to tropical grasslands during the Last Glacial Maximum (LGM) *c.* 21,000 y ago. Community completeness was higher at sites with lower human impact in the vicinity (larger wilderness). Local diversity was associated jointly with wilderness around the study site and current connectivity to mountain grasslands. Dark diversity was higher (i.e. a greater number of potentially suitable taxa were absent) in currently warm conditions.

Species pool size is related to historical biome distribution

The largest AM fungal species pools were identified in eastern and southern Africa and to a certain extent in eastern South America. These areas are dominated by tropical grasslands, which, together with sparse dry forests, form a distinct and diverse system called the tropical grassy biome (Parr *et al*., 2014). We found that AM fungal species pool size was primarily associated with the connectivity to areas of tropical grasslands during the LGM (Ray & Adams,

2001). During the LGM, tropical grasslands covered ca 21 million km^2 (currently ca 20 million km^2), of which 7 million km^2 have remained tropical grassland throughout the past 21000 years and constitute refugia. In fact, parts of the same areas have probably been covered by grasslands since the Miocene (Micheels, 2007). Given that glacial conditions have been more common than interglacials during the Quaternary (Weigelt *et al*., 2016), biome distribution during the LGM is representative of the predominant environmental configuration through much of recent evolutionary time.

The phylogenetic analysis by Davison *et al.* (2015) suggested that the diversification of the majority of current AM fungal VT occurred approximately within the period of 4-30 million years ago, a timing that is corroborated by other molecular clock estimates for particular AM fungal speciation events (reviewed by Öpik & Davison, 2016). This coincides with the appearance and expansion of grasslands (Strömberg, 2011; Strömberg *et al*., 2013; Parr *et al.*, 2014). High diversity of macroorganisms in particular habitats has often been associated with high availability of that habitats area in space and through time (Mittelbach *et al*., 2007). It is possible that developing grasslands created new and spatially (and temporally) very abundant (or 'voluminous', since roots occupy the three-dimensional space) habitat for AM fungi. Although the relative area of grasslands in global vegetation has never been very high, these habitats may be particularly relevant for AM fungi due to the high density and large total abundance of host plant roots. For instance, contemporary grasslands contribute about 68% of the global fine root surface area and 78% of global fine root length (Jackson *et al*., 1997). The difference between forests and grasslands is also evident at small scales: average live fine root length is 4.1 km/m^2 in tropical evergreen forests but 60.4 km/m^2 in tropical grasslands (Jackson *et al*., 1997). The appearance of this vast new grassland habitat may have led to higher diversification rates of AM fungi due to spatial effects (e.g. isolation by distance in a complex three-dimensional habitat), new niches due to the proliferation and spread of grassland plant species, or other mechanisms.

Local diversity is linked both to regional and local factors

In contrast to species pool size, local diversity was most strongly associated with wilderness around study sites. Wilderness is a synthetic measure that is inversely related to human impact (Carver & Fritz, 2016). It incorporates remoteness from modern human infrastructure such as

418 roads, buildings etc., and a lack of strong human influence such as high-input urban and
419 agricultural areas. In this study, we *a priori* omitted sites that were heavily disturbed, but the
420 wilderness index was calculated within radiuses of 5-20 kilometers around study sites. Thus, our
421 measure of wilderness probably reflected human influence on habitat patches neighbouring the
422 local sites under investigation. In this context, the results indicate that human influence can harm
423 meta-community systems and cause loss of taxa in unaffected patches (Lekberg *et al.*, 2007).
424 Recent overviews show a significant decline in global wilderness (Watson *et al*., 2016), which
425 may constitute a threat to local AM fungal diversity. Connectivity to current mountain grasslands
426 also had a positive effect on local diversity. The most plausible explanation for this is that it also
427 reflects relatively low human impact in mountainous areas (Sandel & Svenning, 2013).

428

429 *Higher dark diversity is recorded in warmer climates*

430 High dark diversity of AM fungi was found at lower latitudes: Central America, Sub-Saharan
431 Africa, eastern Asia and eastern Australia. Modelling also identified current annual temperature
432 as the best predictor of dark diversity. Why a greater share of otherwise suitable taxa should be
433 absent in warm areas is not easy to explain, but indicates either more restricted dispersal or more
434 frequent local extinctions. The sites with high dark diversity were often (sub)tropical moist or
435 dry forests, and dark diversity was higher in woodlands compared to grasslands (although this
436 model had low weight compared with the climate model). Woody vegetation in general hinders
437 wind dispersal of plants (Nathan *et al*., 2008) and the same might be true for AM fungi. Indeed,
438 forests exhibited higher spatial turnover of AM fungal communities compared to grasslands in a
439 recent global survey of AM fungal communities, and there was also a trend of decreasing forest
440 beta diversity along a latitudinal gradient (Davison *et al*., 2015). It is conceivable that high
441 spatial heterogeneity in (sub)tropical forests might explain why sampling sites towards the
442 equator lacked a larger number of suitable taxa and dark diversity was consequently higher.
443 However, to properly test this hypothesis we require further empirical studies of spatial structure
444 in AM fungal communities, in particular those inhabiting warmer biomes, such as tropical and
445 subtropical habitats.

446

Community completeness as an indicator of local processes

Community completeness of AM fungi varied among study sites but did not exhibit geographic structure. In contrast to species pool size and to a certain extent also to local diversity, variation in community completeness is not expected to contain the footprint of biogeographic history; rather it is expected to reflect local factors, such as barriers to dispersal, biotic interactions, or disturbances (Pärtel *et al*., 2013; Ronk *et al*., 2015). In our models the best descriptor of AM fungal community completeness was the degree of wilderness around study sites: completeness was high when wilderness was high nearby. Indeed, an adverse impact of intensive land use on AM fungi has been noted in earlier studies (Lopez-Garcia *et al*., 2013; Moora *et al*., 2014). However, further specific case studies are needed to disentangle the types of interaction and disturbance that might be responsible for low completeness of AM fungal communities in particular sites. There is evidence that AM fungal taxa with specific traits (ruderal, measured as ease of sporulation) are more common in anthropogenic habitats (Ohsowski *et al.,* 2014), possibly caused by differences in tolerance to anthropogenic disturbance (Hart & Reader, 2004; Säle et al. 2015). Alternatively, low wilderness may have a cascading effect through loss of functioning meta-communities within highly human-modified areas.

Methodological assumptions and potential limitations

Our findings rest on several methodological assumptions. To identify AM fungi we used phylogroups, in the form of 18S rRNA gene-defined VT, and not traditional taxonomically-defined species. VT are known to merge closely related morphospecies in some, but not all lineages of AM fungi, and across most of the Glomeromycotina phylogeny there is limited information about species boundaries with which to assess the exact taxonomic rank of VT (Öpik et al. 2014; Thiéry et al. 2016). Nonetheless, the rank of VT has been shown to capture ecologically-relevant responses to environmental gradients (Powell et al. 2011), suggesting that VT-based estimates of local diversity are meaningful even if precise species boundaries are unknown. For dark diversity estimates obtained using co-occurrence techniques, we assume that VT have similar ecological properties in distant parts of the globe. We are unaware of published evidence with which to assess this assumption. However, we excluded all successional sites where taxa might not be in equilibrium with their environment. We also assume that our local

and dark diversity measures can be used in parallel. Theoretically, our estimates of extrapolated local and dark diversity might include taxa present at sites but not recorded. In this case, the species pool size would be overestimated and community completeness would be underestimated. However, we do not expect over- or underestimation to be large. Present but unrecorded species are likely to occur at low abundance, and such species would contribute relatively little to local diversity estimates since the Shannon index counts taxa in proportion to their abundance (Chao *et al.*, 2016). However, we excluded sites for which we expected the sampling effort to be seriously limited. Furthermore, rare taxa often have too few co-occurrences to be included in dark diversity calculations (Ronk *et al*., 2016). Using observed rather than extrapolated diversity decreased average species pools from 132 to 112 and increased average community completeness from -0.76 to -0.37. Observed and extrapolated estimates of the species pool size and community completeness were strongly correlated (r=0.89, r=0.97, respectively). We anticipate that the accumulation of highly standardised local sampling data using high-throughput methods will further avoid uncertainty related to sampling adequacy and estimation of local and dark diversity.

Conclusions

Community theory predicts that regional drivers are primarily responsible for shaping species pool size, local drivers shape community completeness, and local diversity contains the footprint of both regional and local drivers (Pärtel *et al*., 2013; Cornell & Harrison, 2014; Zobel, 2016). Nevertheless, comprehensive empirical support for these predictions has been scarce. This study of global diversity patterns in AM fungi provides one of the first large-scale, empirical confirmations of the theory. Furthermore, this study found that the historical distribution of biomes during the LGM was the most important tested regional driver, whereas the degree of wilderness in the vicinity of a study site constituted the most important tested local driver of AM fungal diversity patterns.

Tropical grasslands and savannas harbored the largest species pool of AM fungal species and may thus represent evolutionary hotspots and important refugia. Remoteness from human influence was associated with higher local diversity and greater completeness of AM fungal communities. This is a warning signal that anthropogenic factors have shaped and will continue

to shape AM fungal communities to a significant extent. Although human impact on microbial communities has been reported elsewhere, our study provides the first evidence of potential global impacts.

Acknowledgements

We thank the New Phytologist Trust for funding the 16th New Phytologist Workshop on "Dark diversity of co-occurring arbuscular mycorrhizal fungi and host plants", held in Tartu, Estonia in January 2016. This paper is an outcome of that workshop. We are grateful to Ingolf Kühn, Thomas D. Bruns, Jason Pither and two anonymous referees who provided valuable comments on earlier versions of this work. This research has been supported by the Estonian Ministry of Education and Research (IUT20-28, IUT20-29), and by the European Union through the European Regional Development Fund (Centre of Excellence EcolChange). MCR acknowledges support from the German Federal Ministry of Education and Research (BMBF) within the Collaborative Project "Bridging in Biodiversity Science (BIBS)" (01LC1501A). TH acknowledges support from NERC standard research grant NE/M004864/1. YL is grateful to MPG Ranch for funding.

Author Contribution

All authors discussed the topic during the 16th New Phytologist Workshop and following e-mail exchanges. MÖ coordinated the workshop and the collaboration network. MP performed analyses. MZ coordinated writing of the paper. All authors discussed results and contributed to writing.

References

Barton K. 2016. *MuMIn: Multi-Model Inference. R package version 1.15.6.* [WWW document] URL https://CRAN.R-project.org/package=MuMIn [accessed 15 May 2017]

Beck J, Ballesteros-Mejia L, Buchmann CM, Dengler J, Fritz SA, Gruber B, Hof C, Jansen F, Knapp S, Kreft H et al. 2012. What's on the horizon for macroecology? *Ecography* **35**: 673-683.

Braconnot P, Otto-Bliesner B, Harrison S, Joussaume S, Peterchmitt JY, Abe-Ouchi A, Crucifix M, Driesschaert E, Fichefet T, Hewitt CD, et al. 2007. Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum - Part 1: experiments and large-scale features. *Clim. Past* **3**: 261-277.

Burnham KP, Anderson DR. 2002. *Model Selection and Multi-Model Inference. A Practical Information-Theoretic Approach*. New York, USA: Springer.

Carver SJ, Fritz S. 2016. *Mapping Wilderness: Concepts, Techniques and Applications*. Dordrecht, Netherlands: Springer.

Chao A, Chiu C-H, Jost L. 2016. Statistical challenges of evaluating diversity patterns across environmental gradients in mega-diverse communities. *Journal of Vegetation Science* **27**: 437-438.

Cornell HV, Harrison SP. 2014. What are species pools and when are they important? *Annual Review of Ecology, Evolution and Systematics* **45**: 45-67.

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.

de Bello F, Fibich P, Zelený D, Kopecký M, Mudrák O, Chytrý M, Pyšek P, Wild J, Michalcová D, Sádlo J et al. 2016. Measuring size and composition of species pools: a comparison of dark diversity estimates. *Ecology and Evolution* **6**: 4088-4101.

De Cáceres M, Legendre P. 2008. Beals smoothing revisited. *Oecologia* **156**: 657-669.

Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, De Barba M, Gielly L, Lopes CM, Boyer F, Pompanon F et al. 2015. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources* **15**: 543-556.

Franklin J. 2010. *Mapping species distributions: spatial inference and prediction*. Cambridge, UK: Cambridge University Press.

Freckleton RP. 2011. Dealing with collinearity in behavioural and ecological data: model averaging and the problems of measurement error. *Behavioral Ecology and Sociobiology* **65**: 91-101.

Gaston KJ. 2000. Global patterns in biodiversity. *Nature* **405**: 220-227.

Gelman A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* **27**: 2865-2873.

Grueber CE, Nakagawa S, Laws RJ, Jamieson IG. 2011. Multimodel inference in ecology and evolution: challenges and solutions. *Journal of Evolutionary Biology* **24**: 699-711.

Hanski I. 1994. A practical model of metapopulation dynamics. *Journal of Animal Ecology* **63**: 151-162.

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

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

Hart MM, Reader RJ. **2004.** Do arbuscular mycorrhizal fungi recover from soil disturbance differently? *Tropical Ecology* **45:** 97-111.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965-1978.

Hsieh TC, Ma KH, Chao A. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution* **7**: 1451-1456.

Huston, M.A. 1994. *Biological diversity*. Cambridge, UK; Cambridge University Press.

Jackson RB, Mooney HA, Schulze ED. 1997. A global budget for fine root biomass, surface area, and nutrient contents. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 7362-7366.

