Fungal Community Analysis by Large-Scale Sequencing of Environmental Samples†

Heath E. O’Brien,1* Jeri Lynn Parrent,1 Jason A. Jackson,1 Jean-Marc Moncalvo,2 and Rytas Vilgalys1

Department of Biology, Duke University, Durham, North Carolina 27708;1 and Centre for Biodiversity and Conservation Biology, Royal Ontario Museum, and Department of Botany, University of Toronto, Toronto, Canada2

Received 10 November 2004/Accepted 15 April 2005

Fungi are an important and diverse component of soil communities, but these communities have proven difficult to study in conventional biotic surveys. We evaluated soil fungal diversity at two sites in a temperate forest using direct isolation of small-subunit and internal transcribed spacer (ITS) rRNA genes by PCR and high-throughput sequencing of cloned fragments. We identified 412 sequence types from 863 fungal ITS sequences, as well as 112 ITS sequences from other eukaryotic microorganisms. Equal proportions of Basidiomycota and Ascomycota sequences were present in both the ITS and small-subunit libraries, while members of other fungal phyla were recovered at much lower frequencies. Many sequences closely matched sequences from mycorrhizal, plant-pathogenic, and saprophytic fungi. Compositional differences were observed among samples from different soil depths, with mycorrhizal species predominating deeper in the soil profile and saprophytic species predominating in the litter layer. Richness was consistently lowest in the deepest soil horizon samples. Comparable levels of fungal richness have been observed following traditional specimen-based collecting and culturing surveys, but only after much more extensive sampling. The high rate at which new sequence types were recovered even after sampling 863 fungal ITS sequences and the dominance of fungi in our libraries relative to other eukaryotes suggest that the abundance and diversity of fungi in forest soils may be much higher than previously hypothesized.

Fungi are a diverse component of soil microbial communities, in which they function as decomposers, mycorrhizal mutualists, and pathogens. Studying the ecological interactions of these organisms is challenging because of the extremely high diversity of soil fungi, the complexity of the substrate, and the difficulty of direct observation of these communities (4). Mycologists have traditionally examined soil fungal communities by collecting macroscopic sexual reproductive structures (fruiting bodies) that are produced above ground or by using isolation techniques to obtain pure cultures of fungal strains directly from soil (29). However, these approaches have several limitations. Conventional culturing methods are heavily biased towards fast-growing species, and many fungi have specialized growth requirements, so this approach recovers only a small subset of the community. Similarly, fruiting body collection is limited to the detection of species that frequently reproduce sexually unless long-term studies are conducted (33). Furthermore, since fungi vary dramatically in their resource allocation, there is often a lack of correspondence between above-ground and below-ground views of fungal community structure (15). Finally, many fungi are nonculturable, produce inconspicuous or hypogeous fruiting bodies, or do not reproduce sexually at all and thus are overlooked entirely by these methods.

Sequence-based studies of DNA obtained directly from environmental samples are revolutionizing our view of microbial diversity (18). This technique involves using PCR to amplify target regions of DNA extracted directly from environmental samples. These amplicons are then cloned to isolate individual fragments and sequenced. The application of this approach to a wide range of substrates has resulted in the identification of a large number of new bacterial species that cannot be cultivated, many of which constitute novel lineages or are dominant components of microbial communities (11). This technique has become an indispensable tool in bacterial soil ecology, but its use in studies of eukaryotic microorganisms has been limited. Sequence-based studies have been carried out on fungi from grass leaves (6), plant roots (10, 35, 38), and soil (30), but most studies have been restricted to the use of methods based on fragment length polymorphisms (13, 39) or denaturing gradient gel electrophoresis (2, 37). These methods collectively have the potential to improve our understanding of fungal biodiversity and facilitate ecological studies of soil communities.

