Fungal Diversity in Permafrost and Tallgrass Prairie Soils under Experimental Warming Conditions

C. Ryan Penton, Derek StLouis, James R. Cole, Yiqi Luo, Liyou Wu, E. A. G Schuur, Jizhong Zhou and James M. Tiedje

Published Ahead of Print 6 September 2013.

Updated information and services can be found at:
http://aem.asm.org/content/79/22/7063

These include:

SUPPLEMENTAL MATERIAL
Supplemental material

REFERENCES
This article cites 78 articles, 15 of which can be accessed free at: http://aem.asm.org/content/79/22/7063#ref-list-1

CONTENT ALERTS
Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to another ASM Journal go to: http://journals.asm.org/site/subscriptions/
Fungal Diversity in Permafrost and Tallgrass Prairie Soils under Experimental Warming Conditions

C. Ryan Penton,a Derek St. Louis,a James R. Cole,a Yiqi Luo,b Liyou Wu,b,c E. A. G Schuur,d Jizhong Zhou,c,e,f James M. Tiedjea

Center for Microbial Ecology, Michigan State University, East Lansing, Michigan, USAa; Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma, USAa; Institute for Environmental Genomics, University of Oklahoma, Norman, Oklahoma, USAa; Department of Biology, University of Florida, Gainesville, Florida, USAa; State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, Chinaab; Earth Science Division, Lawrence Berkeley National Laboratory, Berkeley, California, USAa

Soil fungi play a major role in terrestrial ecosystem functioning through interactions with soil structure, plants, micro- and mesofauna, and nutrient cycling through predation, pathogenesis, mutualistic, and saprotrophic roles. The diversity of soil fungi was assessed by sequencing their 28S rRNA gene in Alaskan permafrost and Oklahoma tallgrass prairie soils at experimental sites where the effect of climate warming is under investigation. A total of 226,695 reads were classified into 1,063 genera, covering 62% of the reference data set. Using the Bayesian Classifier offered by the Ribosomal Database Project (RDP) with 50% bootstrap-classification confidence, approximately 70% of sequences were returned as “unclassified” at the genus level, although the majority (~65%) were classified at the class level, which provided insight into these lesser-known fungal lineages. Those unclassified at the genus level were subjected to BLAST analysis against the ARB-SILVA database, where ~50% most closely matched nonfungal taxa. Compared to the more abundant sequences, a higher proportion of rare operational taxonomic units (OTU) were successfully classified to genera at 50% bootstrap confidence, indicating that the fungal rare biosphere in these sites is not composed of sequencing artifacts. There was no significant effect after 1 year of warming on the fungal community structure at both sites, except perhaps for a few minor members, but there was a significant effect of sample depth in the permafrost soils. Despite overall significant community structure differences driven by variations in OTU dominance, the prairie and permafrost soils shared 90% and 63% of all fungal sequences, respectively, indicating a fungal “seed bank” common between both sites.

Fungi represent a large, still incompletely discovered, repository of biological diversity (1) whose role is crucial in ecosystem functioning. By directly influencing the food web as plant pathogens, predators of micro- and mesofauna, mutualists, and decomposers, fungi play a key role in soil function (1). As one of the most diverse groups of the Eukarya, only a small fraction of the estimated total fungal diversity (2) has been assessed to date. Fungal communities are influenced by spatiotemporal as well as biotic and abiotic factors. For example, edaphic characteristics such as soil type (3), substrate composition and pH (4, 5, 6), as well as seasonality (7, 8, 9), have been found to shape community composition. Plant community composition may also play an important role (8, 10) through a direct influence on endophytes, pathogens, and mycorrhizal fungi or on saprotrophs through differences in litter quality and source.

The diversity of the fungal component in microbial communities has become more accessible through high-throughput marker gene sequencing. Among the primer sets proposed for the amplification of the 28S gene, the LR3/LR0R primer combination that spans D1/D2 is suitable for both classification accuracy and resolution to the genus level (11). While 28S gene sequence assignment has typically been performed using iterations of BLAST (12, 13), a fungal 28S naive Bayesian classifier (NBC) was recently made available at the Ribosomal Database Project (RDP) that allows for direct, rapid sequence classification without clustering. The classifier has been independently validated (14) and is based on a manually archived large subunit (LSU) gene training set (11). An advantage of this fungal 28S classifier is the ability to select bootstrap values from 0% to 95%, which indicates the number of times a genus is selected out of 100 bootstrap trials using a subset of one-eighth of the eight-character subsequences (words) comprising each query sequence. Hence, higher bootstrap values result in a larger proportion of sequences binned as “unclassified.” These unclassified bins occur at various taxonomic ranks; an unclassified sequence at the genus level may remain classified at coarser taxonomic levels. This group can be further investigated by examining the composition at different taxonomic levels to give more insight into the lesser-known fungal lineages, including their distribution relative to environmental attributes.