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

Kreft H, Jetz W. 2007. Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 5925-5930.

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

Lessard JP, Belmaker J, Myers JA, Chase JM, Rahbek C. 2012. Inferring local ecological processes amid species pool influences. *Trends in Ecology & Evolution* **27**: 600-607.

Lesslie RG, Taylor SG. 1985. The wilderness continuum concept and its implications for Australian wilderness preservation policy. *Biological Conservation* **32**: 309-333.

Lewis RJ, Szava-Kovats R, Pärtel M. 2016. Estimating dark diversity and species pools: An empirical assessment of two methods. *Methods in Ecology and Evolution* **7**: 104-113.

Lopez-Garcia A, Hempel S, Miranda JD, Rillig MC, Barea JM, Azcon-Aguilar C. 2013. The influence of environmental degradation processes on the arbuscular mycorrhizal fungal community associated with yew (*Taxus baccata* L.), an endangered tree species from Mediterranean ecosystems of Southeast Spain. *Plant and Soil* **370**: 355-366.

Louca S, Parfrey LW, Doebeli M. 2016. Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**: 1272-1277.

Maestre FT, Delgado-Baquerizo M, Jeffries TC, Eldridge DJ, Ochoa V, Gozalo B, Quero JL, García-Gómez M, Gallardo A, Ulrich W et al. 2015. Increasing aridity reduces soil microbial diversity and abundance in global drylands. *Proceedings of the National Academy of Sciences* **112**: 15684-15689.

Meiser A, Baílint M, Schmitt I. 2014. Meta-analysis of deep-sequenced fungal communities indicates limited taxon sharing between studies and the presence of biogeographic patterns. *New Phytologist* **201**: 623-635.

Micheels A, Bruch AA, Uhl D, Utescher T, Mosbrugger V. 2007. A Late Miocene climate model simulation with ECHAM4/ML and its quantitative validation with terrestrial proxy data. *Palaeogeography, Palaeoclimatology, Palaeoecology* **253**: 251-270.

Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB, Harrison SP, Hurlbert AH, Knowlton N, Lessios HA et al. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters* **10**: 315-331.

Moilanen A, Nieminen M. 2002. Simple connectivity measures in spatial ecology. *Ecology* **83**: 1131-1145.

Moora M, Davison J, Öpik M, Metsis M, Saks Ü, Jairus T, Vasar M, Zobel M. 2014. Anthropogenic land use shapes the composition and phylogenetic structure of soil arbuscular mycorrhizal fungal communities. *FEMS Microbiology Ecology* **90**: 609-621.

Nathan R, Schurr FM, Spiegel O, Steinitz O, Trakhtenbrot A, Tsoar A. 2008. Mechanisms of long-distance seed dispersal. *Trends in Ecology & Evolution* 23: 638-647.

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

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'Amico JA, Itoua I, Strand HE, Morrison JC et al. 2001. Terrestrial ecoregions of the worlds: a new map of life on Earth. *BioScience* **51**: 933-938.

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, Zobel M. 2014. DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences. *Botany* **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 ecosystem distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytologist* **188**: 223-241.

Öpik M, Zobel M, Cantero JJ, Davison J, Facelli JM, Hiiesalu I, Jairus T, Kalwij JM, Koorem K, Leal ME et al. 2013. Global sampling of plant roots expands the described molecular diversity of arbuscular mycorrhizal fungi. *Mycorrhiza* **23**: 411-430.

Orme CDL, Davies RG, Burgess M, Eigenbrod F, Pickup N, Olson VA, Webster AJ, Ding T-S, Rasmussen PC, Ridgely RS et al. 2005. Global hotspots of species richness are not congruent with endemism or threat. *Nature* **436**: 1016-1019.

Parr CL, Lehmann CER, Bond WJ, Hoffmann WA, Andersen AN. 2014. Tropical grassy biomes: misunderstood, neglected, and under threat. *Trends in Ecology & Evolution* **29**: 205-213.

Pärtel M, Szava-Kovats R, Zobel M. 2013. Community completeness: linking local and dark diversity within the species pool concept. *Folia Geobotanica* **48**: 307-317.

Pärtel M, Bennett JA, Zobel M. 2016. Macroecology of biodiversity: disentangling local and regional effects. *New Phytologist* **211**: 404-410.

Pärtel M, Szava-Kovats R, Zobel M. 2011. Dark diversity: shedding light on absent species. *Trends in Ecology & Evolution* **26**: 124-128.

Pärtel M. Zobel M, Öpik M, Tedersoo L. 2017. Global patterns in local and dark diversity, species pool size and community completeness in ectomycorrhizal fungi. In: Tedersoo L, ed. *Biogeography of Mycorrhizal Symbiosis. Ecological Studies 230*. Cham, Switzerland: Springer, 395-406.

Põlme S, Bahram M, Yamanaka T, Nara K, Dai YC, Grebenc T, Kraigher H, Toivonen M, Wang P-H, Matsuda Y, Naadel T, Kennedy PG, Kõljalg U, Tedersoo L. 2013. Biogeography of ectomycorrhizal fungi associated with alders (*Alnus* spp.) in relation to biotic and abiotic variables at the global scale. *New Phytologis*t **198**: 1239–1249.

Pozo MJ, Lopez-Raez JA, Azcon-Aguilar C, Garcia-Garrido JM. 2015. Phytohormones as integrators of environmental signals in the regulation of mycorrhizal symbioses. *New Phytologist* **205**: 1431-1436.

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.

Ricklefs RE. 2004. A comprehensive framework for global patterns in biodiversity. *Ecology Letters* **7**: 1-15.

Ricklefs RE. 2007. History and diversity: Explorations at the intersection of ecology and evolution. *American Naturalist* **170**: S56-S70.

Ronk A, Szava-Kovats R, Pärtel M. 2015. Applying the dark diversity concept to plants at the European scale. *Ecography* **38**: 1015-1025.

Ronk A, de Bello F, Fibich P, Pärtel M. 2016. Large-scale dark diversity estimates: new perspectives with combined methods. *Ecology and Evolution* **6**: 6266-6281.

Sahr K, White D, Kimerling AJ. 2003. Geodesic Discrete Global Grid Systems. *Cartography and Geographic Information Science* **30**: 121-134.

Säle V, Aguilera P, Laczko E, Mäder P, Berner A, Zihlmann U, van der Heijden MGA, Oehl F. **2015.** Impact of conservation tillage and organic farming on the diversity of arbuscular mycorrhizal fungi. *Soil Biology and Biochemistry* **84:** 38-52.

Sandel B, Svenning J-C. 2013. Human impacts drive a global topographic signature in tree cover. *Nature Communications* **4**: 2474.

Smith SE, Read DJ. 2008. *Mycorrhizal symbiosis*. 3. Amsterdam, Netherlands: Academic Press.

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

Strömberg CAE. 2011. Evolution of grasses and grassland ecosystems. *Annual Review of Earth and Planetary Sciences* **39**: 517-544.

Strömberg CAE, Dunn RE, Madden RH, Kohn MJ, Carlini AA. 2013. Decoupling the spread of grasslands from the evolution of grazer-type herbivores in South America. *Nature Communications* **4**: 1478

692 **Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH**. 2012. Environmental DNA. *Molecular*
693 *Ecology* **21**: 1789-1793.

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

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

701 **Wardle DA, Lindahl BD**. 2014. Disentangling global soil fungal diversity. *Science* **346**: 1052-
702 1053.

703 **Watson JEM, Shanahan DF, Di Marco M, Allan J, Laurance WF, Sanderson EW, Mackey**
704 **B, Venter O**. 2016. Catastrophic Declines in Wilderness Areas Undermine Global Environment
705 Targets. *Current Biology* **26**: 2929-2934.

706 **Weigelt P, Steinbauer MJ, Cabral JS, Kreft H**. 2016. Late Quaternary climate change shapes
707 island biodiversity. *Nature* **532**: 99-102.

708 **Wood SN**. 2003. Thin plate regression splines. *Journal of the Royal Statistical Society: Series B*
709 *(Statistical Methodology)* **65**: 95-114.

710 **Yang HS, Zang YY, Yuan YG, Tang JJ, Chen X**. 2012. Selectivity by host plants affects the
711 distribution of arbuscular mycorrhizal fungi: evidence from ITS rDNA sequence metadata. *BMC*
712 *Evolutionary Biology* **12**.

713 **Zobel M, Öpik M**. 2014. Plant and arbuscular mycorrhizal fungal (AMF) communities - which
714 drives which? *Journal of Vegetation Science* **25**: 1133-1140.

715 **Zobel M**. 2016. The species pool concept as a framework for studying patterns of plant diversity.
716 *Journal of Vegetation Science* **27**: 8-18.

717

Figure legends:

Fig.1. (a) Sampling locations of AM fungal communities from the MaarjAM database. We excluded sites where the number of recorded sequences was <20. Locations are slightly jittered to show overlapping points. (b, c) Current (Olson et al., 2001) and Last Glacial Maximum (LGM, ca 21000 yrs before present; Ray & Adams, 2001) distribution of biomes: 1: Tropical & Subtropical Moist Broadleaf Forests; 2: Tropical & Subtropical Dry Broadleaf Forests; 3: Tropical & Subtropical Coniferous Forests; 4: Temperate Broadleaf & Mixed Forests; 5: Temperate Conifer Forests; 6: Boreal Forests/Taiga; 7: Tropical & Subtropical Grasslands, Savannas & Shrublands; 8: Temperate Grasslands, Savannas & Shrublands; 9: Flooded Grasslands & Savannas; 10: Montane Grasslands & Shrublands; 11: Tundra; 12: Mediterranean Forests, Woodlands & Scrub; 13: Deserts & Xeric Shrublands; 14: Mangroves; 15: Not vegetated. (d) Wilderness (the degree to which a place is remote from and undisturbed by the influences of modern technological society; UNEP-WCMC). (e, f, g) Global smoothed maps of AM fungal species pool size (GAM, $R^2 = 0.34$), local diversity ($R^2 = 0.12$) and dark diversity ($R^2 = 0.45$). (h) Distribution of AM fungal community completeness across study sites. A smoothed prediction of is not presented because the predictive power of the corresponding model was low. Locations are slightly jittered to distinguish immediately neighbouring points. Colours indicate quantiles (e – h).

Fig. 2. Relationships between AM fungal local (a, c), dark diversity (b, c), and species pool size (a, b) at 128 sites worldwide. Local diversity was estimated as the asymptotic Shannon index-based effective number of taxa using coverage-based rarefaction and extrapolation from site records. Dark diversity was estimated based on VT co-occurrences globally (absent VT which generally co-occur with locally present VT and therefore likely fit local ecological conditions). AM fungal species pool (the theoretical set of VT that can inhabit a study site) is calculated by summing AM fungal local and dark diversity. Lines indicate the 1:1 relationship, i.e. the upper limit that local or dark diversity can have. Semi-transparent symbols are used to show overlapping values. The two outliers with large species pools originate from tropical rainforest in French Guiana, and temperate beech forest in Georgia. Local and dark diversity are negatively correlated (c, Spearman $r = -0.45$, $P<0.001$). Local vegetation type is shown (grasslands or woodlands).

Fig.3. Importance of potential drivers (sum of Akaike weights in models where the driver was included) determining AM fungal species pool size, local and dark diversity, and community completeness (a, c, e, g). Details on the best supported models are presented in Table S7. Scatter plots show relationships with the most significant drivers from model averaging (Table S8). Species pool size is related to the connectivity of LGM tropical grasslands (b, bivariate relationship: R2=0.17, P=<0.001), local diversity is related to wilderness in the vicinity (d, R2=0.08, P=0.002), dark diversity is related to current temperature (f, R2=0.14, P<0.001), community completeness is related to wilderness in the vicinity (h, R2=0.07, P=0.004). Species pool size, local and dark diversity are ln-transformed, completeness is the logratio of local vs. dark diversity. Connectivity, wilderness and climate PC1 have relative values without units.

759 Table S1. Summary of data used in analyses. Geographical coordinates, local vegetation type, number of
760 records (representative sequences from a sampling unit), number of Virtual Taxa (VT), primers and
761 sequencing platform used, and sources.

