

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

Endemism and functional convergence across the North American soil mycobiome

### Permalink

<https://escholarship.org/uc/item/3439v1ww>

### Journal

Proceedings of the National Academy of Sciences of the United States of America,  
111(17)

### ISSN

0027-8424

### Authors

Talbot, Jennifer M  
Bruns, Thomas D  
Taylor, John W  
[et al.](#)

### Publication Date

2014-04-29

### DOI

10.1073/pnas.1402584111

Peer reviewed

# Endemism and functional convergence across the North American soil mycobiome

Jennifer M. Talbot<sup>a</sup>, Thomas D. Bruns<sup>b</sup>, John W. Taylor<sup>b</sup>, Dylan P. Smith<sup>a</sup>, Sara Branco<sup>b</sup>, Sydney I. Glassman<sup>b</sup>, Sonya Erlandson<sup>a</sup>, Rytas Vilgalys<sup>c</sup>, Hui-Ling Liao<sup>c</sup>, Matthew E. Smith<sup>d</sup>, and Kabir G. Peay<sup>a,1</sup>

<sup>a</sup>Department of Biology, Stanford University, Stanford, CA 94305; <sup>b</sup>Department of Plant and Microbial Biology, University of California, Berkeley, CA 94702; <sup>c</sup>Department of Biology, Duke University, Durham, NC 27708; and <sup>d</sup>Department of Plant Pathology, University of Florida, Gainesville, FL 32611

Edited by Susan E. Trumbore, Max Planck Institute for Biogeochemistry, Jena, Germany, and approved March 14, 2014 (received for review February 10, 2014)

**Identifying the ecological processes that structure communities and the consequences for ecosystem function is a central goal of ecology. The recognition that fungi, bacteria, and viruses control key ecosystem functions has made microbial communities a major focus of this field. Because many ecological processes are apparent only at particular spatial or temporal scales, a complete understanding of the linkages between microbial community, environment, and function requires analysis across a wide range of scales. Here, we map the biological and functional geography of soil fungi from local to continental scales and show that the principal ecological processes controlling community structure and function operate at different scales. Similar to plants or animals, most soil fungi are endemic to particular bioregions, suggesting that factors operating at large spatial scales, like dispersal limitation or climate, are the first-order determinants of fungal community structure in nature. By contrast, soil extracellular enzyme activity is highly convergent across bioregions and widely differing fungal communities. Instead, soil enzyme activity is correlated with local soil environment and distribution of fungal traits within the community. The lack of structure–function relationships for soil fungal communities at continental scales indicates a high degree of functional redundancy among fungal communities in global biogeochemical cycles.**

biogeochemistry | biogeography | forest | soil carbon | ectomycorrhizal fungi

The structure and function of ecological communities are intimately linked, such that the number and identity of species within a community often affect the key ecosystem properties of primary productivity (1), resistance and resilience to disturbance (2), and rates of nutrient cycling (3). However, understanding the extent to which structure–function relationships hold across communities and over large spatial scales continues to be a major goal of ecological research. Identifying these relationships for microbial organisms is particularly critical, because these organisms control rates of key ecosystem processes (the cycling of nitrogen, phosphorus, and carbon) (4) and directly affect the community structure of plants and animals through pathogenic or mutualistic interactions (5). As such, microbial activity is also intrinsic to Earth system models that inform citizens and policy makers of ecosystem dynamics and energy exchange between the biosphere and the atmosphere (6). As in plant communities of tropical rainforests (7), the incredible number of microbial taxa on Earth has been a challenge for understanding the link between diversity and function. Advances in DNA sequencing technology have recently allowed for a robust characterization of bacterial biogeographic patterns (8); however, to date, studies have examined structure–function relationships at a fixed scale (9–12). As a result, it is not yet clear how microbial function is linked to large-scale biogeographic patterns, whether or not this link is a more reliable determinant of microbial function in global biogeochemical cycles than other environmental factors, or how these relationships vary across geographic regions.

Certain ecological processes are only apparent or important at a particular scale (13), so a comprehensive understanding of microbial systems requires observation of community–environment–function interactions at multiple scales. A popular hypothetical framework that integrates spatial scale with ecological processes is a filter-type model (14, 15), where species pools are initially determined by processes at large spatial scales (16, 17), like evolutionary history or the presence of major dispersal barriers (dispersal filter). Environmental conditions then determine which species within the pool are able to colonize a particular habitat (environmental filter), and coexistence at smaller spatial scales is determined by niche differences and competition (biotic filter) (18, 19). Historically, a widely invoked assumption has been that the function of soil microbial communities is set primarily by the environmental filter, such that abiotic factors determine the physiology of whole microbial communities (20, 21). Recently, studies that have explicitly considered the composition and diversity of microbial communities at the local scale have also observed relationships between community composition and biochemical function of microbes in soils (9, 22), suggesting that the structure of microbial communities, per se, may be important. However, it is unclear if a link between community structure and function for microbial communities operates on a larger scale, independent of environment. If so, then knowing the structure of regional species pools will be important for understanding the function of microbial communities over large geographic regions. Alternatively, structure–function relationships may break down if there is high functional redundancy among the many microbial

## Significance

**Microbes control vital ecosystem processes like carbon storage and nutrient recycling. Although megadiversity is a hallmark of microbial communities in nature, we still do not know how microbial diversity determines ecosystem function. We addressed this issue by isolating different geographic and local processes hypothesized to shape fungal community composition and activity in pine forests across the continental United States. Although soil enzyme activity varied across soils according to resource availability, enzyme activity was similar across different fungal communities. These observations indicate that much of fungal diversity plays an equal role in soil biogeochemical cycles. However, soil fungal communities vary dramatically in space, indicating that individual species are endemic to bioregions within the North American continent.**

Author contributions: J.M.T., T.D.B., J.W.T., R.V., and K.G.P. designed research; J.M.T., T.D.B., J.W.T., D.P.S., S.B., S.I.G., S.E., R.V., H.-L.L., M.S., and K.G.P. performed research; J.M.T., T.D.B., D.P.S., and K.G.P. contributed new reagents/analytic tools; J.M.T., D.P.S., and K.G.P. analyzed data; and J.M.T. and K.G.P. wrote the paper.