With the exception of a few deep sequencing studies (7, 10, 15, 16), most have not addressed the extent or composition of the unclassified members other than establishing them as nonfungal reads. The definition of unclassified reads is largely dependent on the classification method. In most internal transcribed spacer (ITS)-based studies where BLAST or phylogenetic relationships inferred through trees based on alignments are used, an unclassified bin or cluster would result from a closest sequence match to an unclassified or uncultured sequence or no matches above a predetermined similarity or E value. The abundance of fungal unclassified reads, even at higher subphylum or class levels (7),
illustrates the dearth of sequences available from classified fungi. As such, the composition of any fungal database precludes the identification of the majority of sequences (15, 17, 18). As Lentendu et al. (15) pointed out, the effect of the definition “unclassified” on fungal diversity estimates has been largely ignored.

In this study, we used 28S gene-based pyrosequencing coupled with direct classification using the RDP-naive Bayesian fungal classifier of fungal communities from Alaskan permafrost samples (AK samples) and Oklahoma tallgrass prairie samples (OK samples), both of which have plots at natural temperatures and warmed 1.5 to 2°C above the ambient temperature to simulate global warming effects. Our objectives were (i) to gain more insight into the difficult to classify taxa by using the classifier’s ability to resolve these sequences at different taxonomic levels, (ii) to determine if short-term warming resulted in a shift in community structure, (iii) to assess if depth in the Alaska soil (through the active layer and into permafrost) affected community structure, and (iv) to establish what proportion of the known fungal taxa resided in these two very different ecosystem types and which taxa they had in common.

**MATERIALS AND METHODS**

**Site descriptions.** A permafrost warming experiment, the Carbon in Permafrost Experimental Heating Research project (CiPEHR), is located in the northern foothills of the Alaska Range near Denali National Park and Preserve in the region of Eight Mile Lake, Alaska (63°52’59”N, 149°13’32”W). Established in September 2008, the goal of the study site is to determine the effects of soil, air, and permafrost warming on tundra ecosystems (19). The site lies within a region of discontinuous permafrost where thawing and thermokarst (areas with standing water due to permafrost thaw and resulting soil depression) have been occurring over several decades. Winter soil temperatures were passively increased by accumulation of snow with snow fences in the winter and removal of snow in the spring. Warming has increased the soil temperature by 1.5°C (integrated over a depth of 5 to 40 cm) (19) and melted the upper permafrost layers.

Soils are classified in the soil order Gelsol, and vegetation is dominated by the deciduous shrub Vaccinium uliginosum and the tussock-forming sedge Eriophorum vaginatum. Other common vascular plants include Betula nana and Carex bigelowii, with the nonvascular Sphagnum spp., feather moss, and several lichen species also present (19, 20, 21, 22). Other common vascular plants include Carex bigelowii, with the nonvascular Sphagnum spp., feather moss, and several lichen species also present. The Oklahoma samples originated from the unclipped (no biomass harvest) plots at the Great Plain Apiaries site (34°58’54”N, 97°31’14”W). The study site was implemented in November 1999 to characterize soil responses to experimental warming (22, 23, 24, 25, 26, 27). Soils are classified as silt loams, and the grassland is dominated by the C3 grasses Schizachyrium scoparium, Sorghastrum nutans, and Eragrostis curvula and the C4 forbs Ambrosia psilostachya and Xanthocephalum texanum. Plots were warmed with an infrared heater, while the control plot contained a “dummy” heater to simulate the shading effect. Experimental warming increased mean soil temperatures by 1.8 to 2.7°C. Soil cores were collected in October 2010 from the topsoil (0 to 15 cm) of six replicates in both warming and control plots.

All samples from both sites were kept frozen until DNA extraction. Plant roots were not removed. DNA was extracted and purified using the freeze-grinding mechanical lysis method (28).

**454 pyrosequencing.** Amplification was performed using fungal 28S gene primers LR3/LR0R (<http://www.biology.duke.edu/fungi/mycolab/primers.htm>), which results in 625-bp fragments (11). Bar codes were included on both the forward and reverse primers to facilitate sample discrimination. Quadruplicate 20-μl PCRs were performed with the following materials: 4 μl Promega GoTaq buffer, 0.5 μl GoTaq DNA polymerase, 1.5 μl Roche 25 mM MgCl₂, 1 μl Invitrogen 10 mM deoxynucleo-side triphosphate (dNTP) mix, 1 μl each primer (10 pmol μl⁻¹), 0.2 μl New England Biolabs 10 mg ml⁻¹ bovine serum albumin (BSA), 1 μl (10 ng μl⁻¹) template, and 9.8 μl H₂O. The cycling conditions were an initial denaturation of 94°C for 3 min followed by 30 cycles of 94°C for 1 min, 51°C for 40 s, and 72°C for 1 min, followed by an extension at 72°C for 10 min. Replicates were pooled and gel purified using the Qiagen gel purification kit following the manufacturer’s protocol. Products were further purified using the Qiagen PCR purification kit. Following adapter ligation, amplicons were sequenced by Macrogen (Seoul, South Korea) using Lib-L kits and processed using the shotgun protocol.