No.	Lat.	Lon.	Veg. type	rec	VT	Primers	Seq. Platform	Source
1	69.8	27.2	woodland	101	57	F: NS31 R: AML2	454 sequencing & Sanger	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
2	69.8	27.1	woodland	129	61	F: NS31 R: AML2	454 sequencing & Sanger	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
3	61.3	73.1	woodland	75	44	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
4	61.3	73.2	woodland	200	76	F: NS31 R: AML2	454 sequencing & Sanger	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
5	59.8	18.0	grassland	61	23	F: NS31 R: AM1 & F: NS31 R: AM1+AM2+AM3	Sanger	Santos-Gonzalez et al. 2007 Applied and Environmental Microbiology & Santos et al. 2006 New Phytologist
6	59.2	10.4	woodland	28	11	F: NS31 R: AM1	454 sequencing	Moora et al. 2011 Journal of Biogeography
7	59.0	26.1	woodland	263	40	F: NS31 R: AM1	Sanger	Davison et al. 2011 FEMS Microbiology Ecology & Opik et al. 2008 New Phytologist
8	58.6	23.6	grassland	135	58	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
9	58.6	23.6	grassland	142	87	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
10	58.6	23.6	grassland	88	21	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
11	58.4	25.3	woodland	27	11	F: NS31 R: AM1	Sanger	Opik et al. 2003 New Phytologist
12	58.3	27.3	woodland	78	25	F: NS31 R: AML2	454 sequencing	Davison et al. 2012 PLoS ONE
13	58.2	26.6	grassland	28	14	F: NS31 R: AM1	Sanger	Opik et al. 2003 New Phytologist
14	56.1	159.9	woodland	94	56	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
15	56.1	159.9	woodland	102	58	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
16	56.1	159.9	woodland	40	15	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
17	55.5	-2.2	grassland	57	29	F: NS31 R: AM1	Sanger	Vandenkoornhuyse et al. 2007 Proceedings of the National Academy of Sciences of the United States of America
18	54.1	-0.9	woodland	79	33	F: NS31 R: AM1	Sanger	Helgason et al. 1998 Nature & Helgason et al. 1999 Molecular Ecology & Helgason et al. 2002 Journal of Ecology & Helgason et al. 2007 Journal of Ecology
19	53.9	-1.4	grassland	36	26	F: NS31 R: AM1	Sanger	Dumbrell et al. 2010 Journal of Ecology
20	53.0	158.7	woodland	54	32	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
21	53.0	158.7	woodland	77	41	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
22	53.0	158.7	woodland	55	14	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
23	52.7	4.7	grassland	36	16	F: NS31 R: AM1	Sanger	Scheublin et al. 2004 Applied and Environmental Microbiology
24	50.8	-104.6	grassland	509	115	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
25	48.5	-79.3	woodland	24	11	F: NS31 R: AM1	Sanger	DeBellis & Widden 2006 FEMS Microbiology Ecology
26	47.8	107.1	grassland	206	67	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
27	47.8	107.1	grassland	261	93	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
28	47.5	10.1	grassland	106	63	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
29	47.5	10.1	grassland	101	60	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
30	46.6	16.0	grassland	20	16	F: NS31 R: AM1	Sanger	Macek et al. 2011 Applied and Environmental Microbiology
31	44.8	-0.4	woodland	175	69	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science

No.	Lat.	Lon.	Veg. type	rec	VT	Primers	Seq. Platform	Source
32	43.6	-1.2	woodland	262	95	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
33	43.5	104.1	grassland	239	78	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
34	43.0	104.1	grassland	179	69	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
35	42.0	116.3	grassland	27	20	F: NS31 R: AML2	Sanger	Chen et al. 2014 Soil Biology and Biochemistry
36	41.9	43.4	woodland	68	41	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
37	41.9	43.4	woodland	53	21	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
38	41.9	43.4	woodland	73	58	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
39	41.6	-79.5	woodland	25	7	F: NS31 R: AM1	Sanger	Burke 2008 American Journal of Botany
40	40.2	-111.1	grassland	22	8	F: VANS1 or GEOA2 or GEO11 R: GLOM1311R or SS1492	Sanger	Winther & Friedman 2007 American Journal of Botany
41	39.2	-86.2	woodland	90	49	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
42	39.2	-86.2	woodland	95	56	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
43	39.1	-96.6	grassland	37	15	F: NS31 R: AM1	Sanger	Jumpponen et al. 2005 Biology and Fertility of Soils
44	39.0	-123.1	grassland	35	14	F: NS31 R: AM1	Sanger	Hausmann & Hawkes 2009 New Phytologist
45	38.7	140.7	grassland	51	30	F: AMV4.5NF R: AMV4.5NR	Sanger	Saito et al. 2004 Mycorrhiza
46	38.7	-0.9	woodland	76	29	F: NS31 R: AM1 & F: NS31 R: AM1+AM2+AM3	Sanger	Alguacil et al. 2009 Environmental Microbiology & Alguacil et al. 2009 Microbial Ecology
47	38.2	-1.2	woodland	150	32	F: AML1 R: AML2	Sanger	Alguacil et al. 2011 Science of the Total Environment & Alguacil et al. 2011 Soil Biology and Biochemistry & Torrecillas et al. 2012 Applied and Environmental Microbiology
48	38.2	-1.8	woodland	25	10	F: NS31 R: AM1+AM2+AM3	Sanger	Alguacil et al. 2009 Applied and Environmental Microbiology
49	37.7	-1.7	woodland	71	21	F: AML1 R: AML2	Sanger	Alguacil et al. 2012 Soil Biology and Biochemistry
50	37.4	-2.8	woodland	726	71	F: NS31 R: AM1 & F: NS31 R: AML2	454 sequencing & Sanger	Palenzuela et al. 2012 Journal of Arid Environments & Sanchez-Castro et al. 2012 Mycorrhiza & Varela-Cervero et al. 2015 Environmental Microbiology
51	36.0	101.9	grassland	146	39	F: NS31 R: AML2	Sanger	Liu et al. 2012 New Phytologist
52	35.6	-116.2	grassland	61	24	F: NS31 R: AM1	Sanger	Schechter, S. P.; Bruns, T. D. 2013 PLoS ONE & Schechter, S.P.; Bruns, T.D. 2008 Molecular Ecology
53	35.2	135.4	woodland	29	8	F: NS31 R: AM1	Sanger	Yamato & Iwase 2005 Mycoscience
54	35.0	102.9	grassland	47	23	F: NS31 R: AML2	Sanger	Shi et al. 2014 PLoS ONE
55	33.7	101.9	grassland	68	33	F: NS31 R: AML2	Sanger	Shi et al. 2014 PLoS ONE
56	30.6	34.7	woodland	96	67	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
57	30.6	34.7	woodland	95	57	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
58	30.6	34.7	woodland	66	35	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
59	29.5	118.1	woodland	42	18	F: NS31 R: AM1 & F: NS31 R: AML2	454 sequencing	Moora et al. 2011 Journal of Biogeography & Opik et al. 2013 Mycorrhiza
60	29.5	118.1	woodland	47	20	F: NS31 R: AM1 & F: NS31 R: AML2	454 sequencing	Moora et al. 2011 Journal of Biogeography & Opik et al. 2013 Mycorrhiza
61	29.4	79.6	woodland	153	72	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
62	29.4	79.6	woodland	162	77	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science

No.	Lat.	Lon.	Veg. type	rec	VT	Primers	Seq. Platform	Source
63	29.4	118.2	woodland	63	28	F: NS31 R: AM1 & F: NS31 R: AML2	454 sequencing	Moora et al. 2011 Journal of Biogeography & Opik et al. 2013 Mycorrhiza
64	28.7	77.2	woodland	27	12	F: NS31 R: AM1	Sanger	Deepika & Kothamasi 2015 Mycorrhiza
65	22.4	81.9	woodland	158	83	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
66	22.4	81.9	woodland	169	76	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
67	20.1	-75.1	grassland	28	8	F: AML1 R: AML2	Sanger	Alguacil et al. 2012 PLoS ONE
68	16.9	100.5	woodland	215	99	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
69	16.9	100.5	woodland	77	28	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
70	15.2	-23.7	woodland	61	21	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
71	14.6	-17.0	grassland	136	81	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
72	14.6	-17.0	grassland	137	74	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
73	9.2	-79.9	woodland	63	34	F: NS31 R: AM1	Sanger	Husband et al. 2002 Molecular Ecology & Husband et al. 2002 FEMS Microbiology Ecology
74	9.0	38.6	woodland	23	12	F: GlomerWT0 R: one of either GlomerWT1, GlomerWT2, GlomerWT3, or GlomerWT4	Sanger	Wubet et al. 2006 Canadian Journal of Botany & Wubet et al. 2006 Mycological Research
75	5.3	-52.9	woodland	34	27	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
76	5.3	-52.9	woodland	65	57	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
77	5.3	-52.9	woodland	61	25	F: NS31 R: AML2	454 sequencing	Opik et al. 2013 Mycorrhiza
78	5.0	9.6	woodland	23	9	F: NS1 R: ITS4 & F: NS31 R: AM1	Sanger	Franke et al. 2006 Mycological Progress & Merckx & Bidartondo 2008 Proceedings of The Royal Society B
79	4.6	-52.2	woodland	44	34	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
80	4.6	-52.2	woodland	55	44	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
81	4.6	-52.2	woodland	66	32	F: NS31 R: AML2	454 sequencing	Opik et al. 2013 Mycorrhiza
82	0.6	10.4	woodland	297	82	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
83	0.6	10.4	woodland	249	93	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
84	-1.8	35.2	grassland	46	34	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
85	-1.8	35.2	grassland	75	60	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
86	-2.1	35.0	grassland	86	64	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
87	-2.3	34.5	grassland	90	59	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
88	-2.6	35.1	grassland	75	53	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
89	-2.7	35.1	grassland	141	68	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
90	-5.9	145.1	woodland	37	21	F: SSU817F R: SSU1196ngs	454 sequencing	Tedersoo et al. 2015 Science
91	-7.3	147.1	woodland	92	47	F: SSU817F R: SSU1196ngs	454 sequencing	Tedersoo et al. 2015 Science
92	-9.4	147.4	woodland	127	65	F: SSU817F R: SSU1196ngs	454 sequencing	Tedersoo et al. 2015 Science
93	-18.9	34.4	grassland	27	15	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
94	-18.9	34.4	grassland	54	27	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
95	-18.9	34.5	grassland	37	17	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
96	-18.9	34.5	grassland	57	28	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
97	-18.9	34.5	grassland	33	19	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New

No.	Lat.	Lon.	Veg. type	rec	VT	Primers	Seq. Platform	Source
								Phytologist
98	-18.9	34.4	grassland	71	34	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
99	-18.9	34.5	grassland	95	42	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
100	-18.9	34.4	grassland	119	52	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
101	-19.0	34.4	grassland	67	44	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
102	-19.0	34.4	grassland	180	84	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
103	-19.0	34.2	grassland	150	74	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
104	-19.0	34.2	grassland	181	94	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
105	-19.0	34.2	grassland	122	66	F: NS31 R: AML2	454 sequencing	Rodriguez-Echeverria et al. 2017 New Phytologist
106	-23.8	133.9	woodland	58	14	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
107	-23.8	133.9	woodland	156	70	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
108	-23.8	133.9	woodland	157	82	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
109	-24.7	28.7	grassland	222	76	F: NS31 R: AML2	454 sequencing & Sanger	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
110	-24.8	28.6	grassland	234	100	F: NS31 R: AML2	454 sequencing & Sanger	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
111	-28.6	-51.6	grassland	298	76	F: NS31 R: AML2	454 sequencing	Zobel et al., in prep.
112	-30.1	-51.7	grassland	487	103	F: NS31 R: AML2	454 sequencing	Zobel et al., in prep.
113	-31.2	-64.3	woodland	100	49	F: NS31 R: AML2	454 sequencing	Grilli et al. 2015 Environmental Microbiology
114	-32.8	-64.9	grassland	261	85	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
115	-32.8	-64.9	grassland	287	84	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
116	-33.7	151.2	woodland	42	12	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
117	-33.7	151.2	woodland	55	38	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
118	-33.7	151.2	woodland	34	23	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
119	-34.0	19.0	woodland	108	44	F: NS31 R: AML2	454 sequencing & Sanger	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
120	-34.0	19.0	woodland	100	41	F: NS31 R: AML2	454 sequencing & Sanger	Davison et al. 2015 Science & Opik et al. 2013 Mycorrhiza
121	-35.1	138.7	woodland	85	32	F: NS31 R: AML2	454 sequencing	Opik et al. 2013 Mycorrhiza
122	-35.1	138.7	woodland	227	86	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
123	-37.3	142.2	grassland	71	21	F: NS31 R: AML2	Sanger	Opik et al. 2013 Mycorrhiza
124	-37.3	142.2	grassland	271	71	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
125	-39.0	-71.4	woodland	778	75	F: NS31 R: AML2	454 sequencing	Gazol et al. 2016 FEMS Microbiology Ecology
126	-39.0	-71.4	woodland	815	81	F: NS31 R: AML2	454 sequencing	Gazol et al. 2016 FEMS Microbiology Ecology
127	-52.1	-71.4	grassland	190	79	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science
128	-52.1	-71.4	grassland	223	75	F: NS31 R: AML2	454 sequencing	Davison et al. 2015 Science

Fig. S1. (a) Shannon index based effective number of species for sites with varying numbers of records (number of representative sequences from a sampling unit in a site). Red lines show rarefaction and blue lines extrapolations. We used estimated local diversity extrapolated to the asymptote, i.e. full

767 sample coverage *sensu* Hsieh *et al.* (2016). (b) Increase due to extrapolation (extrapolated / observed
768 local diversity) and sequencing platform within study sites. Locations are slightly jittered to show
769 overlapping points.

770 Table S2. Homogenization of biome classifications between current and Last Glacial Maximum (LGM)
771 maps.

ID	Current	LGM
1	Tropical & Subtropical Moist Broadleaf Forests	Tropical rainforest
2	Tropical & Subtropical Dry Broadleaf Forests	Tropical woodland Monsoon or dry forest Tropical thorn scrub and scrub woodland
3	Tropical & Subtropical Coniferous Forests	Montane tropical forest
4	Temperate Broadleaf & Mixed Forests	Broadleaved temperate evergreen forest
5	Temperate Conifer Forests	---
6	Boreal Forests/Taiga	Open boreal woodlands Main Taiga
7	Tropical & Subtropical Grasslands, Savannas & Shrublands	Tropical grassland Savanna
8	Temperate Grasslands, Savannas & Shrublands	Temperate steppe grassland Forest steppe Dry steppe
9	Flooded Grasslands & Savannas	---
10	Montane Grasslands & Shrublands	Alpine tundra Montane Mosaic Subalpine parkland
11	Tundra	Tundra Steppe-tundra Polar and alpine desert
12	Mediterranean Forests, Woodlands & Scrub	Semi-arid temperate woodland or scrub
13	Deserts & Xeric Shrublands	Tropical semi-desert Tropical extreme desert Temperate desert Temperate semi-desert
14	Mangroves	---
15	Not vegetated	Not vegetated

772

773 Table S3. Correlation matrix of Bioclimatic PCA from current and Last Glacial Maximum predictions
774 (LGM). Very high correlations r>0.9 are indicated by coloured backgrounds.