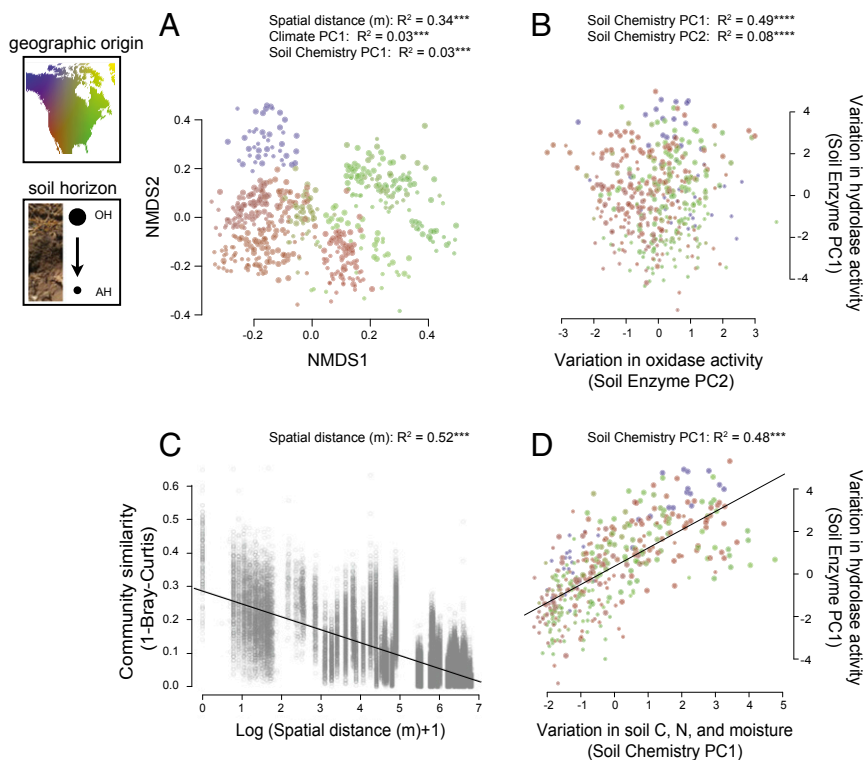
The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the NCBI Short Read Archive (accession no. [SRP040314](https://www.ncbi.nlm.nih.gov/short-read-archive/lookup/SRP040314)).

<sup>1</sup>To whom correspondence should be addressed. E-mail: [kpeay@stanford.edu](mailto:kpeay@stanford.edu).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1402584111/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1402584111/-DCSupplemental).



**Fig. 1.** Biogeography and functional geography maps of soil fungal community structure and function in pine biomes across North America. Patterns of fungal community composition (A) and soil enzymatic activity (B) are illustrated by coloring samples by geographic location (*Inset*). Symbol size indicates whether the sample originated from the organic (large circle) or mineral (small circle) soil horizon. Lines represent single-factor least-squares regression across all samples for factors explaining the most variation in fungal community composition (C) or activity of enzymes responsible for fast carbon and nutrient cycling (D). Statistics are derived from multiple regression analysis (A and B) or single-factor regression analysis (C and D). Asterisks represent significance of correlation (\*\*\*\* $P < 0.0001$ ; \*\*\* $P < 0.001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$ ). Points represent individual organic and mineral horizon samples collected from each sampling location at each plot in each field site (fungal community composition,  $n = 551$ ; enzyme activity,  $n = 253$ ).

taxa that exist, as suggested by observations that the rate of soil biogeochemical cycles saturates with the addition of many microbial species (23, 24).

To determine how microbial community structure and function are linked from the local to continental scale, we characterized soil fungal species composition, environmental conditions, and extracellular enzyme activity in over 600 soil samples taken from 25 plots spanning the geographic extent of the continental United States (Fig. S1). To focus our study on fungal-dominated soil communities, and to reduce confounding effects of dominant vegetation type that could drive cross-biome differences in fungal community structure–function relationships, we established plots in forests dominated by single host species from the plant family, Pinaceae (Table S1). To analyze turnover in fungal communities, environmental factors, and enzyme activity continuously from the local scale to the continental scale, we used a spatially explicit, nested design so that between-sample distances ranged over four orders of magnitude, from  $\sim 0.1$  m to 5,900 km. To decouple spatial and environmental factors, we took advantage of the strong changes in the soil chemical environment that occur vertically along the soil profile by separately identifying the fungi present in the organic and mineral horizons of each soil sample. To characterize community function, we measured the activity of extracellular enzymes involved in nutrient cycling. The activity of extracellular enzymes is a major aggregate function of decomposer communities that correlates with rates of carbon and nutrient cycling through dead organic matter (25). Whereas bacteria are a major component of the soil microbiome, fungal abundance and activity peak in low-pH (26), carbon-rich soils (27), as typified by temperate coniferous forests (Fig. S2). Under these conditions, fungi are the major producers of extracellular hydrolytic and oxidative enzymes (28).

## Results and Discussion

We found that within this single forest type, fungal community composition is delineated strongly by geographic regions within North America (Fig. 1A). This result shows that geographic endemism is a key feature of fungal communities, as seen with

plants and animals, and broadly parallels biogeographic provinces previously described for North America (29). As a consequence of geographic endemism, soil fungal communities displayed a significant distance–decay relationship spanning the meter to continental scale (Fig. 1C) that explained  $>50\%$  of total variation in species composition. This pattern is most consistent with a strong role for dispersal limitation as a driver of community turnover (8). Whereas climate, host plant identity, and local environment varied across our sites (Table S1) and can play an important role in structuring fungal communities (30), we found that only a small proportion of community variation was explained by these factors relative to spatial distance per se (Fig. 1A and Table S2). With respect to climate, the weak correlation with community composition is likely because a significant portion of the distance decay of fungal communities (Fig. 1C) occurs at the local (0–40 m) and landscape (1–100 km) scales over which climate is relatively invariant (Table S3). Variation in fungal community composition was saturated among samples that were taken from widely differing climates, indicating that fungal communities turn over in space much faster than climate. Thus, physiological and experimental approaches, rather than community-based analyses, will ultimately be necessary to determine how sensitive these communities are to climatic variables.