Raw sequences read from both the A and B adapters were quality processed and sorted through the RDP pyrosequencing pipeline (<http://pyro.cme.msu.edu>). The following filter parameters were used: 0 mismatches to forward primer, 400-bp-length filter, ambiguous base N count = 0, and quality score (Q) ≥ 20. The sequences that fully matched the bar codes were selected and then subjected to the RDP fungal LSU classifier (<http://rdp.cme.msu.edu/classifier>) using training set 1 at both 0% (best match) and 50% bootstrap confidence levels. In both cases, nonfungal and Eukaryota incertae sedis reads were removed for downstream statistical analyses. For the 50% bootstrap confidence classification, taxa were identified as “unclassified” at a particular taxonomic level when the classification confidence was <50%. This confidence level was chosen because it preceded large incremental increases in the percentage of unclassified taxa at higher confidence levels. The abundances of sequences belonging to taxa classified at the phylum, class, order, family, and genus levels were treated as bins for downstream statistical analysis. Approximately 1.7% of all sequences were identified as chloramycetes based on UCHIME (29). Putative chimeras were not removed, as this frequency falls within the false-positive detection rate (30).

**Data analysis.** Raw abundances were normalized by Hellinger transformation (square root of relative abundance), and a Bray-Curtis dissimilarity matrix was constructed for nonmetric multidimensional scaling (NMDS). Significant community differences were tested using analysis of similarity (ANOSIM) (31) and permutational multivariate analysis of variance (PERMANOVA) (32). Cluster analyses were performed with the similarity profile analysis (SIMPROF) test (33), and similarity percentages (SIMPER) analysis (34) was used to identify the operational taxonomic unit (OTU) that contributed to the discrimination among treatments. Sample dispersion was tested by permutational analysis of multivariate dispersions (PERMDISP) (35) using PRIMER-E (36). Significant differences in species richness (Margalef; d), Shannon diversity (H’), and Pielou’s evenness (J’) were assessed with one-way analysis of variance (ANOVA) using StatPlus 5.8.2.0 (AnalystSoft, Inc.).

Random resampling was performed to 5,180 Oklahoma and 3,522 Alaska sequences per sample for downstream analyses. For direct comparison between sites, 3,522 sequences were randomly resampled from both data sets.

**Nucleotide sequence accession numbers.** Nucleotide sequences have been deposited in the European Nucleotide Archive (<http://www.ebi.ac.uk/ena/>) under study no. PRJEB3968 with accession no. ERS250779 to ER250807.

**RESULTS**

**Oklahoma site.** From a total of 147,434 sequences subjected to filtering, the Oklahoma (OK) samples yielded a total of 96,669 best-match (0% bootstrap confidence) classified sequences containing 936 genera in 276 families, 96 orders, 31 classes, and 7 phyla. A total of 130 genus bins were singletons. Approximately 95% of the reads were successfully classified as Fungi or Eukaryota incertae sedis. Rarefaction curves were not saturated (see Fig. S1A in the supplemental material). Communities were dominated by the fungal classes Agaricomycetes, Sordariomycetes, Lecanoro-
mycetes, and Glomeromycetes. The top 10 and 20 genera accounted for 36.3% and 46.9% of all sequences, respectively, as evidenced by the species accumulation curve (Fig. 1A). The OK fungal communities were dominated by the Agaricales (Moniliophthora, Leucopaxillus, Camarophyllus, and Camarophyllopsis), the Hysteriales (Hysteropatella), Chaothryiales (Sorocybe), Onygenales (Arachnomycetes), Ascomycota incertae sedis (Everhartia), Lecanorales (Boreoplacea), Pezizales (Phialea), Schizosaccharomycetales (Schizosaccharomyces), and the Sordiales (Cephalora) (Fig. 2), while two predominant soil molds of warm prairie soils, Penicillum and Aspergillus, constituted <0.1% of all sequences. There was no consistent difference among genus-level fungal communities between the control and warming treatments according to NMDS ordination (see Fig. S2 in the supplemental material) and as tested through ANOSIM (P = 0.64) and PERMANOVA (P = 0.66). This was consistent among all taxonomic levels. Only 10 mostly lower-abundance genera changed significantly with treatment (t test, P < 0.05).