	Current climate				LGM climate			
Climatic parameter	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
BIO1 = Annual Mean Temperature	0.94	-0.26	-0.09	0.15	0.95	-0.23	-0.14	0.05
BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))	0.11	-0.68	0.13	0.18	-0.45	0.24	0.43	0.66
BIO3 = Isothermality (BIO2/BIO7)	0.85	-0.09	-0.15	-0.25	0.6	0.28	0.23	0.65
BIO4 = Temperature Seasonality (standard deviation *100)	-0.88	-0.07	0.30	0.30	-0.84	-0.2	0.01	-0.35
BIO5 = Max Temperature of Warmest Month	0.68	-0.50	0.05	0.50	0.82	-0.32	-0.10	0.13
BIO6 = Min Temperature of Coldest Month	0.96	-0.04	-0.25	-0.02	0.97	-0.13	-0.20	0.04
BIO7 = Temperature Annual Range (BIO5-BIO6)	-0.8	-0.27	0.35	0.35	-0.87	-0.05	0.23	0.04
BIO8 = Mean Temperature of Wettest Quarter	0.72	-0.30	0.40	0.30	0.85	-0.37	0.00	-0.17
BIO9 = Mean Temperature of Driest Quarter	0.86	-0.16	-0.42	0.01	0.92	-0.11	-0.28	0.16
BIO10 = Mean Temperature of Warmest Quarter	0.76	-0.42	0.06	0.45	0.87	-0.36	-0.18	-0.07
BIO11 = Mean Temperature of Coldest Quarter	0.97	-0.14	-0.19	0.00	0.97	-0.13	-0.12	0.13
BIO12 = Annual Precipitation	0.63	0.68	0.30	-0.05	0.73	0.58	0.27	-0.13
BIO13 = Precipitation of Wettest Month	0.72	0.38	0.49	-0.20	0.83	0.25	0.41	-0.17
BIO14 = Precipitation of Driest Month	0.07	0.92	-0.09	0.29	0.09	0.94	-0.17	-0.09
BIO15 = Precipitation Seasonality (Coefficient of Variation)	0.31	-0.72	0.42	-0.36	0.37	-0.78	0.39	-0.07
BIO16 = Precipitation of Wettest Quarter	0.73	0.40	0.47	-0.17	0.82	0.29	0.40	-0.17
BIO17 = Precipitation of Driest Quarter	0.14	0.91	0.01	0.32	0.19	0.94	-0.17	-0.12
BIO18 = Precipitation of Warmest Quarter	0.35	0.43	0.69	0.00	0.51	0.27	0.62	-0.33
BIO19 = Precipitation of Coldest Quarter	0.24	0.79	-0.35	0.19	0.27	0.84	-0.33	-0.01

775

776

777

Table S4. Correlation between connectivity of biomes using different distances of influence (500, 1000 and 2000 km). We show only connectivity of biomes that had high importance: cur.10 – current mountain grasslands and shrublands, lgm.7 – Last Glacial Maximum tropical grasslands and savannas.

[Uploaded as a separate file]

Table S5. Correlation between wilderness measures using different radiuses (5, 10 and 20 km) around study sites.

[Uploaded as a separate file]

Table S6. Correlations between independent variables used in models: absolute latitude (abs.lat), connectivity to current and Last Glacial Maximum (LGM) biomes (cur# and lgm#, respectively: see numerical codes of biomes in Fig 1 or Table S1), four current and LGM climate principal components (PC#, PC#lgm, see Table S2 for numerical codes), wilderness and local vegetation type (grassland vs. woodland). For connectivity of biomes we included only the mean distance of influence 1000 km; other distances were highly correlated (see Table S4). For Wilderness we included here only radius of 10 km; other radiuses gave highly correlated values (see Table S5).

[Uploaded as a separate file]

Table S7. Top-ranked models (delta AICc < 4). All variables were standardized with 2 sd values. Polynomial fits are indicted by "+". See model averaging and details about variables in Table S8.

Study variable	Absolute latitude	Conn. current biomes	Conn. LGM biomes	Current climate	LGM climate	Wilderness	Vegetation type = grassland	adjR²	df	logLik	AICc	Delta AICc	Akaike weight
Species pool size	+		0.43					0.27	5	-77.9	166.3	0.00	0.23
			0.35	+				0.26	5	-78.1	166.7	0.38	0.19
			0.31		+			0.26	5	-78.2	166.9	0.56	0.17
			0.41					0.22	3	-80.5	167.1	0.80	0.16
			0.42				-0.1	0.23	4	-80.1	168.5	2.22	0.08
			0.38			0.07		0.23	4	-80.2	168.7	2.36	0.07
		0.0	0.42					0.22	4	-80.5	169.2	2.92	0.05
					+			0.22	4	-80.6	169.5	3.21	0.05
Local diversity				0.22		0.24		0.16	4	-83.9	176.2	0.00	0.73
			0.18			0.20		0.13	4	-85.5	179.3	3.07	0.16
				-0.1		0.25		0.13	4	-85.9	180.1	3.86	0.11
Dark diversity	-0.1	-0.4	0.28	0.57	+	-0.2	-0.1	0.38	10	-70.8	163.4	0.00	0.77
				0.44		-0.2		0.24	4	-79.4	167.2	3.76	0.12
				0.36			-0.2	0.24	4	-79.5	167.3	3.92	0.11
Community completeness			0.21			0.22		0.14	4	-85.1	178.5	0.00	0.25
		0.2				0.23		0.14	4	-85.2	178.7	0.22	0.23
	0.19	0.09	0.19	-0.1	-0.1	0.28	0.07	0.23	9	-80.1	179.7	1.21	0.14
					-0.2	0.26		0.12	4	-86.1	180.5	1.94	0.10
				-0.1		0.22		0.11	4	-86.6	181.5	2.97	0.06
			0.22	-0.2				0.11	4	-86.7	181.8	3.30	0.05
		0.17	0.19					0.11	4	-86.7	181.8	3.30	0.05
						0.23	0.14	0.11	4	-86.8	181.9	3.37	0.05
			0.22				0.16	0.11	4	-86.8	182.0	3.49	0.04
						0.26		0.09	3	-88.1	182.3	3.81	0.04

Table S8. Averaged models (full average) from top-ranked models (delta AICc<4, see Table S7). All variables were standardized with 2 sd values. Variables with P<0.1 are marked by bold font.

Study variable	Predictors	Coef.	Adj. SE	z value	P
Species pool size	**Connectivity to LGM tropical grasslands**	**0.37**	**0.16**	**2.29**	**0.022**
	Absolute latitude	0.01	0.44	0.02	0.982
	Absolute latitude ^{2}	0.24	0.48	0.49	0.626
	Current climate PC1 (temperature)	0.08	0.38	0.20	0.845
	Current climate PC1 (temperature)^2	0.18	0.43	0.43	0.667
	LGM climate PC1 (temperature)	0.21	0.58	0.35	0.725
	LGM climate PC1 (temperature)^2	0.22	0.46	0.47	0.640
	Vegetation type (grassland)	-0.01	0.03	0.18	0.859
	Wilderness	0.01	0.03	0.16	0.873
	Connectivity to current tropical moist forests	0.00	0.03	0.02	0.988
Local diversity	Connectivity to current mountain grasslands	0.16	0.12	1.33	0.184
	Wilderness	**0.23**	**0.09**	**2.63**	**0.009**
	Connectivity to LGM tropical grasslands	0.03	0.08	0.38	0.706
	Current climate PC4 (temp. warm periods)	-0.02	0.05	0.29	0.770
Dark diversity	Absolute latitude	-0.11	0.24	0.45	0.650
	Current climate PC1 (temperature)	**0.53**	**0.27**	**2.00**	**0.046**
	Connectivity to current mangroves	-0.28	0.21	1.32	0.188
	Connectivity to LGM tropical dry forests	0.20	0.15	1.51	0.130
	LGM climate PC1 (temperature)	-0.39	1.42	0.28	0.781
	LGM climate PC1 (temperature)^2	0.71	0.64	1.11	0.268
	Vegetation type (grassland)	-0.13	0.09	1.37	0.170
	Wilderness	-0.18	0.12	1.54	0.124
Community completeness	Connectivity to LGM deserts	0.11	0.12	0.90	0.368
	Wilderness	**0.22**	**0.12**	**1.73**	**0.083**
	Connectivity to current mountain grasslands	0.07	0.10	0.64	0.519
	Absolute latitude	0.03	0.08	0.35	0.727
	Current climate PC4 (temp. warm periods)	-0.03	0.07	0.40	0.687
	LGM climate PC4 (prec. dry periods)	-0.03	0.07	0.40	0.693
	Vegetation type (grassland)	0.02	0.06	0.37	0.712

Table S9. Details all models tested. Four dependent diversity measures (AM fungal species pool size, local diversity, dark diversity, and community completeness) are related to seven driver types: absolute latitude, connectivity to current and LGM biomes (see biome numbers from Tables S1, three distance of influence are used, 500 km, 1000 km and 2000 km, models with coefficient >0 are given since the negative connectivity has no biological meaning here), current and LGM climate (four principal components, PC1…PC4), wilderness index (mean value in radiuses 5 km 10 km and 20 km) and local vegetation type (grassland vs. woodland). For latitude, climate and wilderness both linear and polynomial models have been considered. Coefficients are comparable since all variables were standardized with 2 sd.

Study variable	Driver type	predictors	Coef	SE	t value	P	AICc	R^2
sp.pool.size	abs.lat	abs.lat	-0.37	0.08	-4.4	<0.001	172.4	0.14
sp.pool.size	abs.lat	poly(abs.lat, 2)1	-2.07	0.46	-4.5	<0.001	171.3	0.16
sp.pool.size	abs.lat	poly(abs.lat, 2)2	0.82	0.46	1.8	0.077	171.3	0.16
sp.pool.size	cur.biomes	cur.13.500	0.00	0.09	0.0	0.983	191.0	0.00
sp.pool.size	cur.biomes	cur.2.500	0.26	0.09	3.1	0.003	181.7	0.07
sp.pool.size	cur.biomes	cur.2.1000	0.23	0.09	2.6	0.011	184.3	0.05
sp.pool.size	cur.biomes	cur.2.2000	0.14	0.09	1.6	0.108	188.4	0.02
sp.pool.size	cur.biomes	cur.14.500	0.29	0.09	3.4	0.001	179.8	0.08
sp.pool.size	cur.biomes	cur.14.1000	0.27	0.09	3.2	0.002	181.3	0.07
sp.pool.size	cur.biomes	cur.14.2000	0.23	0.09	2.6	0.010	184.2	0.05
sp.pool.size	cur.biomes	cur.7.500	0.31	0.08	3.7	<0.001	177.8	0.10
sp.pool.size	cur.biomes	cur.7.1000	0.31	0.08	3.7	<0.001	177.8	0.10
sp.pool.size	cur.biomes	cur.7.2000	0.31	0.08	3.6	<0.001	178.2	0.10
sp.pool.size	cur.biomes	cur.1.500	0.34	0.08	4.1	<0.001	175.3	0.12
sp.pool.size	cur.biomes	cur.1.1000	0.33	0.08	4.0	<0.001	176.0	0.11
sp.pool.size	cur.biomes	cur.1.2000	0.30	0.09	3.5	0.001	179.2	0.09
sp.pool.size	cur.biomes	cur.10.500	0.27	0.09	3.1	0.002	181.6	0.07
sp.pool.size	cur.biomes	cur.10.1000	0.23	0.09	2.6	0.010	184.2	0.05
sp.pool.size	cur.biomes	cur.10.2000	0.12	0.09	1.3	0.186	189.2	0.01
sp.pool.size	cur.biomes	cur.9.500	0.02	0.09	0.2	0.866	191.0	0.00
sp.pool.size	cur.biomes	cur.9.1000	0.05	0.09	0.6	0.573	190.7	0.00
sp.pool.size	cur.biomes	cur.9.2000	0.08	0.09	1.0	0.343	190.1	0.01
sp.pool.size	lgm.biomes	lgm.12.500	0.02	0.09	0.2	0.833	191.0	0.00
sp.pool.size	lgm.biomes	lgm.13.500	0.14	0.09	1.5	0.128	188.6	0.02
sp.pool.size	lgm.biomes	lgm.13.1000	0.16	0.09	1.8	0.080	187.9	0.02
sp.pool.size	lgm.biomes	lgm.13.2000	0.16	0.09	1.8	0.073	187.7	0.03
sp.pool.size	lgm.biomes	lgm.2.500	0.05	0.09	0.5	0.603	190.7	0.00
sp.pool.size	lgm.biomes	lgm.2.1000	0.09	0.09	1.0	0.314	190.0	0.01
sp.pool.size	lgm.biomes	lgm.2.2000	0.11	0.09	1.2	0.234	189.5	0.01
sp.pool.size	lgm.biomes	lgm.1.500	0.32	0.08	3.7	<0.001	177.6	0.10
sp.pool.size	lgm.biomes	lgm.1.1000	0.27	0.09	3.1	0.002	181.3	0.07
sp.pool.size	lgm.biomes	lgm.1.2000	0.17	0.09	2.0	0.050	187.1	0.03
sp.pool.size	lgm.biomes	lgm.7.500	0.38	0.08	4.6	<0.001	170.8	0.15
sp.pool.size	lgm.biomes	lgm.7.1000	0.41	0.08	5.1	<0.001	167.1	0.17
sp.pool.size	lgm.biomes	lgm.7.2000	0.40	0.08	4.9	<0.001	169.1	0.16
sp.pool.size	lgm.biomes	lgm.3.500	0.24	0.09	2.7	0.007	183.6	0.06
sp.pool.size	lgm.biomes	lgm.3.1000	0.18	0.09	2.0	0.047	187.0	0.03
sp.pool.size	lgm.biomes	lgm.3.2000	0.13	0.09	1.5	0.138	188.8	0.02
sp.pool.size	lgm.biomes	lgm.4.500	0.08	0.09	0.9	0.383	190.2	0.01
sp.pool.size	lgm.biomes	lgm.4.1000	0.08	0.09	0.9	0.382	190.2	0.01