Most recent studies of uncultured microbial diversity have sequenced genes coding for the small-subunit (SSU) rRNA. Sequencing of SSU rRNA is highly effective in bacteria, but this region does not evolve rapidly enough in eukaryotes for use in more than higher level taxonomic assignments (38). The internal transcribed spacers (ITS) that separate the SSU and large-subunit (LSU) rRNAs in eukaryotes evolve at a much faster rate and can be used to identify fungi to genus and often to species level. This region has become a critical marker for species identification from fungal vegetative structures such as mycorrhizae (14, 35). Due in large part to the use of ITS sequence data for fungal molecular systematic studies (5), a large database of publicly available ITS sequences from diverse

* Corresponding author. Mailing address: Department of Biology, Duke University, Durham NC 27708. Phone: (919) 660-7285. Fax: (919) 660-7293. E-mail: heo3@duke.edu.
† Supplemental material for this article may be found at http://aem.asm.org.
fungal taxa now exists (40). If the taxonomic coverage of this database is sufficient such that all sequences from environmental samples will have a close match in the database, the ITS rRNA will provide high-level resolution among different members of fungal communities while simultaneously providing information on taxonomic composition.

Our objectives in this study were (i) to determine the feasibility of using sequence-based biotic inventories for soil fungi to quantify species richness and to detect correlations between the taxonomic composition of soil communities and environmental variables and (ii) to compare the utility of the ITS and SSU rRNA regions for characterizing soil fungi. We hypothesized that a sequence-based approach would detect higher levels of fungal diversity in the soil, representing a greater taxonomic breadth than is usually detected with traditional survey methods, and that distinct fungal communities could be detected at different depths in the soil profile. We expected the ITS region to provide finer-scale taxonomic resolution of community composition than the SSU while still permitting taxonomic placement of sequences because of the extensive sequence database available. If this approach is successful, it can be used to greatly increase our understanding of the dynamics of soil fungal communities through rapid and unbiased inventories of soil fungi.

MATERIALS AND METHODS

Sampling. Two plots in Duke Forest, Durham, N.C., were sampled: a early successional well-drained, upland stand dominated by loblolly pine (Pinus taeda) containing a few hardwood species regenerating in the understory and a thermic Ultic Hapludalf soil in the Enon series; and a late successional, low-lying, mature, mixed hardwood stand with an alluvial Fluvaquentic Dystroelpt soil in the Chewacla soil series. Each plot was 0.1 ha; the two plots are approximately 6.5 km apart and are two of approximately 250 plots originally established by Peet and Christensen in the 1970s to study forest succession and community diversity (8).

At each site, 1-m² quadrats were sampled at 5-m intervals along transects spaced approximately 5 m apart, for a total of 50 soil samples per plot. For each quadrat, all of the litter was collected and one soil core 2.5 cm in diameter and 20 cm deep was taken. In each soil core distinct zonation corresponding to O (organic material and detritus), A (organic matter-rich mineral soil), and B (clay-rich mineral soil) horizons could be visually detected. Each soil core was subdivided by zone, and all samples within a soil horizon type were composited to overcome any within-plot spatial heterogeneity. Litter samples were air dried for 24 h and homogenized by grinding in a Wiley mill (Thomas Scientific, Swedesboro, NJ). Organic and mineral soil samples were sieved using a 2-mm sieve to remove large rocks and roots. Samples were stored at 4°C for 1 week or less before processing. Subsamples (50 ml) were finely ground using a sterile mortar and pestle and liquid nitrogen and stored at -80°C until DNA was extracted. Samples were designated according to plot (P, pine; MH, mixed hardwood) and horizon (L, litter; O, organic; A, A horizon; and B, B horizon).

Molecular methods. DNA was isolated from 0.5 g of the finely ground subsamples using the Mo Bio UltraClean soil DNA extraction kit (Mo Bio Laboratories, Solana Beach, CA) according to the manufacturer’s instructions. The ITS and SSU rRNA regions were amplified from these extracts using universal primers (ITS1, ITS4, NS1, and NS2) (42) as well as taxon-specific primers ITS1F (fungus-specific) and ITS4B (Basidiomycota specific) (16). PCR was performed using the following protocol: initial denaturation at 94°C for 3 min, followed by 35 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min, with a final extension at 72°C for 10 min. A modified PCR protocol consisting of an initial denaturation at 94°C for 3 min, followed by 20 cycles of 94°C for 1 min, 50°C for 30 s, and 72°C for 2 min, with a final extension at 72°C for 7 min was used for some PCRs to assess the effect of cycle number and extension time on the generation of chimeric sequences and other PCR artifacts (Table 1). Each 25-μl PCR consisted of 25 μl bovine serum albumin, 0.625 U Taq DNA polymerase (Abgene, Rochester, NY), 1.5 mM MgCl₂, deoxynucleoside triphosphates (0.2 mM each), primers (0.5 μM each), and PCR buffer.