**Alaska site.** Of a total of 153,403 sequences, the Alaska (AK) samples generated a total of 130,026 classified sequences at best match, containing 743 genera in 276 families, 101 orders, 33 classes, and 9 phyla. Rarefaction curves showed higher coverage than OK samples, with saturation approached in many of the deeper soil samples (see Fig. S1B in the supplemental material). A total of 157 genera were singletons. The distribution of genera revealed a dominant community, with the top 10 and 20 genera accounting for 54.7% and 69.9% of all sequences, respectively (Fig. 1B). Fungal communities were dominated by the Helotiales genera Godronia, Chloroscypha, Hyaloscypha, Hyalodendriella, Sclerotinia, Lachnum, Cinula, Vibliasea, Ascoscypha, and Fulvoflamma, as well as Chaothryiales (Sorocybe), Saccharomycetales (Dipodascus), Russulales (Russula), Polyporales (Abortiporus), Chytridiales (Karlovyomyces), and Zoogopagales (Kuzuharae), which comprised an average of 62.2% of all sequences.

There was no significant effect of warming (PERMANOVA and ANOSIM), with only seven significantly different genera (t test, P < 0.05) that comprised 5.9 and 14.9% of all sequences at the control and warming sites. However, there was a significant difference in fungal community structure with depth by parsing of samples between 0 and 20 cm (shallow) from the 20- to 50-cm (deep) samples (ANOSIM, global R = 0.42, P < 0.01; PERMANOVA, F = 4.53, P < 0.01) (see Fig. S2B in the supplemental material). Differences in fungal community structures between the shallow and deep samples were present from the class to genus taxonomic levels in both relative abundance and presence or absence (see Fig. S3 and S4 in the supplemental material). While both depth increments were dominated by the fungal class Leotiomycetes, the identity and abundance of the lesser classes changed with depth. SIMPER analysis for this comparison was weighted heavily toward a few genera, with the top 10 comprising 52.7% of the dissimilarity and 53.1% of all sequences (Fig. 3). In all, 18 genera were significantly different (t test, P < 0.05) between the shallow and deeper layers, generally corresponding to the more abundant taxa in each layer.

**Composition of the unclassified taxa.** Classification was performed for the entire data set at a 50% bootstrap confidence level to gain insight into the composition of the underrepresented taxa within the current database. For the OK samples, at this confidence, “unclassified” taxa ranged from 8.4% (kingdom) to 67.1% (genus) of all sequences, depending on the taxonomic level and were contained within 268 family-level bins (Fig. 4A), leaving 390 classified genera, a reduction of the original genus-level diversity by 68.3%. At the phylum level, an average of 32.9% of sequences were unclassified, with the majority (25.1%) most closely assigned to the Ascomycota. At the order level, 47.2% were unclassified, with the majority belonging to the Sordariomycetes (6.0%), Dothideomycetes (7.4%), Eurotiomycetes (6.2%), and Pezizomycetes (4.9%). At the genus level, unclassified sequences were heavily weighted toward relatively few taxa (see Fig. S5 in the supplemental material). Both NMDS ordination and ANOSIM results were largely unaffected when unclassified bins were included for the Oklahoma data (data not shown). Overall, the abundances of the unclassified genera between the AK shallow, AK deep, and OK samples at all taxonomic levels were significantly different (two-way ANOVA, P < 0.05).

In the Alaska data set, unclassified taxa ranged from an average of 36.8% to 88.1% of sequences from the phylum to genus level (Fig. 4B). At the order level, 38.9% were unclassified—the majority in the Chytridiomycetes (19.5%), Fungi incertae sedis (16.9%), Leotiomycetes (9.4%), and Sordariomycetes (8.5%). A total of 241 genera were placed into unclassified bins at the corresponding family level, which were again heavily weighted toward a few taxa (see Fig. S6 in the supplemental material). This resulted in a 67.6% reduction of classified genera compared to the original data set. Of
the top 30 unclassified family bins, only 7 were shared between OK and AK samples. The majority of the top 20 genera obtained by best match remained prevalent, although at lower abundances due to binning into members of associated unclassified families. A higher proportion of unclassified Fungi at the family level were associated with the deeper permafrost samples (27.8% versus 16.6%).