Study variable	Driver type	predictors	Coef	SE	t value	P	AICc	R^2
sp.pool.size	lgm.biomes	lgm.4.2000	0.08	0.09	0.9	0.382	190.2	0.01
sp.pool.size	cur.climate	PC1	0.36	0.08	4.3	<0.001	173.4	0.13
sp.pool.size	cur.climate	poly(PC1, 2)1	2.02	0.46	4.4	<0.001	170.9	0.16
sp.pool.size	cur.climate	poly(PC1, 2)2	0.99	0.46	2.2	0.034	170.9	0.16
sp.pool.size	cur.climate	PC2	0.00	0.09	-0.1	0.963	191.0	0.00
sp.pool.size	cur.climate	poly(PC2, 2)1	-0.02	0.50	-0.1	0.963	193.1	0.00
sp.pool.size	cur.climate	poly(PC2, 2)2	-0.02	0.50	0.0	0.973	193.1	0.00
sp.pool.size	cur.climate	PC3	0.14	0.09	1.6	0.114	188.5	0.02
sp.pool.size	cur.climate	poly(PC3, 2)1	0.79	0.50	1.6	0.114	189.3	0.03
sp.pool.size	cur.climate	poly(PC3, 2)2	-0.55	0.50	-1.1	0.267	189.3	0.03
sp.pool.size	cur.climate	PC4	-0.10	0.09	-1.2	0.242	189.6	0.01
sp.pool.size	cur.climate	poly(PC4, 2)1	-0.59	0.50	-1.2	0.242	190.4	0.02
sp.pool.size	cur.climate	poly(PC4, 2)2	-0.57	0.50	-1.2	0.252	190.4	0.02
sp.pool.size	lgm.climate	PC1	0.36	0.08	4.3	<0.001	173.2	0.13
sp.pool.size	lgm.climate	poly(PC1, 2)1	2.03	0.46	4.4	<0.001	169.5	0.17
sp.pool.size	lgm.climate	poly(PC1, 2)2	1.10	0.46	2.4	0.018	169.5	0.17
sp.pool.size	lgm.climate	PC2	-0.03	0.09	-0.4	0.720	190.9	0.00
sp.pool.size	lgm.climate	poly(PC2, 2)1	-0.18	0.50	-0.4	0.722	193.0	0.00
sp.pool.size	lgm.climate	poly(PC2, 2)2	-0.01	0.50	0.0	0.990	193.0	0.00
sp.pool.size	lgm.climate	PC3	0.07	0.09	0.8	0.400	190.3	0.01
sp.pool.size	lgm.climate	poly(PC3, 2)1	0.42	0.50	0.9	0.400	191.0	0.02
sp.pool.size	lgm.climate	poly(PC3, 2)2	-0.58	0.50	-1.2	0.248	191.0	0.02
sp.pool.size	lgm.climate	PC4	-0.11	0.09	-1.3	0.212	189.4	0.01
sp.pool.size	lgm.climate	poly(PC4, 2)1	-0.63	0.50	-1.3	0.212	190.3	0.02
sp.pool.size	lgm.climate	poly(PC4, 2)2	-0.56	0.50	-1.1	0.263	190.3	0.02
sp.pool.size	wild	wild.5	0.19	0.09	2.2	0.028	186.1	0.04
sp.pool.size	wild	poly(wild.5, 2)1	1.09	0.49	2.2	0.029	188.2	0.04
sp.pool.size	wild	poly(wild.5, 2)2	-0.03	0.49	-0.1	0.945	188.2	0.04
sp.pool.size	wild	wild.10	0.20	0.09	2.2	0.027	186.0	0.04
sp.pool.size	wild	poly(wild.10, 2)1	1.10	0.49	2.2	0.028	188.0	0.04
sp.pool.size	wild	poly(wild.10, 2)2	-0.22	0.49	-0.4	0.663	188.0	0.04
sp.pool.size	wild	wild.20	0.23	0.09	2.7	0.009	184.0	0.05
sp.pool.size	wild	poly(wild.20, 2)1	1.30	0.49	2.7	0.009	185.9	0.05
sp.pool.size	wild	poly(wild.20, 2)2	-0.24	0.49	-0.5	0.629	185.9	0.05
sp.pool.size	veg.type	veg.type = grassl.	-0.02	0.09	-0.3	0.792	190.9	0.00
local.diversity	abs.lat	abs.lat	-0.16	0.09	-1.8	0.080	187.9	0.02
local.diversity	abs.lat	poly(abs.lat, 2)1	-0.87	0.49	-1.8	0.079	187.8	0.04
local.diversity	abs.lat	poly(abs.lat, 2)2	0.72	0.49	1.5	0.146	187.8	0.04
local.diversity	cur.biomes	cur.13.500	0.05	0.09	0.6	0.561	190.7	0.00
local.diversity	cur.biomes	cur.13.1000	0.02	0.09	0.3	0.789	190.9	0.00
local.diversity	cur.biomes	cur.13.2000	0.02	0.09	0.3	0.784	190.9	0.00
local.diversity	cur.biomes	cur.12.500	0.03	0.09	0.3	0.771	190.9	0.00
local.diversity	cur.biomes	cur.12.1000	0.02	0.09	0.2	0.822	190.9	0.00
local.diversity	cur.biomes	cur.12.2000	0.01	0.09	0.1	0.890	191.0	0.00
local.diversity	cur.biomes	cur.2.500	0.11	0.09	1.3	0.212	189.4	0.01
local.diversity	cur.biomes	cur.2.1000	0.12	0.09	1.3	0.191	189.3	0.01
local.diversity	cur.biomes	cur.2.2000	0.10	0.09	1.2	0.240	189.6	0.01
local.diversity	cur.biomes	cur.14.500	0.10	0.09	1.1	0.257	189.7	0.01
local.diversity	cur.biomes	cur.14.1000	0.07	0.09	0.8	0.409	190.3	0.01
local.diversity	cur.biomes	cur.14.2000	0.05	0.09	0.6	0.581	190.7	0.00
local.diversity	cur.biomes	cur.7.500	0.20	0.09	2.3	0.026	186.0	0.04
local.diversity	cur.biomes	cur.7.1000	0.20	0.09	2.3	0.021	185.5	0.04
local.diversity	cur.biomes	cur.7.2000	0.23	0.09	2.7	0.008	183.8	0.05

Study variable	Driver type	predictors	Coef	SE	t value	P	AICc	R^2
local.diversity	cur.biomes	cur.1.500	0.18	0.09	2.1	0.041	186.7	0.03
local.diversity	cur.biomes	cur.1.1000	0.19	0.09	2.2	0.028	186.1	0.04
local.diversity	cur.biomes	cur.1.2000	0.20	0.09	2.3	0.024	185.8	0.04
local.diversity	cur.biomes	cur.10.500	0.25	0.09	2.9	0.004	182.5	0.06
local.diversity	cur.biomes	cur.10.1000	0.26	0.09	3.0	0.003	182.0	0.07
local.diversity	cur.biomes	cur.10.2000	0.23	0.09	2.6	0.010	184.3	0.05
local.diversity	cur.biomes	cur.9.500	0.13	0.09	1.5	0.131	188.7	0.02
local.diversity	cur.biomes	cur.9.1000	0.17	0.09	1.9	0.058	187.3	0.03
local.diversity	cur.biomes	cur.9.2000	0.20	0.09	2.3	0.022	185.7	0.04
local.diversity	lgm.biomes	lgm.13.500	0.21	0.09	2.4	0.017	185.2	0.04
local.diversity	lgm.biomes	lgm.13.1000	0.24	0.09	2.8	0.006	183.3	0.06
local.diversity	lgm.biomes	lgm.13.2000	0.25	0.09	2.9	0.005	183.0	0.06
local.diversity	lgm.biomes	lgm.1.500	0.11	0.09	1.2	0.217	189.4	0.01
local.diversity	lgm.biomes	lgm.1.1000	0.09	0.09	1.0	0.301	189.9	0.01
local.diversity	lgm.biomes	lgm.1.2000	0.05	0.09	0.6	0.584	190.7	0.00
local.diversity	lgm.biomes	lgm.7.500	0.23	0.09	2.7	0.008	183.8	0.05
local.diversity	lgm.biomes	lgm.7.1000	0.25	0.09	2.9	0.004	182.5	0.06
local.diversity	lgm.biomes	lgm.7.2000	0.27	0.09	3.1	0.002	181.6	0.07
local.diversity	lgm.biomes	lgm.3.500	0.05	0.09	0.5	0.602	190.7	0.00
local.diversity	lgm.biomes	lgm.3.1000	0.01	0.09	0.1	0.946	191.0	0.00
local.diversity	lgm.biomes	lgm.4.500	0.12	0.09	1.3	0.183	189.2	0.01
local.diversity	lgm.biomes	lgm.4.1000	0.12	0.09	1.3	0.183	189.2	0.01
local.diversity	lgm.biomes	lgm.4.2000	0.12	0.09	1.3	0.183	189.2	0.01
local.diversity	cur.climate	PC1	0.09	0.09	1.1	0.296	189.9	0.01
local.diversity	cur.climate	poly(PC1, 2)1	0.52	0.50	1.1	0.292	188.8	0.03
local.diversity	cur.climate	poly(PC1, 2)2	0.88	0.50	1.8	0.079	188.8	0.03
local.diversity	cur.climate	PC2	-0.12	0.09	-1.3	0.192	189.3	0.01
local.diversity	cur.climate	poly(PC2, 2)1	-0.65	0.50	-1.3	0.193	191.3	0.01
local.diversity	cur.climate	poly(PC2, 2)2	0.13	0.50	0.3	0.791	191.3	0.01
local.diversity	cur.climate	PC3	0.02	0.09	0.3	0.794	190.9	0.00
local.diversity	cur.climate	poly(PC3, 2)1	0.13	0.50	0.3	0.794	192.5	0.01
local.diversity	cur.climate	poly(PC3, 2)2	0.38	0.50	0.8	0.454	192.5	0.01
local.diversity	cur.climate	PC4	-0.20	0.09	-2.3	0.025	185.9	0.04
local.diversity	cur.climate	poly(PC4, 2)1	-1.12	0.49	-2.3	0.025	186.6	0.05
local.diversity	cur.climate	poly(PC4, 2)2	-0.58	0.49	-1.2	0.236	186.6	0.05
local.diversity	lgm.climate	PC1	0.14	0.09	1.6	0.115	188.5	0.02
local.diversity	lgm.climate	poly(PC1, 2)1	0.79	0.50	1.6	0.116	190.1	0.02
local.diversity	lgm.climate	poly(PC1, 2)2	0.35	0.50	0.7	0.486	190.1	0.02
local.diversity	lgm.climate	PC2	-0.12	0.09	-1.4	0.163	189.0	0.02
local.diversity	lgm.climate	poly(PC2, 2)1	-0.70	0.50	-1.4	0.164	190.5	0.02
local.diversity	lgm.climate	poly(PC2, 2)2	0.38	0.50	0.8	0.444	190.5	0.02
local.diversity	lgm.climate	PC3	0.07	0.09	0.8	0.404	190.3	0.01
local.diversity	lgm.climate	poly(PC3, 2)1	0.42	0.50	0.8	0.404	191.6	0.01
local.diversity	lgm.climate	poly(PC3, 2)2	0.45	0.50	0.9	0.370	191.6	0.01
local.diversity	lgm.climate	PC4	0.08	0.09	0.9	0.386	190.2	0.01
local.diversity	lgm.climate	poly(PC4, 2)1	0.44	0.50	0.9	0.385	190.9	0.02
local.diversity	lgm.climate	poly(PC4, 2)2	-0.60	0.50	-1.2	0.229	190.9	0.02
local.diversity	wild	wild.5	0.25	0.09	3.0	0.004	182.5	0.06
local.diversity	wild	poly(wild.5, 2)1	1.43	0.49	2.9	0.004	184.6	0.06
local.diversity	wild	poly(wild.5, 2)2	-0.08	0.49	-0.2	0.866	184.6	0.06
local.diversity	wild	wild.10	0.28	0.09	3.2	0.002	180.8	0.08
local.diversity	wild	poly(wild.10, 2)1	1.56	0.48	3.2	0.002	181.9	0.08
local.diversity	wild	poly(wild.10, 2)2	-0.50	0.48	-1.0	0.306	181.9	0.08