Fragments were cloned into Escherichia coli plasmids using the Topo-TA 5-minute PCR cloning kit (Invitrogen, Carlsbad, CA) following the manufacturer’s instructions. The maximum recommended incubation times were used for each step with 4 μl of PCR product. For SSU fragments, the L, O, A, and B PCR products were pooled and used to generate a single library.

White colonies were screened for inserts by adding whole cells directly to 10 ml of 2x YT medium, incubated at 37°C for 24 h, and grown on agar plates. Positive clones were used to prepare template DNA for sequencing.

Molecular methods. Fragments were cloned into plasmids using the Topo-TA cloning kit following the manufacturer’s instructions. The maximum recommended incubation times were used for each step with 4 μl of PCR product. For SSU fragments, the L, O, A, and B PCR products were pooled and used to generate a single library.

White colonies were screened for inserts by adding whole cells directly to 10 ml of 2x YT medium, incubated at 37°C for 24 h, and grown on agar plates. Positive clones were used to prepare template DNA for sequencing.

Molecular methods. Fragments were cloned into plasmids using the Topo-TA cloning kit following the manufacturer’s instructions. The maximum recommended incubation times were used for each step with 4 μl of PCR product. For SSU fragments, the L, O, A, and B PCR products were pooled and used to generate a single library.

White colonies were screened for inserts by adding whole cells directly to 10 ml of 2x YT medium, incubated at 37°C for 24 h, and grown on agar plates. Positive clones were used to prepare template DNA for sequencing.
an alignment score over 600 could be identified to the level of genus, but this was not possible for some taxonomically problematic groups. We did not attempt to identify OTUs to the level of species. The taxonomic placement of sequences with high similarity to anamorphic fungi were determined either by looking at the best BLAST match for a teleomorph or by consulting phylogenetic studies which included that species (23, 31).

All unique SSU sequences were queried against GenBank using BLAST (1). Chimeric sequences were identified using the program Chimera Check from the Ribosomal Database Project II (25) and eliminated from subsequent analyses. The remaining sequences were subjected to automated multiple alignment using ClustalW (36) and the resulting alignment was edited by eye using MacClade 4.0 (24). Ambiguously aligned regions were excluded from the alignment. Phylogenetic analysis was conducted with the neighbor joining method using PAUP*4.0b10 (34).

**RESULTS**

**Description of libraries.** Fifteen ITS clone libraries were generated for this study; six from MH-O (designated MH-O-1 through MH-O-6), two from MH-L and MH-A (designated MH-L-1, MH-L-2, MH-A-1, and MH-A-2), and one each from MH-B, P-L, P-O, P-A, and P-B. All libraries were generated using the primers ITS1 and ITS4 for initial PCR amplification, except for MH-O-5 and MH-O-6, for which ITS1F/ITS4 (specific to fungi) and ITS1F/ITS4B (specific to basidiomycete fungi) were used, respectively (Table 1). The number of colonies sequenced per library varied from 18 to 123, with a minimum of 62 sequences (P-O) and a maximum of 217 sequences (MH-O) per soil sample (Table 1). In total, 1,015 sequences were obtained, 40 of which were determined to be chimeric and removed from the data set.

The 975 nonchimeric sequences comprised 454 unique OTUs, 301 of which were singletons occurring only once in the entire data set. The remaining 153 OTUs ranged in abundance from 2 to 51 sequences (0.2% and 5.2% of the data set, respectively). Since primer specificity affects the breadth of phylogenetic diversity recovered, only libraries generated with the universal primer pair ITS1 and ITS4 were considered when comparing diversity and community composition among soil horizons and between plots.

One SSU library was generated for each plot, and 100 and 92 clones were sequenced for the P and MH plots, respectively. Eleven of these sequences were determined to be chimeric and removed from subsequent analyses. A total of 153 unique sequence types were recovered, with each occurring between 1 and 12 times in the data set. The phylogenetic analysis of SSU sequences confirmed the results of BLAST searches as all Ascomycota and Basidiomycota sequences formed monophyletic groups (Fig. 1). Four sequences that matched members of the Chytridiomycota and Zygomycota fell outside their corresponding clusters, but all fungal sequences grouped together.