**Community comparison between OK and AK samples.** Not surprisingly, fungal communities were fundamentally different between the Alaska permafrost and Oklahoma grassland samples. The OK samples exhibited significantly higher diversity and evenness than the AK shallow (0- to 20-cm) and deep (20- to 65-cm) samples (ANOVA, \( P < 0.01 \)) (Fig. 5). Significant differences in abundance between the sites were most pronounced at the higher taxonomic levels (\( t \) test, \( P < 0.05 \)) (see Tables S1 and S2 in the supplemental material) when all depths were included. Of the top 50 most abundant genera at 0% bootstrap confidence from both sites at all depths comprising 56% (AK) and 76% (OK) of all sequences, 11 genera were shared. A total of 105 genera (~10% of sequences, 49 singletons) were unique to the OK samples, while 503 genera (~37% of sequences, 320 singletons) were unique to AK. Both NMDS for shallow sample comparison (Fig. 6) and cluster analysis for all depths (see Fig. S7 in the supplemental material) showed a clear delineation between AK and OK samples (PERMANOVA, \( P = 0.001 \); ANOSIM, \( P = 0.001 \)). As evidenced from the NMDS, sample dispersion was significantly different between the shallow AK and OK samples (PERMDISP, \( F = 25.26, P = 0.001 \)). The top 29 genera reported from SIMPER analysis contributed to 21.2% of the dissimilarity (Fig. 7). The Leotiomycetes and Chytridiomycetes were most associated with the Alaska samples, while the Agaricomycetes, Eurotiomycetes, Sordariomycetes, and Lecanoromycetes dominated at the Oklahoma site.

![Composition of the top 10 taxa at the (A) class, (B) order, and (C) family levels and the top 20 taxa at the (D) genus level for Oklahoma samples. The total percentages of all sequences are (A) 97%, (B) 75.1%, (C) 49.3%, and (D) 46.9%.

**FIG 2** Composition of the top 10 taxa at the (A) class, (B) order, and (C) family levels and the top 20 taxa at the (D) genus level for Oklahoma samples. The total percentages of all sequences are (A) 97%, (B) 75.1%, (C) 49.3%, and (D) 46.9%. 

Penton et al. 7066 aem.asm.org Applied and Environmental Microbiology
ARBSILVA LSU reference database comparison. To investigate more deeply the composition of the abundant unclassified genus-level bins and any bias associated with the RDP fungal classifier, we subjected these 158,860 “unclassified” sequences to BLAST against the ARB-SILVA LSU reference database (release 111; www.arb-silva.de). This resulted in hits to 917 taxa, of which 364 were classified as Fungi. These 80,312 sequences classified as Fungi comprised significantly different (ANOVA, \( P < 0.05 \)) relative abundances of the previously unclassified taxa, with 32%, 48%, and 62% of the AK deep, AK shallow, and OK totals, respectively. The majority of new fungal hits were to relatively few taxa (see Table S3 in the supplemental material). The remaining RDP unclassified hits were mostly to Metazoa (especially *Gomphiocephalus hodgsoni*; 7.9% of all “unclassified” reads), cyanobacteria, and green algae (Viridiplantae) sequences. The deep AK samples were dominated by hits to the yeast *Cyberlindnera meyerae* (80% of sequences, 88.1% average identity), while the shallow AK and OK samples each contained ~25% of hits to the brown rot fungus *Serpula lacrymans* and 21% and 12% of hits, respectively, to *Phaeosphaeria avenaria*. Between sites, AK samples shared 47.5% of all SILVA previously unclassified best matches, while OK samples shared 67.4%.

To verify that the inclusion of these SILVA fungal hits would not alter our overall results, we repeated our analyses with these data in combination with the 50% RDP classifier data with the “unclassified” bins removed to create a new RDP_SILVA best-match data set. Differences between the shallow AK versus OK samples remained significant (ANOSIM, \( P < 0.05 \); PERMANOVA, \( P < 0.01 \)), as did the comparisons between the shallow and deep AK samples (ANOSIM, \( P < 0.05 \); PERMANOVA, \( P < 0.01 \)). Warming treatment effects remained nonsignificant. The addition of the SILVA hits did not influence the significant varia-

**FIG 3** SIMPER analysis showing the top contributing genera to the difference between the 0- to 20-cm and 20- to 50-cm soil layers in the Alaska samples. Bars indicate average percent relative abundances in the resampled data set. Percentages to the right indicate the percent dissimilarity contribution of each genus.

**FIG 4** Percent unclassified as determined by classification at 50% bootstrap confidence using the RDP fungal LSU classifier at each taxonomic level in the (A) Oklahoma and (B) Alaska samples.
tion previously found with species richness, evenness, and Shannon diversity between the AK shallow, AK deep, and OK samples. However, overall richness and Shannon diversity decreased in the AK sites and increased in the OK site.

**DISCUSSION**

**Diversity profiling of grassland and permafrost soil fungi.** Direct classification of 226,695 filtered 28S sequences yielded 1,063 genera from the two ecosystems, covering 62% of the reference data set’s genera. The number of genera recovered per sample ranged from 351 to 594 in the OK grassland samples to 87 to 351 in AK permafrost samples. After normalization for reads per sample, our genus-level diversity was higher than other 97 to 98% cluster-based estimates of species richness from a variety of environments determined by cloning and sequencing (8, 37), ITS pyrosequencing (10, 38, 39), and 28S pyrosequencing but with 95% clustering (40). However, our results agreed with the ITS “species”-level OTU in recent studies from various Mediterranean land use categories (41) and neotropical rain forests (42) but were lower than that found in alpine tundra habitats (15). Such comparisons are difficult to interpret, as the classification resolution differs between the ITS and 28S genes.