Study variable	Driver type	predictors	Coef	SE	t value	P	AICc	R^2
local.diversity	wild	wild.20	0.25	0.09	2.9	0.004	182.5	0.06
local.diversity	wild	poly(wild.20, 2)1	1.43	0.49	2.9	0.004	184.4	0.07
local.diversity	wild	poly(wild.20, 2)2	-0.22	0.49	-0.4	0.660	184.4	0.07
local.diversity	veg.type	veg.type = grassl.	0.13	0.09	1.5	0.137	188.7	0.02
dark.diversity	abs.lat	abs.lat	-0.27	0.09	-3.2	0.002	181.0	0.08
dark.diversity	abs.lat	poly(abs.lat, 2)1	-1.54	0.48	-3.2	0.002	182.8	0.08
dark.diversity	abs.lat	poly(abs.lat, 2)2	0.28	0.48	0.6	0.568	182.8	0.08
dark.diversity	cur.biomes	cur.2.500	0.18	0.09	2.0	0.045	186.9	0.03
dark.diversity	cur.biomes	cur.2.1000	0.12	0.09	1.4	0.165	189.0	0.02
dark.diversity	cur.biomes	cur.2.2000	0.03	0.09	0.3	0.745	190.9	0.00
dark.diversity	cur.biomes	cur.14.500	0.23	0.09	2.6	0.010	184.3	0.05
dark.diversity	cur.biomes	cur.14.1000	0.23	0.09	2.7	0.009	184.0	0.05
dark.diversity	cur.biomes	cur.14.2000	0.20	0.09	2.3	0.023	185.7	0.04
dark.diversity	cur.biomes	cur.7.500	0.14	0.09	1.6	0.110	188.4	0.02
dark.diversity	cur.biomes	cur.7.1000	0.15	0.09	1.7	0.095	188.2	0.02
dark.diversity	cur.biomes	cur.7.2000	0.11	0.09	1.2	0.225	189.5	0.01
dark.diversity	cur.biomes	cur.1.500	0.19	0.09	2.1	0.034	186.4	0.04
dark.diversity	cur.biomes	cur.1.1000	0.16	0.09	1.8	0.072	187.7	0.03
dark.diversity	cur.biomes	cur.1.2000	0.11	0.09	1.2	0.237	189.6	0.01
dark.diversity	cur.biomes	cur.10.500	0.03	0.09	0.4	0.720	190.9	0.00
dark.diversity	lgm.biomes	lgm.12.500	0.12	0.09	1.4	0.162	189.0	0.02
dark.diversity	lgm.biomes	lgm.12.1000	0.10	0.09	1.2	0.253	189.7	0.01
dark.diversity	lgm.biomes	lgm.12.2000	0.05	0.09	0.5	0.600	190.7	0.00
dark.diversity	lgm.biomes	lgm.2.500	0.21	0.09	2.4	0.019	185.4	0.04
dark.diversity	lgm.biomes	lgm.2.1000	0.24	0.09	2.8	0.007	183.4	0.06
dark.diversity	lgm.biomes	lgm.2.2000	0.25	0.09	2.9	0.004	182.6	0.06
dark.diversity	lgm.biomes	lgm.1.500	0.25	0.09	2.8	0.005	183.1	0.06
dark.diversity	lgm.biomes	lgm.1.1000	0.20	0.09	2.3	0.022	185.7	0.04
dark.diversity	lgm.biomes	lgm.1.2000	0.13	0.09	1.5	0.137	188.7	0.02
dark.diversity	lgm.biomes	lgm.7.500	0.20	0.09	2.3	0.026	185.9	0.04
dark.diversity	lgm.biomes	lgm.7.1000	0.21	0.09	2.4	0.017	185.2	0.04
dark.diversity	lgm.biomes	lgm.7.2000	0.18	0.09	2.0	0.046	186.9	0.03
dark.diversity	lgm.biomes	lgm.3.500	0.22	0.09	2.5	0.015	184.9	0.05
dark.diversity	lgm.biomes	lgm.3.1000	0.19	0.09	2.2	0.033	186.4	0.04
dark.diversity	lgm.biomes	lgm.3.2000	0.17	0.09	2.0	0.050	187.1	0.03
dark.diversity	cur.climate	PC1	0.38	0.08	4.5	<0.001	171.6	0.14
dark.diversity	cur.climate	poly(PC1, 2)1	2.11	0.47	4.5	<0.001	173.7	0.14
dark.diversity	cur.climate	poly(PC1, 2)2	0.07	0.47	0.1	0.888	173.7	0.14
dark.diversity	cur.climate	PC2	0.20	0.09	2.4	0.020	185.5	0.04
dark.diversity	cur.climate	poly(PC2, 2)1	1.15	0.49	2.3	0.021	187.6	0.04
dark.diversity	cur.climate	poly(PC2, 2)2	-0.13	0.49	-0.3	0.796	187.6	0.04
dark.diversity	cur.climate	PC3	0.14	0.09	1.5	0.125	188.6	0.02
dark.diversity	cur.climate	poly(PC3, 2)1	0.77	0.48	1.6	0.114	181.5	0.09
dark.diversity	cur.climate	poly(PC3, 2)2	-1.47	0.48	-3.1	0.003	181.5	0.09
dark.diversity	cur.climate	PC4	0.10	0.09	1.1	0.265	189.7	0.01
dark.diversity	cur.climate	poly(PC4, 2)1	0.56	0.50	1.1	0.267	191.8	0.01
dark.diversity	cur.climate	poly(PC4, 2)2	-0.11	0.50	-0.2	0.826	191.8	0.01
dark.diversity	lgm.climate	PC1	0.33	0.08	3.9	<0.001	176.5	0.11
dark.diversity	lgm.climate	poly(PC1, 2)1	1.84	0.47	3.9	<0.001	174.5	0.14
dark.diversity	lgm.climate	poly(PC1, 2)2	0.95	0.47	2.0	0.045	174.5	0.14
dark.diversity	lgm.climate	PC2	0.20	0.09	2.3	0.023	185.7	0.04
dark.diversity	lgm.climate	poly(PC2, 2)1	1.13	0.49	2.3	0.023	186.3	0.05
dark.diversity	lgm.climate	poly(PC2, 2)2	-0.60	0.49	-1.2	0.226	186.3	0.05

Study variable	Driver type	predictors	Coef	SE	t value	P	AICc	R^2
dark.diversity	lgm.climate	PC3	0.02	0.09	0.3	0.783	190.9	0.00
dark.diversity	lgm.climate	poly(PC3, 2)1	0.14	0.48	0.3	0.775	182.0	0.08
dark.diversity	lgm.climate	poly(PC3, 2)2	-1.62	0.48	-3.4	0.001	182.0	0.08
dark.diversity	lgm.climate	PC4	-0.26	0.09	-3.0	0.003	182.3	0.07
dark.diversity	lgm.climate	poly(PC4, 2)1	-1.45	0.49	-3.0	0.004	184.4	0.07
dark.diversity	lgm.climate	poly(PC4, 2)2	0.03	0.49	0.1	0.948	184.4	0.07
dark.diversity	wild	wild.5	-0.07	0.09	-0.8	0.428	190.4	0.00
dark.diversity	wild	poly(wild.5, 2)1	-0.40	0.50	-0.8	0.430	192.4	0.01
dark.diversity	wild	poly(wild.5, 2)2	0.12	0.50	0.2	0.809	192.4	0.01
dark.diversity	wild	wild.10	-0.09	0.09	-1.0	0.325	190.0	0.01
dark.diversity	wild	poly(wild.10, 2)1	-0.49	0.50	-1.0	0.326	191.8	0.01
dark.diversity	wild	poly(wild.10, 2)2	0.27	0.50	0.5	0.595	191.8	0.01
dark.diversity	wild	wild.20	-0.01	0.09	-0.1	0.937	191.0	0.00
dark.diversity	wild	poly(wild.20, 2)1	-0.04	0.50	-0.1	0.937	193.1	0.00
dark.diversity	wild	poly(wild.20, 2)2	-0.12	0.50	-0.2	0.819	193.1	0.00
dark.diversity	veg.type	veg.type = grassl.	-0.22	0.09	-2.6	0.011	184.4	0.05
comm.compl.	abs.lat	abs.lat	-0.03	0.09	-0.4	0.723	190.9	0.00
comm.compl.	abs.lat	poly(abs.lat, 2)1	-0.18	0.50	-0.4	0.723	192.0	0.01
comm.compl.	abs.lat	poly(abs.lat, 2)2	0.49	0.50	1.0	0.328	192.0	0.01
comm.compl.	cur.biomes	cur.8.500	0.00	0.09	0.0	0.992	191.0	0.00
comm.compl.	cur.biomes	cur.8.1000	0.00	0.09	0.0	0.987	191.0	0.00
comm.compl.	cur.biomes	cur.8.2000	0.01	0.09	0.1	0.890	191.0	0.00
comm.compl.	cur.biomes	cur.13.500	0.07	0.09	0.8	0.437	190.4	0.00
comm.compl.	cur.biomes	cur.13.1000	0.05	0.09	0.5	0.599	190.7	0.00
comm.compl.	cur.biomes	cur.13.2000	0.05	0.09	0.5	0.587	190.7	0.00
comm.compl.	cur.biomes	cur.12.500	0.06	0.09	0.6	0.529	190.6	0.00
comm.compl.	cur.biomes	cur.12.1000	0.06	0.09	0.7	0.497	190.5	0.00
comm.compl.	cur.biomes	cur.12.2000	0.05	0.09	0.6	0.553	190.6	0.00
comm.compl.	cur.biomes	cur.2.500	0.03	0.09	0.3	0.745	190.9	0.00
comm.compl.	cur.biomes	cur.2.1000	0.05	0.09	0.6	0.559	190.6	0.00
comm.compl.	cur.biomes	cur.2.2000	0.08	0.09	0.9	0.398	190.3	0.01
comm.compl.	cur.biomes	cur.14.500	0.00	0.09	0.1	0.963	191.0	0.00
comm.compl.	cur.biomes	cur.7.500	0.11	0.09	1.3	0.212	189.4	0.01
comm.compl.	cur.biomes	cur.7.1000	0.12	0.09	1.3	0.195	189.3	0.01
comm.compl.	cur.biomes	cur.7.2000	0.15	0.09	1.8	0.083	187.9	0.02
comm.compl.	cur.biomes	cur.1.500	0.08	0.09	0.9	0.354	190.1	0.01
comm.compl.	cur.biomes	cur.1.1000	0.10	0.09	1.2	0.247	189.6	0.01
comm.compl.	cur.biomes	cur.1.2000	0.13	0.09	1.4	0.154	188.9	0.02
comm.compl.	cur.biomes	cur.10.500	0.20	0.09	2.2	0.027	186.0	0.04
comm.compl.	cur.biomes	cur.10.1000	0.22	0.09	2.5	0.012	184.6	0.05
comm.compl.	cur.biomes	cur.10.2000	0.23	0.09	2.6	0.009	184.1	0.05
comm.compl.	cur.biomes	cur.9.500	0.17	0.09	1.9	0.058	187.3	0.03
comm.compl.	cur.biomes	cur.9.1000	0.19	0.09	2.2	0.030	186.2	0.04
comm.compl.	cur.biomes	cur.9.2000	0.22	0.09	2.5	0.014	184.8	0.05
comm.compl.	lgm.biomes	lgm.11.500	0.02	0.09	0.2	0.821	190.9	0.00
comm.compl.	lgm.biomes	lgm.11.1000	0.04	0.09	0.4	0.688	190.8	0.00
comm.compl.	lgm.biomes	lgm.8.500	0.05	0.09	0.6	0.560	190.7	0.00
comm.compl.	lgm.biomes	lgm.8.1000	0.03	0.09	0.4	0.696	190.8	0.00
comm.compl.	lgm.biomes	lgm.8.2000	0.02	0.09	0.2	0.861	191.0	0.00
comm.compl.	lgm.biomes	lgm.13.500	0.22	0.09	2.5	0.014	184.9	0.05
comm.compl.	lgm.biomes	lgm.13.1000	0.24	0.09	2.8	0.006	183.1	0.06
comm.compl.	lgm.biomes	lgm.13.2000	0.24	0.09	2.8	0.006	183.2	0.06
comm.compl.	lgm.biomes	lgm.1.500	0.00	0.09	0.1	0.960	191.0	0.00