Other types of environmental variation had similarly small effects on fungal community structure. At the local scale, we sampled across steep soil chemical gradients (i.e., between the organic and mineral soil horizons) but found little consistent variation in fungal community composition (Fig. 1A). In fact, samples from different horizons in the same soil core often showed the highest degree of species overlap (Fig. 1C). Although vertical stratification of fungal communities is widely reported in the literature (31–33), the primary differences in fungal community composition generally occur between litter and soil, rather than between the two soil horizons (organic and mineral) that we sampled. Previous studies that have sampled across different forest types (34) or across fine scales at the local level (35) have clearly demonstrated the importance of environmental or habitat filtering of fungi. However, using our study design to

compare the role of local soil environment directly with other factors hypothesized to shape communities, we find that local environmental conditions play a secondary role to large-scale geographic processes in structuring fungal communities.

Rapid community turnover at local scales and high degrees of endemism at the largest scales are most consistent with dispersal as an important driver of local and regional species pools. Studies of airborne fungal spores show that dispersal can be a limiting factor at the meter to kilometer scale (36, 37). Similarly, evidence from population genetic (38) and phylogeographic (39) studies shows that large geographic features (e.g., oceans, mountains) are effective dispersal barriers to fungal populations. Limited dispersal in tandem with the strong priority effects that have been observed for fungi (40) may give rise to historically contingent fungal communities (41). Such historical contingencies may explain the high levels of community divergence we observed across plots within the same region.

Our results contrast with those of other studies of widely dispersing organisms for which environmental conditions can be the primary limits on community assembly (27, 42). Minimum winter temperatures and maximum summer temperatures constrain the migration and total population sizes of migrating birds, insects, and parasites (42), whereas pH is a strong predictor of soil bacterial community structure across biomes (43). This discrepancy may be the result of the poor dispersal ability of fungi, or the wider climatic tolerances of fungi compared with macroorganisms. For example, many ectomycorrhizal fungi persist in novel climates but first require introductions to establish (44). It is possible that host species range accounts for differences in ectomycorrhizal community structure across large geographic regions. However, strong host specificity of mycorrhizal fungi is not common below the plant family level (45). Collectively, these observations indicate that large-scale geographic processes like dispersal limitation are first-order determinants of both regional species pools and community composition of soil fungi at landscape scales. These results are consistent with the general filter-type models of community assembly often applied to plant and animal communities (16), which hypothesize that dispersal sets the regional species pools, whereas environmental factors and organismal traits define abundance of species at local scales.

In contrast to fungal communities, soil enzyme activity did not differ across regions of the continent (Fig. 1B). Instead, enzyme activity varied primarily at the local scale, correlating most strongly with soil chemistry. Activity of extracellular enzymes responsible for fast carbon and nutrient cycling (i.e., carbohydrases, phosphatases, chitinases, proteases) was best explained by availability of resources (e.g., carbon) to saprotrophs, which varied across soil horizons and accounted for 49% of the variation in enzyme activity (Fig. 1D, Table S2, and Dataset S1). Enzymes responsible for the slow release of carbon from recalcitrant plant and soil material (phenol oxidases and peroxidases) correlated with climate, soil moisture, and pH (Table S2 and Dataset S1), consistent with other studies on hydrolase and oxidase activity in bulk soil both within and across biomes (12, 21). Although saprotrophic bacteria can contribute to extracellular enzyme activity in bulk soil (46) and may be responsible for some of the activity we observed, the relative abundance of soil fungi is highest in temperate forest soils characterized by low pH and high carbon/nitrogen ratios, such as ours (27) (Fig. S2). Metaproteomic (28), metatranscriptomic (47) and metagenomic studies (48) show that in these systems, fungi are the dominant producers of extracellular enzymes involved in decomposition. Although aggregate measures, such as enzyme activity, cannot assign functions to individual organisms, our analysis showed that models best explaining the activity of both fast and slow carbon cycling enzymes included a measure of fungal functional trait composition. Specifically, peroxidase activity correlated positively with the ratio of ectomycorrhizal to saprotrophic diversity, whereas cellobiohydrolases showed the reverse trend, correlating positively with the ratio of saprotrophic to ectomycorrhizal

diversity (Dataset S1). These observations are consistent with the classical notion of saprotrophs as principal degraders of readily available carbon, as well as recent hypotheses that ectomycorrhizal fungi target nutrients bound in complex soil organic matter (49, 50). Because species-level composition of fungal communities does not correlate strongly with enzyme activity at the scale of this study, functional trait or phylogenetic approaches targeting key resource acquisition strategies may offer a better predictive framework for understanding fungal function across systems that have few species in common.

Our examination of fungal community structure and key ecosystem processes that are affected by fungi across a range of scales provides an important window into functional structure of highly diverse communities. Soils on different sides of the continent can have no fungal species in common and yet function very similarly with respect to the key enzymes driving biogeochemical cycles. Because fungi are major contributors to the decomposition process, this divergence of community structure and convergence of function demonstrate that much of fungal diversity in pine forests across North America may be functionally redundant from the perspective of nutrient cycling and decomposition. This functional redundancy is likely due to broad convergence in resource capture strategies across lineages of soil fungi. Common selection pressure has led to convergence on a core set of functional strategies across plant lineages (51) and leads to structural similarity in plant communities across different biogeographic regions. Similarly, the primary fungal trophic strategies are highly polyphyletic and distributed broadly across the fungal tree of life (52). For example, the ectomycorrhizal habit has evolved independently at least 66 times across the kingdom (53). Although most species appear to have restricted ranges in our study, the major fungal lineages are present in all of the geographic regions we surveyed (Table S4). Similar enzyme activities across regions are likely driven by selection for active taxa within these lineages by local gradients in resource availability. New efforts to represent microbes in biogeochemical models should therefore focus on trophic groups, rather than taxonomic diversity, to simulate short-term and long-term carbon storage on land. Our results lend empirical support to recent theoretical and empirical work showing that explicitly representing ectomycorrhizal and saprotrophic fungi in biogeochemical models alters simulations of soil carbon fluxes (54, 55) and will likely improve quantitative predictions of global carbon balance.