Rarefaction curves were nearly saturated in the permafrost samples, especially in the deeper samples, where 4,000 to 8,000 reads was sufficient to cover the majority of fungal diversity present. Comparatively, the sequencing depth was not sufficient to cover the majority of the higher alpha diversity in the Oklahoma prairie samples. As illustrated by lower Pioulou’s evenness, relatively few taxa dominated the permafrost soils. Together these results suggest that fungal diversity is constrained in permafrost samples, as observed in another study from high alpine soils (43) and especially in our deep soils, where plant influences are limited. Despite the significant differences in overall community structure, the sharing of genera encompassing 90% (OK) and 63% (AK) of all fungal sequences (at 0% bootstrap) between the two sites indicated the possible presence of a fungal “seed bank” analogous to that identified with bacteria in a marine system (44).

**Prairie fungal communities.** The dominance of the orders Agaricales, Pleosporales, and Sordariales and the most abundant genus *Montiophthora* (4.5% of all sequences, 84.1% ± 0.3% average classification confidence) in the Oklahoma prairie samples is in agreement with findings from healthy grass root zones in semiarid (45) and desert (46) grassland. *Montiophthora* is related to the causative agents of witches’ broom and frosty pod rot in cacao (47). As noted by Khidir et al. (45), this may be due to an ecological shift from pathogen to endophyte that is not yet understood. We also identified *Cercophora* sp. (2.3%), an abundant endophyte found colonizing *Bouteloua gracilis* (46) and *Phialea* sp. (4.2%), the causative agent of blind-seed disease of ryegrass (48). As a possible artifact of seasonal arbuscular mycorrhizal (AM) fungi, we detected very little *Glomus* sp. (0.06%), an AM genus identified as being dominant in grasslands (49) and found to be widespread among different global ecosystems (50). Primer coverage of the Glomeromycota was verified with the NCBI primer tool. Matches were identified to 2,103 of the 9,441 Glomeromycota 28S sequences in NCBI. Approximately 8,600 of the deposited sequences are of partial length and therefore may or may not contain both primer binding regions. Little is known about the abundant *Arachronymies* (6.5%) species, other than they have been found associated with dung and decaying wood (51) as well as implicated in human nail infections (52). The most abundant basidiomycete was *Leucopaxillus* (3.2%), described as non ectomy corrhizal but saprotrophic in grasslands and forests (53) and which includes *L. giganteus*, which produces grassland “fairy rings.”

**Permafrost fungal communities.** As found elsewhere (3, 7) in tundra soils, Ascomycota (63.5% of reads) dominated over Basidiomycota (11.5%) in our samples. However, this did not agree with the findings of Wallenstein et al. (8), who found higher levels of Basidiomycota in Arctic tundra tussock soils. This may be largely dependent on the composition of the plant community. The dominance of the order Helotiales (average abundance, 43.3%), which is also abundant in other Arctic studies (3, 54), points to its broad niche range (55) and likely reflects substrate heterogeneity within the deep soil profiles. The mostly aquatic Chytridiomycetes, which constitute 8.5% of all sequences, were found at higher dominance in high-elevation sites in Nepal and Colorado (56) and were easily isolated from other periglacial soils (57).

According to the RDP classifier results, among the most abundant genera, which comprised 41% of all sequences, were the *Kuzuhuaea* (11.2%), belonging to the amoeba predatory order *Zoopagales*. The amoebae, which comprised the majority of nonfungal sequences, with the genera *Amoebidium* and *Nuclearia* accounting for 3.1% of all sequences in the raw data set, are known to be an important component of the C cycle at high latitudes (58). The *Karlingiomycetes* (7.3%), comprised of chitinophilic and cellulostic species (39), and *Godronia* (7.2%), which is known to infect
blueberry and has also been found on lingonberry (Vaccinium vitus-idaea) branches (60), which populate this site (19), were also abundant. Following in abundance are Chloroscypha (5.7%), characterized by endophytic habits (61), the dematiaceous hypomycete Sorocybe (resin fungus) (5.6%), the wood endophyte Hyalodendriella (4.6%), and the putative biotrophic parasite or saprobic Hyaloscypha (4.0%) (62), as well as the boreal ectomycorrhizal genus Russula (1.8%) (63), which were abundant in Arctic tundra associated with Betula nana, which was also abundant at our site (54).