Study variable	Driver type	predictors	Coef	SE	t value	P	AICc	R^2
comm.compl.	lgm.biomes	lgm.1.1000	0.01	0.09	0.1	0.953	191.0	0.00
comm.compl.	lgm.biomes	lgm.7.500	0.12	0.09	1.4	0.170	189.1	0.01
comm.compl.	lgm.biomes	lgm.7.1000	0.13	0.09	1.5	0.132	188.7	0.02
comm.compl.	lgm.biomes	lgm.7.2000	0.16	0.09	1.8	0.079	187.8	0.02
comm.compl.	lgm.biomes	lgm.4.500	0.12	0.09	1.3	0.187	189.2	0.01
comm.compl.	lgm.biomes	lgm.4.1000	0.12	0.09	1.3	0.187	189.2	0.01
comm.compl.	lgm.biomes	lgm.4.2000	0.12	0.09	1.3	0.188	189.2	0.01
comm.compl.	cur.climate	PC1	-0.05	0.09	-0.6	0.544	190.6	0.00
comm.compl.	cur.climate	poly(PC1, 2)1	-0.30	0.50	-0.6	0.543	190.8	0.02
comm.compl.	cur.climate	poly(PC1, 2)2	0.69	0.50	1.4	0.167	190.8	0.02
comm.compl.	cur.climate	PC2	-0.17	0.09	-1.9	0.061	187.4	0.03
comm.compl.	cur.climate	poly(PC2, 2)1	-0.93	0.50	-1.9	0.062	189.5	0.03
comm.compl.	cur.climate	poly(PC2, 2)2	0.15	0.50	0.3	0.759	189.5	0.03
comm.compl.	cur.climate	PC3	-0.03	0.09	-0.3	0.752	190.9	0.00
comm.compl.	cur.climate	poly(PC3, 2)1	-0.16	0.50	-0.3	0.751	190.3	0.02
comm.compl.	cur.climate	poly(PC3, 2)2	0.82	0.50	1.6	0.103	190.3	0.02
comm.compl.	cur.climate	PC4	-0.20	0.09	-2.2	0.027	186.0	0.04
comm.compl.	cur.climate	poly(PC4, 2)1	-1.10	0.49	-2.2	0.027	187.3	0.04
comm.compl.	cur.climate	poly(PC4, 2)2	-0.44	0.49	-0.9	0.375	187.3	0.04
comm.compl.	lgm.climate	PC1	0.00	0.09	0.0	0.994	191.0	0.00
comm.compl.	lgm.climate	poly(PC1, 2)1	0.00	0.50	0.0	0.994	193.1	0.00
comm.compl.	lgm.climate	poly(PC1, 2)2	-0.05	0.50	-0.1	0.928	193.1	0.00
comm.compl.	lgm.climate	PC2	-0.17	0.09	-2.0	0.054	187.2	0.03
comm.compl.	lgm.climate	poly(PC2, 2)1	-0.96	0.49	-2.0	0.054	188.2	0.04
comm.compl.	lgm.climate	poly(PC2, 2)2	0.52	0.49	1.1	0.295	188.2	0.04
comm.compl.	lgm.climate	PC3	0.05	0.09	0.6	0.558	190.6	0.00
comm.compl.	lgm.climate	poly(PC3, 2)1	0.29	0.50	0.6	0.554	189.2	0.03
comm.compl.	lgm.climate	poly(PC3, 2)2	0.93	0.50	1.9	0.063	189.2	0.03
comm.compl.	lgm.climate	PC4	0.15	0.09	1.7	0.087	188.0	0.02
comm.compl.	lgm.climate	poly(PC4, 2)1	0.86	0.50	1.7	0.087	189.1	0.03
comm.compl.	lgm.climate	poly(PC4, 2)2	-0.50	0.50	-1.0	0.312	189.1	0.03
comm.compl.	wild	wild.5	0.23	0.09	2.7	0.009	183.9	0.05
comm.compl.	wild	poly(wild.5, 2)1	1.31	0.49	2.7	0.009	186.0	0.05
comm.compl.	wild	poly(wild.5, 2)2	-0.11	0.49	-0.2	0.823	186.0	0.05
comm.compl.	wild	wild.10	0.26	0.09	3.0	0.004	182.3	0.07
comm.compl.	wild	poly(wild.10, 2)1	1.44	0.49	3.0	0.004	183.4	0.07
comm.compl.	wild	poly(wild.10, 2)2	-0.50	0.49	-1.0	0.307	183.4	0.07
comm.compl.	wild	wild.20	0.21	0.09	2.4	0.018	185.3	0.04
comm.compl.	wild	poly(wild.20, 2)1	1.18	0.49	2.4	0.018	187.3	0.04
comm.compl.	wild	poly(wild.20, 2)2	-0.14	0.49	-0.3	0.784	187.3	0.04
comm.compl.	veg.type	veg.type = grassl.	0.19	0.09	2.1	0.036	186.5	0.03

Fig. S2. Uncertainty maps for predictions of AM fungal species pool size, local and dark diversity. Global predictions were made using random 80% subsets of the full data. This was repeated 100 times and uncertainty was calculated as the standard deviation of estimates derived from the different iterations.

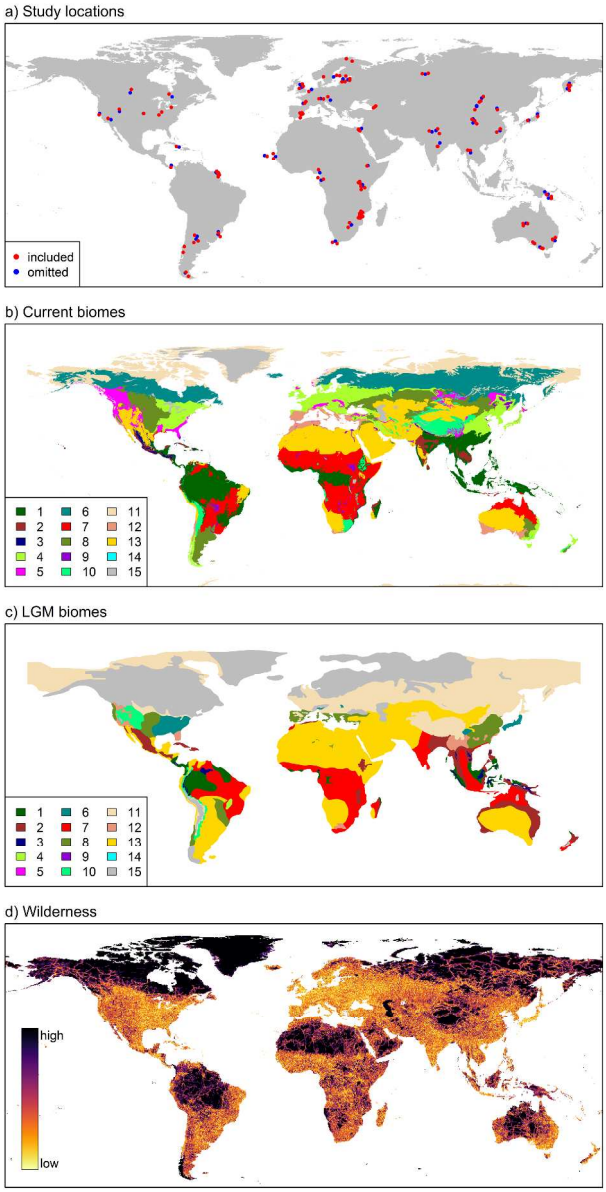
a) Study locations
included
omitted
b) Current biomes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
c) LGM biomes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
d) Wilderness
high
low

Fig 1 a, b, c, d

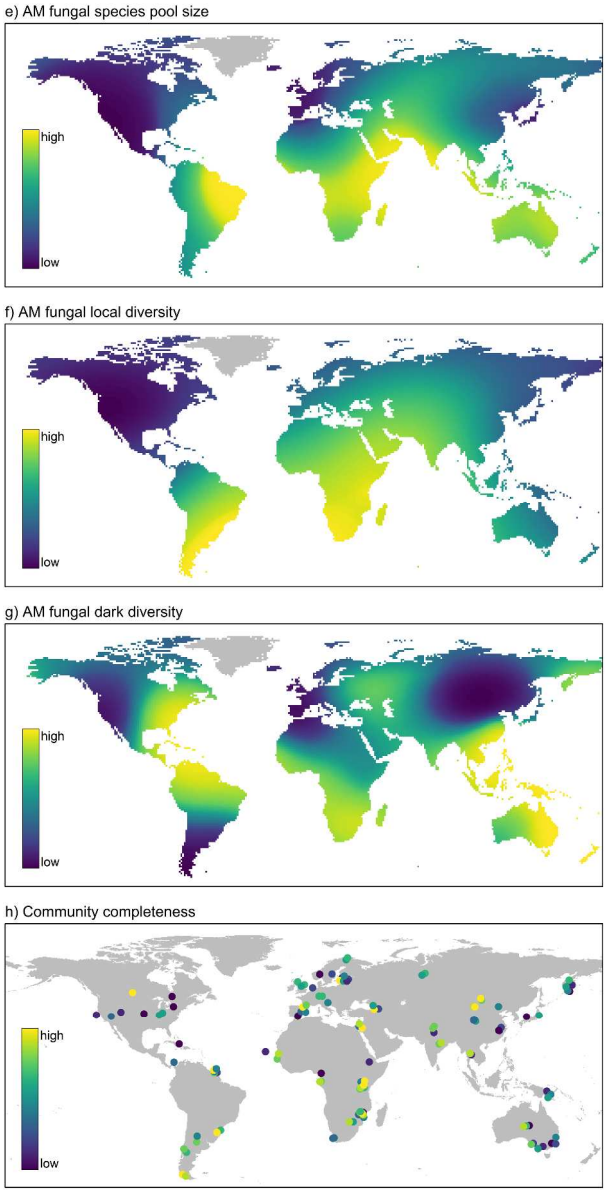
e) AM fungal species pool size
high
low
f) AM fungal local diversity
high
low
g) AM fungal dark diversity
high
low
h) Community completeness
high
low

Fig 1 e, f, g, h

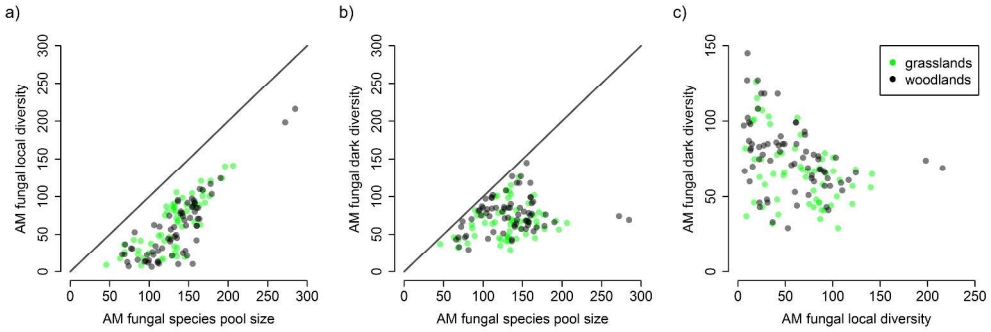

Fig. 2

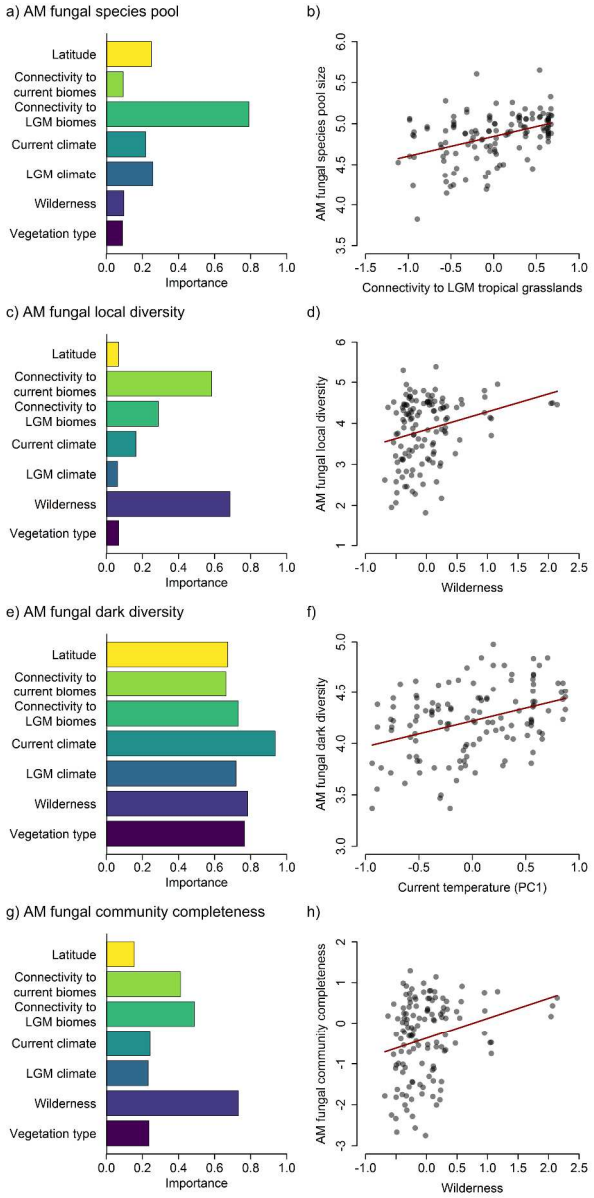
a) AM fungal species pool
Latitude
Connectivity to current biomes
Connectivity to LGM biomes
Current climate
LGM climate
Wilderness
Vegetation type
Importance
b)
AM fungal species pool size
Connectivity to LGM tropical grasslands
c) AM fungal local diversity
Latitude
Connectivity to current biomes
Connectivity to LGM biomes
Current climate
LGM climate
Wilderness
Vegetation type
Importance
d)
AM fungal local diversity
Wilderness
e) AM fungal dark diversity
Latitude
Connectivity to current biomes
Connectivity to LGM biomes
Current climate
LGM climate
Wilderness
Vegetation type
Importance
f)
AM fungal dark diversity
Current temperature (PC1)
g) AM fungal community completeness
Latitude
Connectivity to current biomes
Connectivity to LGM biomes
Current climate
LGM climate
Wilderness
Vegetation type
Importance
h)
AM fungal communiy completeness
Wilderness