Our observation that the key drivers of fungal community composition and ecosystem function in soils operate at disparate scales presents a fundamental insight into the nature of these communities. Because most of the variation in decomposition activity can be found within a single forest, we can understand much of how fungal communities function by working at the local scale. Although predicting microbial function requires knowledge of local environmental conditions and richness estimates of key trophic groups, predicting the species composition of natural fungal communities first requires knowledge of geographical context, such as physical location, climate, and patch size. Regional endemism is a defining characteristic of soil fungal communities, with 85% of taxa in our study only found within a single bioregion (Fig. 2). Such endemism has implications for how society interacts with local natural ecosystems and can inform fields like microbial forensics, where reliable tracers of criminal activity across terrestrial habitats are critical (56). Endemism also allows for more complete descriptions of bioregions that include both aboveground and belowground biota (i.e., the *Pinus ponderosa-Rhizoscyphus* forest of the Sierra Nevada in California as distinct from the *P. ponderosa-Tricholoma* forest of the Cascade Range in Oregon). Conservation efforts to protect natural habitats may need to consider these belowground microbial communities, because endemic fungi risk extinction by habitat destruction and global climate change.

## Materials and Methods

**Soil Sampling.** To look at fungal community and functional turnover across a range of spatial scales, we used a hierarchical sampling design that allowed comparisons of multiple samples within a plot, across landscapes, and across the continent (Table S1). To do this while minimizing the known effects of vegetation biome types on microbial community structure and function (9, 57), we chose to focus on forests dominated by a single plant family, the Pinaceae (Table S1). The Pinaceae are ideal for exploring environment–community–function relationships across the fungal kingdom because they have a broad distribution across North America and show low levels of host specificity for mycorrhizal fungi within the family (47, 58). For example, North American pines readily associate with European ectomycorrhizal fungi (44), and co-occurring Pinaceae and angiosperms often share most common ectomycorrhizal fungi (59). Plots were chosen with the help of local experts to find mature stands with high dominance of a single Pinaceae species at sites spanning the North American continent (Fig. S1).

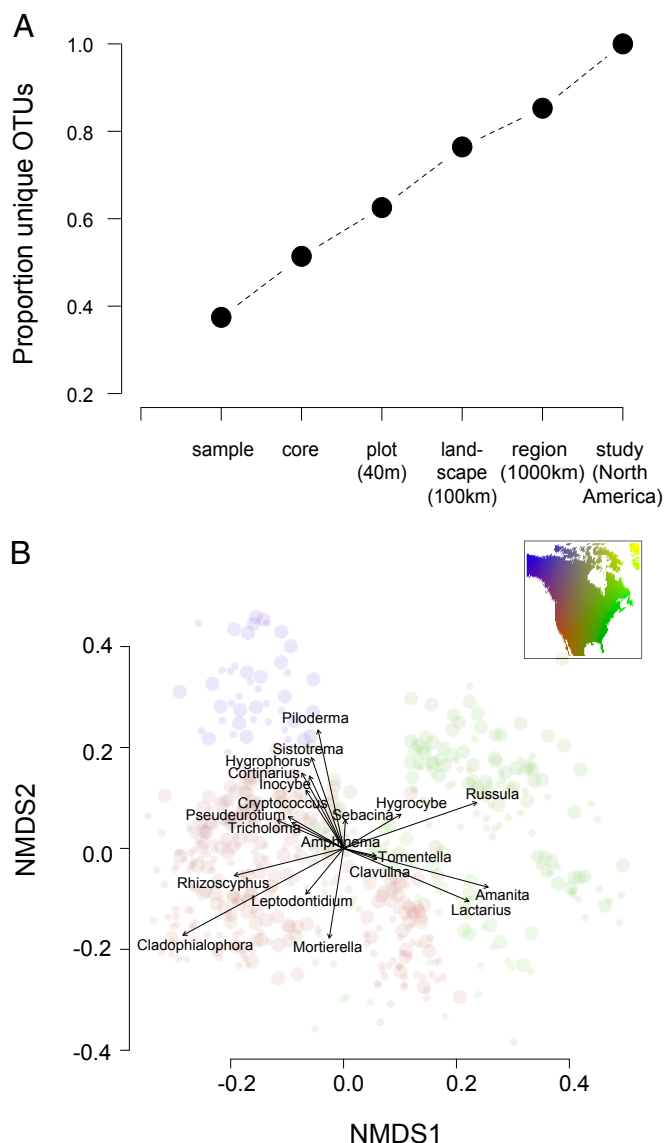
Sampling was carried out in 2011 and 2012 near the period of peak plant biomass production for a given region. In each plot, 13 soil cores were

collected from a 40-m  $\times$  40-m grid (Fig. S1). To look at turnover of community and function at the landscape scale, we ensured that each plot had at least one other plot located within a 1- to 50-km range (Fig. S1B). This design was chosen because it facilitates analysis of turnover at fine spatial scales by enabling multiple comparisons for each sampling distance. At each point in the plot, fresh litter was removed and a 14-cm deep, 7.6-cm diameter soil core was taken and immediately separated into a humic/organic horizon and mineral horizon. Our sampling protocol resulted in a total of 26 soil samples collected per plot (13 sample points  $\times$  two horizons). After removal, soils were kept on ice and transported to the nearest university within 8 h of collection. Soils were sieved through a 2-mm mesh to remove roots and rocks and homogenized by hand. A  $\sim$ 0.15- to 0.25-g subsample was placed directly into a bead tube from the Powersoil DNA Extraction Kit (MoBio), and the samples were stored at 4 °C until DNA extraction. Before extraction, samples were homogenized for 30 s at 75% power using a Mini-Beadbeater (BioSpec). A second subset was used for soil moisture measurements (described below), and a third subset was preserved for extracellular enzyme analysis and characterization of soil chemistry (frozen at  $-80$  °C). Soils were preserved within 48 h of field sampling.