**Community composition.** For all but one ITS library, the proportion of sequences determined to be of fungal origin ranged from 84% to 100% (Fig. 2A). The only exception was MH-A-1, which possessed 65 nonfungal sequences out of 101, 49 of which belonged to a single OTU. Because this OTU was not found at high frequency in MH-A-2, this result was considered an artifact of library construction and MH-A-1 was excluded from diversity and community composition comparisons. Seventy-eight percent of SSU sequences were of fungal origin, while a slightly higher percentage (88%) of fungal ITS sequences could be unambiguously identified to a fungal phylum by BLAST.

The majority of fungal sequences recovered in this study belonged to the Ascomycota and Basidiomycota, with a slightly greater proportion of Ascomycota sequences (Ascomycota relative frequency = 46% ITS, 35% SSU; Basidiomycota relative frequency = 41% ITS, 34% SSU). However, the relative abundance of Ascomycota and Basidiomycota ITS sequences varied greatly among libraries, ranging from 18% Ascomycota and 60% Basidiomycota for P-A to 69% Ascomycota and 30% Basidiomycota for MH-L (Fig. 2A). There were no apparent trends in relative abundance of these phyla with plot type or soil horizon; Ascomycota sequences were most abundant in MH-O, MH-A, MH-L, P-L, and P-B, while Basidiomycota sequences were most abundant in MH-B, P-A, and P-O. Zygomycota and Glomeromycota sequences were present in both SSU and ITS libraries, but were rarer in ITS libraries (Zygo-
mycota relative frequency = 1.5% ITS, 4% SSU; Glomeromy-
cota relative frequency = 0.5% ITS, 1% SSU). Four percent of
SSU sequences matched the Chytridiomycota, but no ITS se-
quences matched members of this phylum.

Composition of Ascomycota sequences. Sixty-six percent of
all Ascomycota ITS sequences recovered in this study could be
identified to subclass or lower. The most abundant subclasses
were the Leotiomycetidae and Dothidomycetidae. Members of
the Sordariomycetidae, Chaetothyriomycetidae, Pezizomyceti-
dae, and Eurotiomycetidae also were found in many samples.
There were no clear compositional differences with depth at
this taxonomic scale, but Dothidiomycetidae sequences were
most abundant in the mixed hardwood plot, while Leotiomy-
cetidae sequences dominated in the pine plot (Fig. 2B). Twenty-
five percent of the Ascomycota ITS sequences could be
identified to genus, with the most abundant genera including
Phialophora (Leotiomycetidae), Hormonema (Dothidiomyceti-
dae), and Alternaria (Dothidiomycetidae) (Table 2).

Composition of Basidiomycota sequences. For Basidiomy-
cota ITS sequences, 64% could be identified to order or lower.
Over half of all identifiable sequences matched members of
two Basidiomycota orders, the Agaricales and Russulales, with
the Thelephorales, Polyporales, and Tremellales constituting
an additional 29%. ITS clone libraries from the L horizons
were enriched in Agaricales sequences relative to those at
greater soil depth, while the proportions of Russulales sequences
were much lower in the L horizons of both plots compared to the
O, A, and B horizons (Fig. 2C). These differences can be attri-

TABLE 2. Number of sequences recovered for the most common genera encountereda

<table>
<thead>
<tr>
<th>Genus (order/subclass)</th>
<th>No. of sequences in library:</th>
<th>MH-L</th>
<th>P-L</th>
<th>MH-O</th>
<th>P-O</th>
<th>MH-A</th>
<th>P-A</th>
<th>MH-B</th>
<th>P-B</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Russula</strong> (Russulales)</td>
<td></td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>12</td>
<td>2</td>
<td>13</td>
<td>35</td>
</tr>
<tr>
<td><strong>Myccena</strong> (Agaricales)</td>
<td></td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td><strong>Atelia</strong> (Polyporales)</td>
<td></td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td><strong>Phialophora</strong> (Leotiomycetidae)</td>
<td></td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td><strong>Tomentella</strong> (Thelephorales)</td>
<td></td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td><strong>Hormonema</strong> (Dothidiomycetidae)</td>
<td></td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td><strong>Alternaria</strong> (Dothidiomycetidae)</td>
<td></td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td><strong>Cryptococcus</strong> (Tremellales)</td>
<td></td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><strong>Hygrocybe</strong> (Agaricales)</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td><strong>Tylopora</strong> (Polyporales)</td>
<td></td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

a Excluding MH-O-5, MH-O-6, and MH-A-1.
uted primarily to the distributions of the most abundant genus in these two orders. The genus *Russula* made up 75% of Russulales sequences in A and B horizon samples and 44% in the O horizon samples, but was absent from the L horizon libraries (Table 2). Similarly, *Mycena* (Agaricales) was one of the most abundant genera overall, but 11 out of 16 *Mycena* sequences occurred in the mixed hardwood L horizon.