Fig 3

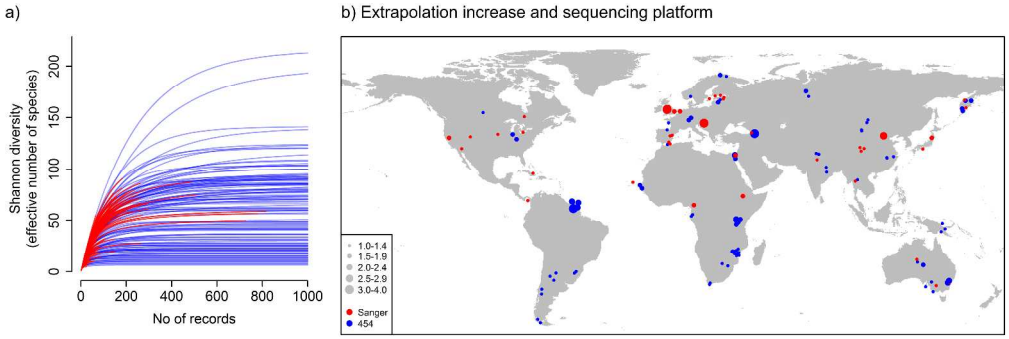

Fig. S1

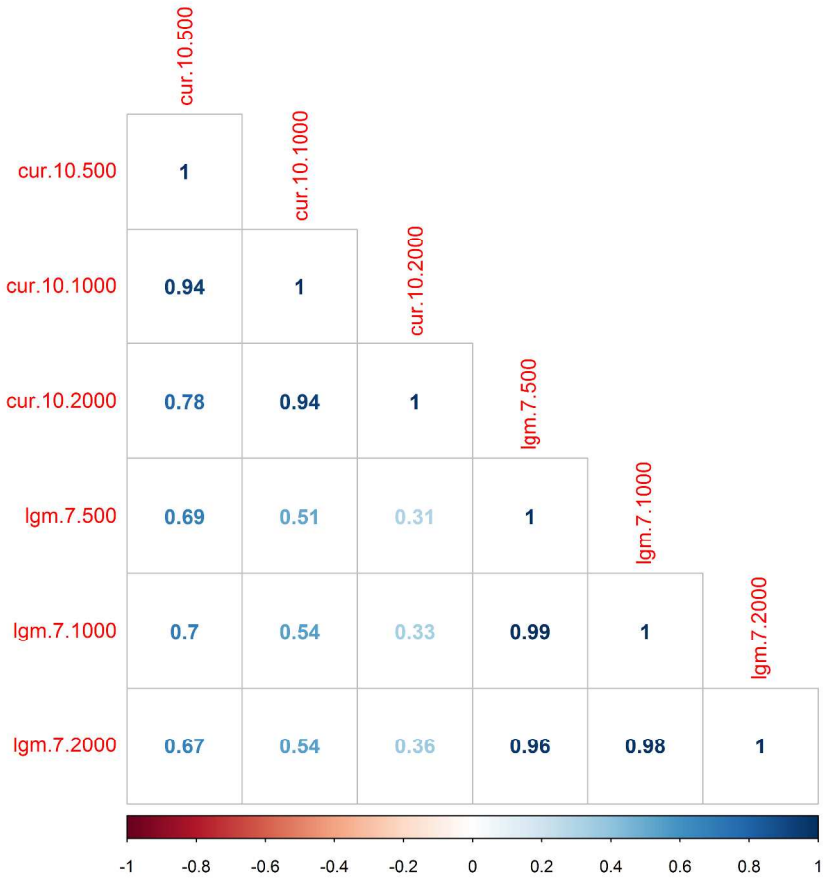
cur.10.500
cur.10.1000
cur.10.2000
lgm.7.500
lgm.7.1000
lgm.7.2000
1
0.94 1
0.78 0.94 1
0.69 0.51 0.31 1
0.7 0.54 0.33 0.99 1
0.67 0.54 0.36 0.96 0.98 1
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1

Table S4

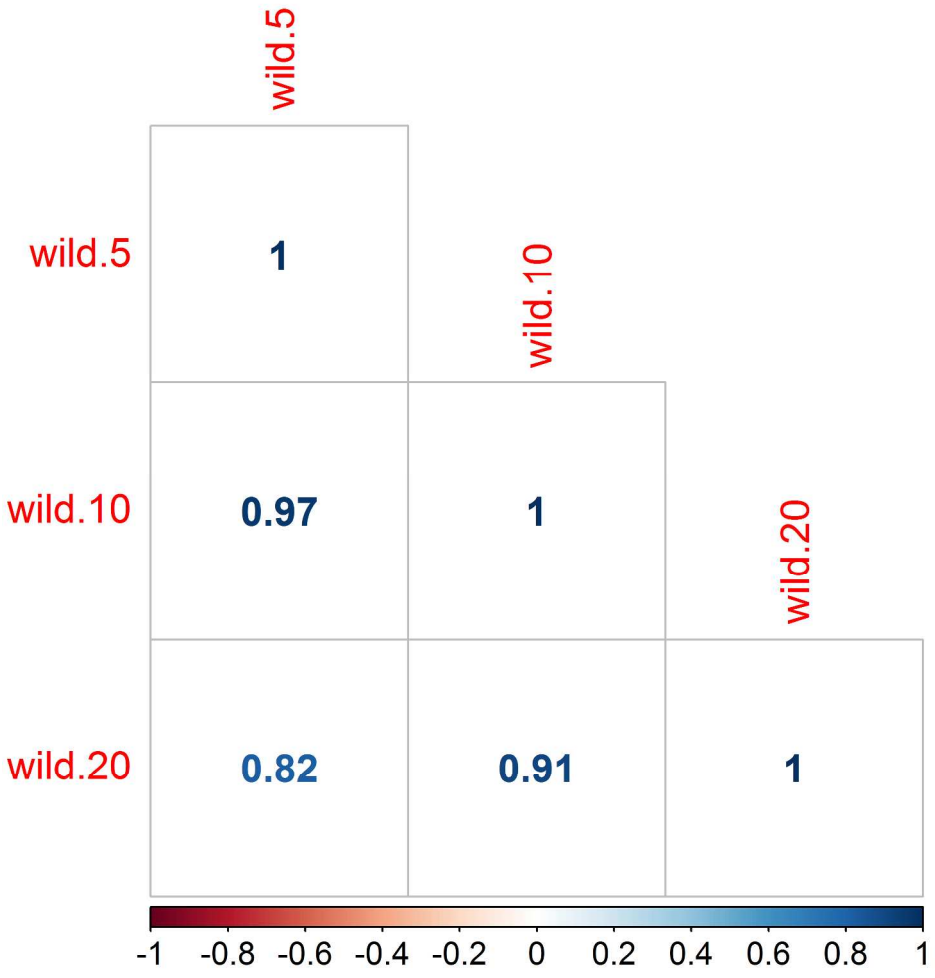
wild.5
wild.10
wild.20
wild.5
1
wild.10
0.97
1
wild.20
0.82
0.91
1
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

Table S5

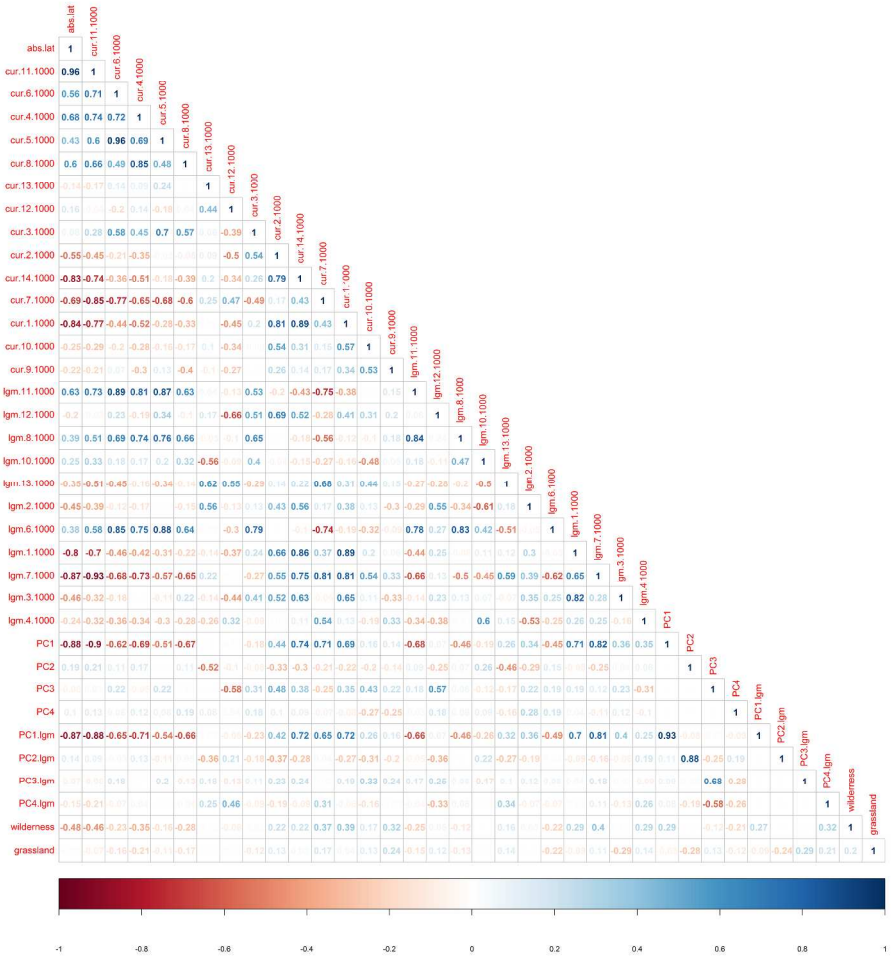
abs.lat
cur.11.1000
cur.6.1000
cur.4.1000
cur.5.1000
cur.8.1000
cur.13.1000
cur.12.1000
cur.3.1000
cur.2.1000
cur.14.1000
cur.7.1000
cur.1.1000
cur.10.1000
cur.9.1000
lgm.11.1000
lgm.12.1000
lgm.8.1000
lgm.10.1000
lgm.13.1000
lgm.2.1000
lgm.6.1000
lgm.1.1000
lgm.7.1000
lgm.3.1000
lgm.4.1000
PC1
PC2
PC3
PC4
PC1.lgm
PC2.lgm
PC3.lgm
PC4.lgm
wilderness
grassland
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

Table S6

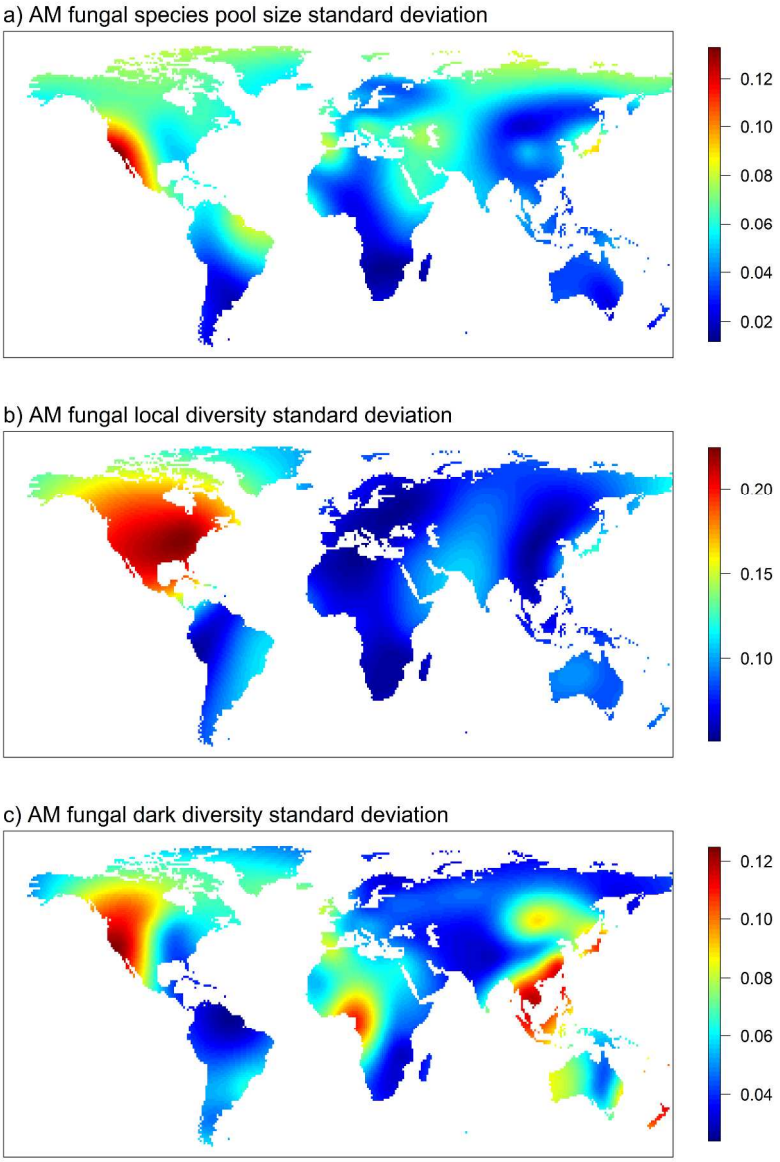
a) AM fungal species pool size standard deviation
0.12
0.10
0.08
0.06
0.04
0.02
b) AM fungal local diversity standard deviation
0.20
0.15
0.10
c) AM fungal dark diversity standard deviation
0.12
0.10
0.08
0.06
0.04

Fig S2