**Fungal Community Composition.** To determine the fungal community composition in each soil sample, we used high-throughput sequencing of fungal DNA amplicons from each soil sample following the procedure of Talbot et al. (49). DNA was extracted from bulk soil and the internal transcribed spacer (ITS) region of the nuclear ribosomal RNA genes was amplified using the fungal-specific primers ITS1f and ITS4. We then identified the fungal community in each soil sample using 454-pyrosequencing of the ITS region with primer barcoding methods following the method of Talbot et al. (49). Individual PCR reactions were cleaned using an Agencourt Ampure XP kit (Beckman Coulter), quantified fluorescently with the Qubit dsDNA HS kit (Life Technologies), and pooled at equimolar concentration before pyrosequencing. Pyrosequencing was performed using a one-eighth or one-fourth run on the Roche Genome Sequencer FLX system at the Duke Institute for Genome Sciences and Policy. Sequence data were analyzed using the Quantitative Insights into Microbial Ecology (QIIME) pipeline (60). Initial sequence processing and sample assignment were done using the `split_libraries.py` command with a minimum/maximum sequence length cutoff of 350/1,200 bp, maximum homopolymer run length of 10 bp, and maximum barcode error number of 1.5. To account for the sequencing error that can arise as part of the PCR and 454-sequencing process, we used flowgram clustering (61) implemented in the `denoise_wrapper.py` command using a titanium error profile. Denoised sequences were chimera-checked and clustered into operational taxonomic units (OTUs) using the `usearch` option (62) in the `pick_otus.py` command. Chimera checking was implemented in `usearch` using both de novo and reference-based methods against the QIIME 12.11 alpha release of curated ITS sequences from the unified system for the DNA-based fungal species linked to the classification (UNITE) database. Only sequences identified as chimeric using both de novo and reference methods were considered chimeric and removed from the analysis.

Sequences were clustered into OTUs in `usearch` using a 97% sequence similarity threshold and a minimal cluster size of one. Taxonomy was assigned by searching representative sequences from each OTU against a previously published fungal ITS database (63), including well-curated sequences from the National Center for Biotechnology Information (NCBI) GenBank and UNITE using BLAST with a minimum expected value of 0.001 with the QIIME `assign_taxonomy.py` script. We chose to use this database instead of the QIIME 12.11 ITS database because the QIIME/UNITE database contained a large number of sequences with uninformative taxonomy. However, hand-checking of a number of the taxonomic assignments showed good concordance between identified taxa across the two databases. For the top 100, the two databases were consistent in 98 of 100 cases. The two disagreements were due to (i) an OTU with no close sequence matches in GenBank (closest match was 80%) and (ii) an error in GenBank that was incorporated into the QIIME/UNITE release. Of the 98 OTUs that were consistent, 45 showed an identical match at the genus level, 51 were not identified to genus (or above) using the UNITE database, and 2 had no BLAST hits with our database. Taxonomic assignments were consistent with GenBank sequences when manually compared using the BLASTn algorithm.

To look at the effects of fungal lifestyle on patterns of community and functional turnover, OTUs were categorized as ectomycorrhizal (ECM) or non-ECM based on current knowledge of metabolic lifestyle of the BLAST matches for each individual taxon (53). Inspection of the non-ECM OTUs shows that they are primarily decomposers of litter and woody plant material, although they also include some root endophytes and pathogens.



**Fig. 2.** Endemism of fungal taxa as shown by the proportion of OTUs unique to samples aggregated across a range of spatial scales (A) and genera (B) in pine forests across North America. (B) Ordination of the 20 most abundant genera across all samples is mapped onto Fig. 1A. (Inset) Samples are colored by geographic location in North America.

To investigate the reliability of 454-sequencing, we repeatedly sequenced independent PCR reactions from the same DNA extraction for four different samples (between three and five replicates per extraction). Samples were prepared and sequenced as above. Although there was variability between sequencing reactions for the same sample, we found that (i) quantitative patterns of taxon abundance were highly repeatable within the same sample and (ii) low-abundance taxa were often present in multiple samples, suggesting that they are not artifacts of the PCR or sequencing process. Based on these findings, we chose to retain low-abundance taxa in our analyses, because they are likely real, but to use the Bray–Curtis dissimilarity metric that weights abundant taxa more heavily. Although sequence abundance is not likely a perfect indication of relative biomass, these data suggest that it is still a better indicator of similarity between fungal communities than an incidence-based metric. In addition, results of community analyses using presence/absence data lead to essentially the same conclusions.

**Soil Chemistry.** Each soil horizon was analyzed for total carbon, total nitrogen, pH, total extractable ammonium and nitrate content, and percentage of soil moisture. Subsamples of fresh soil from each horizon at each sampling point in a plot were analyzed for soil water content by drying at 60 °C for 48 h. Frozen soils were thawed and analyzed for pH in a 1:1 water ratio using a glass electrode. To generate carbon/nitrogen ratios, total carbon and total nitrogen were analyzed using dry combustion on a Vario MAX CNS Elemental Analyzer (Elementar, Inc.) at the University of Minnesota and the Carlo Erba NA 1500 Elemental Analyzer at Stanford University. Total extractable ammonium and nitrate concentrations were analyzed in 2.0 M potassium chloride extracts of each soil sample using a WestCo SmartChem 200 discrete analyzer at Stanford University.