**Richness and diversity estimates.** Shannon-Weiner diversity values for clone libraries ranged from 3.58 for MH-O-6 to 5.61 for P-A (Table 1). Overall, diversity was positively correlated with sample size, except for library MH-A-1 (101 sequences, $H = 3.62$), for which 49% of the sequences belong to a single OTU. Species-effort curves for all samples have a positive slope throughout and show no evidence of approaching saturation, indicating that true richness is much higher than we have captured in our samples (19) (Fig. 3A and B). Similar results were obtained when we examined the data pooled by horizon (Fig. 3C) or plot (Fig. 4).

The curves for all samples had similar slopes except for MH-B, which appeared to be more saturated than other samples from the mixed hardwood plot (Fig. 3A). When data were pooled by horizon, O horizon richness was highest, followed by the A and L horizons, with the B horizon having the lowest richness (Fig. 3C). When sequences were pooled by plot, the curves were indistinguishable (Fig. 4). Pine plot richness was estimated to be 491 using the ACE estimator; the estimated richness of the mixed hardwood plot was 616. Predicted richness continued to climb with increasing sample size (Fig. 4), suggesting that these values are likely to be underestimates of true richness.

**DISCUSSION**

**Methodological considerations.** Two difficulties inherent in molecular diversity surveys are determining an appropriate sequence divergence cutoff for defining OTUs that correspond to traditionally recognized species and elimination of PCR artifacts. PCR artifacts encountered in this study included the generation of chimeric sequences due to PCR recombination and biased amplification of a single sequence type. Approximately 5% of the sequences with high BLAST matches in both spacers showed strong differences between ITS1 and ITS2 BLAST matches and were therefore deemed chimeric sequences. Amplification bias in a single PCR could be detected by comparing the frequency of OTUs in independent libraries (e.g., 49 copies of one OTU in MH-A-1 versus 2 copies in MH-A-2). Since adjustments to template concentrations and thermal cycler conditions and pooling of PCR products were unable to completely eliminate these problems, we recom-

![FIG. 3. Species-effort curves of richness for fungal OTUs. A. Mixed hardwood plot. B. Pine plot. C. Pooled data from both plots.](image)

![FIG. 4. Species-effort curves for observed and estimated fungal OTU richness in each plot.](image)
mend making multiple independent libraries from separate DNA extracts to accurately evaluate OTU abundances.

The large range of intraspecific ITS variation reported in the literature complicates the determination of an appropriate sequence similarity cutoff. In some cases, different species have been found to have ITS similarity greater than 99% (12, 20) while intraspecific ITS similarity of 90% or lower has been reported for others (22, 27). The degree of ITS sequence variation within and between fungal species is an issue that deserves greater attention if we are to continue to rely on this gene region to delimit species, but clustering sequences with 97% or greater sequence similarity is likely a conservative measure of species richness, especially when sampling from a single locality.

**Comparisons of ITS sequencing with traditional fungal survey methods.** Despite methodological limitations, this study demonstrates the potential of large-scale sequencing projects targeting the ITS rRNA for quantifying and characterizing soil fungal diversity. We recovered 412 unique OTUs from a total of 863 fungal sequences. This level of richness is comparable to that found in a 1,500-m² forest plot in Switzerland, where 408 species were recovered from fruiting body samples (33), and to richness levels found for endophytic fungi isolated from leaves of two tropical trees, where 418 morphospecies corresponding to an estimated 347 genetic species were recovered (3). However, the study in Switzerland required 21 years of sampling and identification of 71,222 fruiting bodies (33); likewise, the endophyte study involved examining 1,472 cultures that were isolated from 1,992 leaf segments (3).

We captured this level of richness from 863 fungal sequences isolated from only two 2,000-m² forest plots in a single locale at a single point in time. Given the lack of saturation in our species-effort curves, we would have vastly exceeded their richness estimates if our sampling effort had been comparable to that of these other studies. Using DNA sequences as the primary sampling unit also requires much less processing of each sample than more labor-intensive methods such as fruiting body collection and culture-based census taking and is therefore amenable to automation of data collection and analysis, making large-scale biotic surveys conducted over short time periods feasible.