Significant differences in fungal community compositions between the shallow and deep soil samples were likely attributed to influences of the mineral versus organic soil horizons (3), substrate composition, and pH (4, 5, 6). These factors serve to shape the niche partitioning between fungi and bacteria, resulting in competitive exclusion and thus influencing community structure. Fungal community partitioning with depth was observed elsewhere with fungi (64, 65), as well as with Actinobacteria (66). Within our AK site, shallower samples are mostly contained within the permafrost thaw depth, where soil is saturated, temperatures are warmer, and decomposition processes are more rapid (19).

**Warming effects.** Although warming was not a significant factor in shaping the permafrost or prairie fungal community as a whole, we did identify a few relationships of interest within the permafrost samples. In a fungal clone study, Arctic tundra warming was also shown to influence individual fungal groups but not overall community structure (3). While our observed reduction in Helotiaceae abundances did not agree with a previous Arctic warming study (3), the mycorrhizal fungi Russula, Lactarius, and Cortinarius increased with warming, in agreement with previous studies that found increases in mycorrhizal abundance or diversity with warming (54, 67, 68). Although warming has resulted in enhanced graminoid growth (20), other factors may play a combined role in shaping fungal communities in the permafrost. In addition, random sample noise related to the reproducibility of amplicon pyrosequencing may have negatively affected the discrimination between treatments (69).

For the Oklahoma prairie site, previous studies have found an increase in the ratio of fungi to bacteria and a change in the percentage of signature fungal phospholipid-derived fatty acids (PLFAs) (25) and increased CO$_2$ efflux (26) with warming but little change in fungal enzyme abundances (27). Although warming at this site has resulted in increases in C$_4$ but not C$_3$ biomass (25, 70), which has increased the C/N ratio of litter inputs, there were only 10 low-abundance fungal genera that significantly changed with treatment. Previously, the influences of warming treatment on bacterial community structure at this site have been heavily influenced by precipitation, with the largest changes associated with wet months (71).

Regarding the lack of significant overall warming effects, the Alaska CIPEHR samples only represent 1 year of warming, a short time frame to expect microbial community shifts. However, the Oklahoma samples represent over a decade of warming. This lack of warming influences is likely due to the overwhelming influence of plant communities, as other studies have found that shading, precipitation, and plant composition (8, 72, 73, 74, 75), as well as seasonality (7, 8, 9), are the main drivers of fungal communities, which may overwhelm the marginal (~2 to 3°C) warming effects. Although we did not find significant changes in community structure with warming, it should also be noted that this was based on genus-level resolution, and there are considerable genetic differences that remain at the species level that are not addressed in this study.
**Unclassified fungi.** Most fungal sequencing studies have not addressed the abundance or composition of the unclassified taxa because the use of BLAST or hidden Markov model algorithms (Fungal ITS Pipeline) (18) for closest-match taxonomic assignment provides little information about the unclassified sequences at different taxonomic levels. Lentendu et al. (15) expanded on this method to allow fungus-targeted sequence assignments at the coarse phylum, kingdom, and domain levels. When an unfiltered reference data set was used to assign ITS reads, Buée et al. (10) found that 71.5% of sequences lacked an explicit taxonomic annotation, on par with our estimations at a finer taxonomic resolution. After database filtering of unknown fungal species, their unclassified proportion dropped to 11% of the OTU by forcing assignment to a classified nearest neighbor. As such, taxonomic assignment is dependent on the composition of the reference database, which is further complicated by the high fungal species diversity in soil (10, 15), estimated to be $7.12 \times 10^7$ at the lower boundary (76).

“Unclassified” sequences, defined here as those with <50% classification confidence using the RDP classifier, are a consequence of either (i) high sequence dissimilarity to a nearest match or (ii) equal classification probability to multiple genera. In both cases, BLAST retrieves a best-match hit used for taxonomic assignment without confidence estimates. By invoking a confidence cutoff, we reduced the likelihood of erroneous taxonomic assignment, although there are cases where a high confidence may also be associated with an incorrect call. For example, a query in which the closest genus is not present in the database may still be assigned to the genus with the higher rank-order likelihood score while supported by bootstrap replication (14).

Likely a consequence of the composition of the reference database, the more thoroughly characterized Ascomycota and Basidiomycota contained the smallest proportion of sequences classified at less than 50% confidence. The reference database Fungal_LSU_train_1400bp_8506_mod used for the RDP classifier and tested by Liu et al. (11) contains 8,506 unique sequences distributed into 1,702 genera, 412 families, 124 orders, 43 classes, and 9 phyla, with the majority of sequences belonging to the Basidiomycota (70.8%) and Ascomycota (28.2%). In contrast, of the approximately 100,000 accepted, described species of fungi (77), approximately 64,000 belong to the Ascomycota and 32,000 to Basidiomycota. Although these numbers probably greatly underrepresent the true diversity of these fungi, they are likely a reasonable approximation of the relative diversity of these two phyla (78). As in other ITS databases, several phyla, such as the Glomeromycota, were underrepresented (10, 16, 79, 80) leading to abundant “unclassified” bins. Of the abundant unclassified family bins (see Fig. S5 and S6 in the supplemental material), the unclassified Tricholomataceae within the prairie site (5.4% of all reads) are unique. This family is among the most highly represented in the dataset, with 8.3% of all sequences in 121 genera. Two types of “unclassified” sequences are likely represented here due to the underrepresentation of many genera of Tricholomataceae in the RDP database and the polyphyletic nature of the genera (81) within the family. Conversely, the unclassified Helotiales contain 20.1% of all reads in permafrost but are only represented by 19 sequences covering 13 genera in the database.