**Soil Enzyme Analyses.** We assayed the potential activities of seven extracellular enzymes involved in soil carbon and nutrient cycling: cellobiohydrolase (an exocellulase),  $\beta$ -glucosidase (which hydrolyzes cellobiose into glucose), polyphenol oxidase (which oxidizes phenols), peroxidase (including oxidases that degrade lignin), acid phosphatase (which releases inorganic phosphate from organic matter), *N*-acetyl-glucosaminidase (which breaks down chitin), and leucine-aminopeptidase (which breaks down polypeptides). Potential enzyme activities in bulk soil were measured separately for individual organic and mineral horizon samples using fluorometric and colorimetric procedures (64) on a microplate reader ( $n = 253$ ).

**Climate.** We obtained climate data for each sample using the WorldClim global climate dataset (65). For this purpose, we chose to use the Bioclim variables, which summarize monthly precipitation and temperature into 19 meaningful biological variables, such as temperature seasonality or precipitation of wettest month (Dataset S2). Although other datasets, such as Prism and Daymet, also provide high-resolution climate data, neither was available across all of our study sites. The latitude and longitude for each sample point were used to extract values for all Bioclim variables from North American raster layers with a 30-arc second resolution (~1 km). Bioclim tiles were downloaded, and data were extracted using the Raster package in R.

**Data Analysis and Statistics.** To determine the role of different spatial and environmental factors in determining structure and function of fungal communities, we collapsed environmental variables into vectors using principal components analysis (PCA). The percentages of soil moisture, total soil carbon, and total soil nitrogen were highly correlated, whereas pH was weakly correlated with soil carbon/nitrogen ratio and nitrate-nitrogen (Table S5). Ammonium and nitrate concentrations were not measured at all sites, but because ammonium correlated with percentage of soil nitrogen (Table S5), we chose to omit ammonium and nitrate from the soil chemistry PCA. After examining scree plots, we chose to retain the first two principal components, which explained 79.9% ( $PC1 = 56.1\%$ ,  $PC2 = 23.8\%$ ) of the variation in soil chemistry variables. Climate variables were also highly correlated and separated into three principal components that explained 90.5% ( $PC1 = 47.1\%$ ,  $PC2 = 32.3\%$ ,  $PC3 = 11.1\%$ ) of the variation in climate across sites (Dataset S2). Mantel tests were used to identify spatial autocorrelation in soil chemistry variables (across individual soil samples) and climate variables (across plots). The data reported in this paper are tabulated in Datasets S1 and S2.

**Factors Determining Fungal Community Composition Across All Samples.** To determine the factors controlling fungal community composition in soils across all samples, we used multiple regression on matrices (MRM) tests in the ecodist R package (66). Permutation tests were conducted with spatial

distance (meters), soil chemistry PC1 or PC2, or climate principal component axes as independent variables and with Bray–Curtis community dissimilarity among samples as the dependent variable. For comparative analyses among samples, samples were rarefied to 500 ITS reads ( $n = 551$ ). Bray–Curtis dissimilarity was based on the average of 10 different rarefactions. Community similarity using the Bray–Curtis abundance-based dissimilarity index was highly correlated with the incidence-based Jaccard index (Mantel test:  $r = 0.97$ ,  $P < 0.0001$ ). To determine the relative importance of geographic and local environmental factors in structuring communities, we then conducted multiple regression using MRM, including those variables that showed significant correlation with community composition in the univariate analyses and explained over 2% of variation in community dissimilarity. Stepwise model selection by Akaike's information criterion corrected for small sample sizes ( $AIC_c$ ) was used to determine factors retained in each multiple regression model. Models with the smallest  $AIC_c$  value are considered those best supported by the data. Patterns of community dissimilarity among samples were visualized with nonmetric multidimensional scaling (NMDS).

To visualize the role of geography in structuring soil fungal communities, color was assigned to each sample point based on location in North America following a modified version of the approach outlined by Krefth and Jetz (29). Basically, a color gradient was generated by assigning the colors red, yellow, blue, and green to four corners of a 2D plot and then interpolating across the remaining pixels in steps of 100. The color gradient was then projected onto a 101-cell  $\times$  101-cell raster map of the North American continent spanning 50° of latitude (range: 25–75°) and 100° of longitude (range: –150 to –50°). Each cell in the North American raster was assigned a corresponding value from the color gradient raster. The colors assigned to sample points in Fig. 1 are thus based on the color assigned to the corresponding geographic coordinate. This method provides a way of depicting geographic provenance without relying on a priori assumptions about fungal biogeographic realms.

**Factors Determining Soil Enzyme Activity Across All Samples.** Soil enzymes were highly correlated (Table S6); thus, to reduce the measured enzymes to a reasonable number of predictors, we used a PCA on the log-transformed variables. After examining scree plots, we chose to retain the first two principal components, which explained 70.6% ( $PC1 = 54.7\%$ ,  $PC2 = 15.9\%$ ) of the variation in soil enzymes. The first principal component was associated with variation in carbohydrate-, phosphorus-, and nitrogen-targeting enzymes, and the second was associated with polyphenolic (PPO, PER)-targeting enzymes (Table S6).

To determine the effect of fungal community, resource availability, and climate on enzyme activity, we first conducted univariate analyses with spatial distance, fungal community dissimilarity (Bray–Curtis), ECM species richness, saprotrophic species richness, soil chemistry PC1 or PC2, or climate PC1 or PC2 as independent variables and enzyme activity in bulk soil or enzyme principal components as independent variables. MRM analysis was used for fungal community composition and spatial distance only; otherwise, linear least-squares regression was used. To determine the relative importance of factors in structuring soil enzyme activity, we then conducted multiple regression analysis using only those variables that showed significant correlation with soil enzyme PC1 and PC2 in the univariate analyses and explained more than 2% of variation in soil enzyme activity. Stepwise model selection  $AIC_c$  was used to determine factors retained in each multiple regression model. For instances where climate predicted enzyme activity, we generated slope and  $P$  values using a linear mixed effect model, which adjusts degrees of freedom to account for multiple soil samples collected under the same climate conditions.