**Comparisons of ITS sequencing with other molecular survey methods.** While sequencing rRNA clones is standard practice in environmental bacteriology, most fungal studies using environmental sampling have employed either fingerprinting methods based on banding patterns obtained from restriction site polymorphisms (13) or denaturing gradient gel electrophoresis profiles (2, 37). These methods can be useful for detecting overall community shifts in response to environmental variables and for tracking the presence or absence of individual OTUs in different samples. However, they are limited in their ability to enumerate species richness in complex communities and do not provide any information about the taxonomic affiliation of the OTUs.

When sequencing efforts have been conducted, most studies have targeted either SSU or LSU rRNA. These regions are useful for phylogenetic characterization of novel lineages (30, 38), but lack sufficient resolution to discriminate among closely related species, resulting in underestimation of diversity. While most fungal ITS sequences were identifiable, 12% could not be assigned to a phylum. Many of these unidentifiable sequences may represent novel lineages with no representatives known from cultures or fruiting bodies, but almost all SSU sequences cluster within known fungal groups in phylogenetic analysis, suggesting that they may represent known fungi which are not yet represented in the ITS database. Since the ITS region allows fine-scale taxonomic resolution but lacks sufficient conservation for phylogenetic analyses of unknown sequences, we recommend that future studies coamplify the ITS with either the SSU or LSU rRNA region as a single fragment.

**Taxonomic breadth of libraries and compositional differences with depth in soil profile.** Comparing the relative importance of major groups of fungi and other soil microbes has been challenging for soil ecologists. Because the primers used to generate clone libraries for this study have been shown to have broad specificity across the fungi, metazoa, and protozoa, we would expect the composition of our libraries to be roughly proportional to the number of genomes of each group present in our sample. This is confirmed by the overall consistency found between ITS and SSU clone libraries obtained using independent priming sites despite differences between these regions in selective constraints. We did find the Chytridiomycota and Zygomycota to be underrepresented in the ITS libraries (0% and 1.5%, respectively) relative to SSU libraries (4% each). While this may reflect differences in primer specificity, it should be noted that 7% of our fungal ITS sequences could not be identified to phylum and may belong to these groups which are underrepresented in public databases relative to other fungal phyla (40).

While sample diversity was high relative to our sampling effort, we were able to detect compositional differences at higher taxonomic scales. Many of the taxa found in the litter layer, such as *Mycena* species, are known decomposers, while mycorrhizal Basidiomycota sequences, including *Russula* species, were most common in the O, A, and B layers. The B horizon samples contained a higher proportion of unidentifiable sequences and were considerably less diverse than the other horizons, which may reflect a reduction in niche breadth with soil depth.

**Implications for fungal diversity.** We found a remarkably large number of species from just a few grams of soil, but this is only a fraction of the estimated number of fungal species worldwide. Based on extrapolations from the ratio of fungi to vascular plants in the British Isles to the a worldwide vascular plant species richness estimate of 270,000 species, global fungal richness has been estimated to be 1.5 million species (17). Based on our ITS-based estimates of soil fungal richness (491 and 616 for P and MH, respectively) and the vascular plant richness of these plots (26 and 48, respectively; R. Peet, personal communication), this calculation yields global soil fungal species richness estimates ranging from 3.5 to 5.1 million. These values are almost certainly underestimates because the richness estimators continue to increase with increasing sampling and because of the conservative 97% ITS similarity clustering used to designate OTUs.

Given the results of the present study, along with data suggesting that endophytic (3) and insect-associated (41) fungi also represent vast amounts of formerly uncharted fungal diversity, the true magnitude of fungal diversity could be orders of magnitude higher than previously suggested.
ACKNOWLEDGMENTS

We thank Jason Stajich for assistance with bioinformatics, Dave Carmine for help with the DFM database, and R. Peet, D. Urban, and N. Christensen for providing plant census data from their “Long-term succession on the North Carolina Piedmont” project funded by NSF grants DEB97-07551 and DEB-7804043. This work was supported in part by National Science Foundation grant MCB-00-84207 to R.V., J.M., and J. E. Johnson.

REFERENCES