The results from subjecting our “unclassified” sequences to BLAST against the ARB-SILVA LSU reference database showed that 50.6% of those sequences had best matches to fungi, while the remaining binned to other Eukarya. The wide range of BLAST percent identities to fungi (see Table S3 in the supplemental material) illustrate potential novelty on the low end (e.g., 78.0% to Synchytrium decipiens) and cases in which low RDP confidence was likely based on two or more closely matching sequences on the high end (e.g., 96.6% identity to Chytridium sp. strain JEL341). Reads binned into Eukaryota incertae sedis and Fungi incertae sedis in the original data sets comprised an average of 30.7% of all the unclassified sequences at the genus-level bootstrap, compared to 60% binning into nonfungi after BLAST analyses of the same data set. This increase is most likely due to the composition of the SILVA database; 94% of all reference sequences are nonfungal (1,802 sequences are fungal). Thus, sequences that RDP binned into unclassified fungi due to high divergence likely have a higher probability of a closest match to the abundant nonfungal SILVA sequences. However, the inclusion of additional nonfungal sequences into the RDP database will likely change the binning results in both the best-match and 50% bootstrap databases, as divergent sequences alter the statistical underpinning of the naive Bayesian classifier model.

**Rare biosphere.** Since a significant proportion of rare OTU were successfully assigned to a genus at 50% bootstrap confidence (68.7% of singletons, 59.7% of doubletons, and 70.3% of tripletons), the assignment of singletons as sequencing artifacts is not justified. This is supported by Lentendu et al. (15), who found that 53.3% of rare ITS molecular OTU (MOTU$_{98,98}$) could be successfully assigned to Fungi. Our singletons contained 24.1% of all genera but accounted for less than 0.01% of all sequences. This is lower than the singleton abundances of 38% (15) and 60% (10) reported using ITS cluster-based methods. Differences in the bioinformatic approach, such as clustering, rather than sequencing depth, are likely responsible for a large portion of the differences in singleton abundances among these studies as the addition of singletons decreases proportionally with increased sequencing coverage, as reflected by the flattening of rarefaction curves.

In summary, while there was no effect of artificial warming on fungal communities in AK and OK soils after 1 year of warming, we did observe significant changes between sites and with depth at the AK sites. The ability to directly classify 28S sequences at different confidence levels without the limits imposed by clustering and representative sequence generation allowed us to explore in greater depth the composition of “unclassified” sequences. The majority of “unclassified” reads were assigned to the more abundant families in each site, whereas most of the rare biosphere could be successfully assigned to the genus level. This illustrated the current dearth of fungal sequences from classified taxa, likely a consequence of high community heterogeneity among environments. The implementation of BLAST with the ARB-SILVA database showed instances where low NBC classification confidence could be due to either a large distance to the nearest relative or the presence of an equal probability of classification to two or more taxa. This direct comparison between methods showed limitations regarding the choice of taxonomic assignment method and the coverage of current fungal databases. While a majority of fungal sequencing studies have targeted subpopulations, such as the AM fungi, to address plant-fungus relationships, current and developing sequencing technology continues to reduce the cost, allowing higher coverage, such that whole-community sequencing will increase database coverage while applying environmental re-
lationships to taxa, which is essential in understanding global fung-
gal distribution.

ACKNOWLEDGMENTS

This work is supported by the Department of Energy, Biological Systems Research on the Role of Microbial Communities in Carbon Cycling Program (DE-SC0004601).

We thank Amanda Pham, Christina Hazekamp, and Travis Baes for laboratory assistance, the staff of the Ribosomal Database Project at Michi-
gan State University for implementation of the Bayesian Fungai Classi-
gram (DE-SC0004601).

Research on the Role of Microbial Communities in Carbon Cycling Pro-

REFERENCES


22. Wright ES, Yilmaz LS, Nooguer DR. 2012. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. Appl. Envi-


30. Xu L, Ravnksov S, Larson J, Nicholson M. 2012. Linking fungal communities in roots to rhizosphere, and to the soil health status of Pismum sati-