All statistical tests and graphics were done using the program R, version 2.7.2 (R Core Development Team, 2008). Distance matrices and NMDS ordination were performed with the package Vegan (67), and Mantel tests and MRM were performed with the package Ecodist. In cases where data did not conform to assumptions of normality and homogeneity of variance, values were log-transformed before analysis. All statistical tests were considered significant at  $P < 0.05$ .

**ACKNOWLEDGMENTS.** We thank T. Fukami, J. H. Martiny, R. L. Sinsabaugh, D. Leopold, R. Vannette, H. Moeller, M. Knope, E. J. Forrester, and A. Ravenscroft for feedback on the manuscript. We also thank D. L. Taylor (University of Alaska, Fairbanks), B. Becker (Point Reyes National Seashore), and J. Haas (Stanislaus National Forest) for logistical support and assistance with field collections. This project was funded by the National Science Foundation Dimensions of Biodiversity program (Awards 1045658, 1046115, and 1046052). J.M.T. was supported by the National Oceanic and Atmospheric Administration Climate and Global Change Postdoctoral Research Fellowship program.

1. Tilman D, et al. (1997) The influence of functional diversity and composition on ecosystem processes. *Science* 277(5330):1300–1302.
2. Cardinale BJ, et al. (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* 443(7114):989–992.
3. Strickland MS, Lauber C, Fierer N, Bradford MA (2009) Testing the functional significance of microbial community composition. *Ecology* 90(2):441–451.
4. Falkowski PG, Fenchel T, Delong EF (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* 320(5879):1034–1039.
5. Peay KG, Kennedy PG, Bruns TD (2008) Fungal community ecology: A hybrid beast with a molecular master. *Bioscience* 58(9):799–810.
6. Wieder WR, Bonan GB, Allison SD (2013) Global soil carbon projections are improved by modelling microbial processes. *Nat Clim Change* 3(10):909–912.
7. Hubbell SP (2005) Neutral theory in community ecology and the hypothesis of functional equivalence. *Funct Ecol* 19(1):166–172.
8. Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JB (2012) Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nat Rev Microbiol* 10(7):497–506.
9. Fierer N, et al. (2012) Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc Natl Acad Sci USA* 109(52):21390–21395.
10. Zhou J, Kang S, Schadt CW, Garten CT, Jr. (2008) Spatial scaling of functional gene diversity across various microbial taxa. *Proc Natl Acad Sci USA* 105(22):7768–7773.
11. Dinsdale EA, et al. (2008) Functional metagenomic profiling of nine biomes. *Nature* 452(7187):629–632.
12. Kivlin SN, Treseder KK (2014) Soil extracellular enzyme activities correspond with abiotic factors more than fungal community composition. *Biogeochemistry* 117: 23–37.
13. Levin SA (1992) The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. *Ecology* 73(6):1943–1967.
14. Martiny JBH, et al. (2006) Microbial biogeography: Putting microorganisms on the map. *Nat Rev Microbiol* 4(2):102–112.
15. Loreau M, et al. (2001) Biodiversity and ecosystem functioning: Current knowledge and future challenges. *Science* 294(5543):804–808.
16. Ricklefs RE (1987) Community diversity: Relative roles of local and regional processes. *Science* 235(4785):167–171.
17. Brown JH, Stevens GC, Kaufman DM (1996) The geographic range: Size, shape, boundaries, and internal structure. *Annu Rev Ecol Syst* 27:597–623.
18. Kraft NJB, Ackerly D (2010) Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. *Ecol Monogr* 80:401–422.
19. Fukami T (2010) Community assembly dynamics in space. *Community Ecology: Processes, Models, and Applications*, eds Verhoef HA, Morin PJ (Oxford Univ Press, Oxford), pp 45–54.
20. Sinsabaugh RL (2010) Phenol oxidase, peroxidase and organic matter dynamics of soil. *Soil Biol Biochem* 42(3):391–404.
21. Sinsabaugh RL, et al. (2008) Stoichiometry of soil enzyme activity at global scale. *Ecol Lett* 11(11):1252–1264.
22. McGuiire KL, et al. (2010) Functional diversity in resource use by fungi. *Ecology* 91(8): 2324–2332.
23. Setälä H, McLean MA (2004) Decomposition rate of organic substrates in relation to the species diversity of soil saprophytic fungi. *Oecologia* 139(1):98–107.
24. Nannipieri P, et al. (2003) Microbial diversity and soil functions. *Eur J Soil Sci* 54(4): 655–670.
25. Sinsabaugh RL (2005) Fungal enzymes at the community scale. *Mycology Series* (CRC, Boca Raton, FL), eds Dighton J, White JF, Jr., Oudemans P, Vol 23, pp 349–360.
26. Rousk J, Brookes PC, Bååth E (2009) Contrasting soil pH effects on fungal and bacterial growth suggest functional redundancy in carbon mineralization. *Appl Environ Microbiol* 75(6):1589–1596.
27. Fierer N, Strickland MS, Liptzin D, Bradford MA, Cleveland CC (2009) Global patterns in belowground communities. *Ecol Lett* 12(11):1238–1249.
28. Schneider T, et al. (2012) Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. *ISME J* 6(9):1749–1762.
29. Kreft H, Jetz W (2010) A framework for delineating biogeographical regions based on species distributions. *J Biogeogr* 37(11):2029–2053.
30. Sato H, Tsujino R, Kurita K, Yokoyama K, Agata K (2012) Modelling the global distribution of fungal species: New insights into microbial cosmopolitanism. *Mol Ecol* 21(22):5599–5612.
31. O'Brien HE, Parrent JL, Jackson JA, Moncalvo JM, Vilgalys R (2005) Fungal community analysis by large-scale sequencing of environmental samples. *Appl Environ Microbiol* 71(9):5544–5550.
32. Lindahl BD, et al. (2007) Spatial separation of litter decomposition and mycorrhizal nitrogen uptake in a boreal forest. *New Phytol* 173(3):611–620.
33. Vorišková J, Brabcová V, Cajthaml T, Baldrian P (2014) Seasonal dynamics of fungal communities in a temperate oak forest soil. *New Phytol* 201(1):269–278.
34. Edwards IP, Zak DR (2010) Phylogenetic similarity and structure of Agaricomycotina communities across a forested landscape. *Mol Ecol* 19(7):1469–1482.
35. Taylor DL, et al. (2014) A first comprehensive census of fungi in soil reveals both hyperdiversity and fine-scale niche partitioning. *Ecol Monogr* 84(1):3–20.
36. Peay KG, Schubert MG, Nguyen NH, Bruns TD (2012) Measuring ectomycorrhizal fungal dispersal: Macroecological patterns driven by microscopic propagules. *Mol Ecol* 21(16):4122–4136.
37. Peay KG, Garbelotto M, Bruns TD (2010) Evidence of dispersal limitation in soil microorganisms: Isolation reduces species richness on mycorrhizal tree islands. *Ecology* 91(12):3631–3640.
38. Amend A, Garbelotto M, Fang Z, Keeley S (2010) Isolation by landscape in populations of a prized edible mushroom *Tricholoma matsutake*. *Conserv Genet* 11(3):795–802.
39. Matheny PB, et al. (2009) Out of the Palaeotropics? Historical biogeography and diversification of the cosmopolitan ectomycorrhizal mushroom family *Inocybaceae*. *J Biogeogr* 36(4):577–592.
40. Kennedy P (2010) Ectomycorrhizal fungi and interspecific competition: Species interactions, community structure, coexistence mechanisms, and future research directions. *New Phytol* 187(4):895–910.
41. Fukami T, et al. (2010) Assembly history dictates ecosystem functioning: Evidence from wood decomposer communities. *Ecol Lett* 13(6):675–684.
42. Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annu Rev Ecol Syst* 37:637–669.
43. Lauber CL, Hamady M, Knight R, Fierer N (2009) Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl Environ Microbiol* 75(15):5111–5120.
44. Vellinga EC, Wolfe BE, Pringle A (2009) Global patterns of ectomycorrhizal introductions. *New Phytol* 181(4):960–973.
45. Ishida TA, Nara K, Hogetsu T (2007) Host effects on ectomycorrhizal fungal communities: Insight from eight host species in mixed conifer-broadleaf forests. *New Phytol* 174(2):430–440.
46. Rodríguez H, Fraga R (1999) Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnol Adv* 17(4-5):319–339.
47. Lindahl BD, Kuske CR (2014) Metagenomics for study of fungal ecology. *The Ecological Genomics of Fungi*, ed Martin F (John Wiley & Sons, Inc., Hoboken, NJ), pp 281–303.
48. Damon C, et al. (2012) Metatranscriptomics reveals the diversity of genes expressed by eukaryotes in forest soils. *PLoS ONE* 7(1):e28967.
49. Talbot JM, et al. (2013) Independent roles of ectomycorrhizal and saprotrophic communities in soil organic matter decomposition. *Soil Biol Biochem* 57(0):282–291.
50. Rineau F, et al. (2013) Carbon availability triggers the decomposition of plant litter and assimilation of nitrogen by an ectomycorrhizal fungus. *ISME J* 7(10):2010–2022.
51. Reich PB, Walters MB, Ellsworth DS (1997) From tropics to tundra: Global convergence in plant functioning. *Proc Natl Acad Sci USA* 94(25):13730–13734.
52. James TY, et al. (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443(7113):818–822.
53. Tedersoo L, May TW, Smith ME (2010) Ectomycorrhizal lifestyle in fungi: Global diversity, distribution, and evolution of phylogenetic lineages. *Mycorrhiza* 20(4): 217–263.
54. Averill C, Turner BL, Finzi AC (2014) Mycorrhiza-mediated competition between plants and decomposers drives soil carbon storage. *Nature* 505(7484):543–545.
55. Orwin KH, Kirschbaum MU, St John MG, Dickie IA (2011) Organic nutrient uptake by mycorrhizal fungi enhances ecosystem carbon storage: A model-based assessment. *Ecol Lett* 14(5):493–502.
56. Araujo R (2010) Endemism versus dispersion: Contribution of microbial genetics for forensic evidences. *Open Forensic Sci J* 3:14–21.
57. Tedersoo L, et al. (2012) Towards global patterns in the diversity and community structure of ectomycorrhizal fungi. *Mol Ecol* 21(17):4160–4170.
58. Rusca TA, Kennedy PG, Bruns TD (2006) The effect of different pine hosts on the sampling of *Rhizopogon* spore banks in five Eastern Sierra Nevada forests. *New Phytol* 170(3):551–560.
59. Kennedy PG, Izzo AD, Bruns TD (2003) There is high potential for the formation of common mycorrhizal networks between understorey and canopy trees in a mixed evergreen forest. *J Ecol* 91(6):1071–1080.
60. Caporaso JG, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7(5):335–336.
61. Knight R, Reeder J (2010) Rapid denoising of pyrosequencing amplicon data: exploiting the rank-abundance distribution. *Nat Methods* 7:668–669.
62. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19):2460–2461.
63. Peay KG, Baraloto C, Fine PV (2013) Strong coupling of plant and fungal community structure across western Amazonian rainforests. *ISME J* 7(9):1852–1861.
64. German DP, et al. (2011) Optimization of hydrolytic and oxidative enzyme methods for ecosystem studies. *Soil Biol Biochem* 43(7):1387–1397.
65. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25(15):1965–1978.
66. Goslee SC, Urban LD (2007) The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw* 22(7):1–19.
67. Oksanen JF, et al. (2012) vegan: Community Ecology Package. R package, Version 2.0-4. Available at <http://CRAN.R-project.org/package=vegan>.