REVIEW ARTICLE

Applications
in Plant Sciences

Choices on sampling, sequencing, and analyzing DNA influence the estimation of community composition of plant fungal symbionts

Jonathan A. Cale[1] | Natalie Scott[1] | Gregory J. Pec[1,2] | Simon M. Landhäusser[1] | Justine Karst[1]

[1]Department of Renewable Resources, University of Alberta, 442 Earth Sciences Building, Edmonton, Alberta T6G 2E3, Canada

[2]Department of Biology, University of Nebraska at Kearney, Kearney, Nebraska 68849, USA

Correspondence

Gregory J. Pec, Department of Biology, University of Nebraska at Kearney, Kearney, Nebraska 68849, USA.
Email: pecg@unk.edu

Abstract

Plant root symbionts, namely mycorrhizal fungi, can be characterized using a variety of methods, but most of these rely on DNA. While Sanger sequencing still fulfills particular research objectives, next-generation sequencing currently dominates the field, thus understanding how the two methods differ is important for identifying both opportunities and limitations to characterizing fungal communities. In addition to testing sequencing methods, we also examined how roots and soils may yield different fungal communities and how disturbance may affect those differences. We sequenced DNA from ectomycorrhizal fungi colonizing roots of *Pinus banksiana* and found that operational taxonomic unit richness was higher, and compositional variance lower, for Illumina MiSeq–sequenced communities compared to Sanger-sequenced communities. We also found that fungal communities associated with roots were distinct in composition compared to those associated with soils and, moreover, that soil-associated fungi were more clustered in composition than those of roots. Finally, we found community dissimilarity between roots and soils was insensitive to disturbance; however, rarefying read counts had a sizeable influence on trends in fungal richness. Although interest in mycorrhizal communities is typically focused on the abiotic and biotic filters sorting fungal species, our study shows that the choice of methods to sample, sequence, and analyze DNA can also influence the estimation of community composition.

KEYWORDS

boreal forest, ectomycorrhizal fungi, Illumina, jack pine, roots, Sanger sequencing, soils

Most plant species engage in symbiosis with mycorrhizal fungi, which in addition to affecting plant nutrition, play important roles in carbon and nutrient cycling in terrestrial ecosystems (Smith and Read, 2008; Johnson et al., 2017). A variety of methods are available to characterize the composition of mycorrhizal fungal communities, but most recent studies rely on DNA sequences. Although interest in fungal community ecology is typically on the abiotic and biotic filters sorting species, the specific methods used to prepare and sequence DNA may also influence species detection. For instance, sample storage (Clasen et al., 2020;

Guerrieri et al., 2021), sequencing primers (Bellemain et al., 2010; Yang et al., 2018), and bioinformatic pipelines (Pauvert et al., 2019) can influence the community composition of soil fungi. While next-generation sequencing methods have become almost default in current mycorrhizal research owing to their ability to sequence a mixed-DNA template across many samples simultaneously, and with high sequencing depth (Nilsson et al., 2019), Sanger sequencing (i.e., sequencing DNA from individual mycorrhizal root tips) may be appropriate when the goal is detecting shifts in common fungal species (e.g., Shemesh

et al., 2020). Sanger sequencing fungal DNA from individual colonized roots necessarily targets mycorrhizal fungi, thus reducing time and effort in the analysis of non-mycorrhizal fungi. However, if the study objectives include detecting non-mycorrhizal fungi or cataloging fungal species richness, tools that capture more of the mycobiome are needed (Tedersoo et al., 2014; Pec et al., 2017). Sanger sequencing differs from next-generation sequencing methods in sequencing volume, sequencing only a single DNA fragment at a time versus millions of DNA fragments simultaneously (Slatko et al., 2018), a feature that could give rise to differences in community composition estimation. Although Sanger sequencing still fulfills particular research objectives, next-generation sequencing is currently the dominant technology and thus understanding how the two methods compare is important for identifying opportunities and limitations to characterizing mycorrhizal fungal communities.

Ectomycorrhizal (EM) fungi, which dominate boreal forests worldwide, acquire resources from living roots and soils (Smith and Read, 2008). They acquire carbon from sugars in the roots of their plant partner and mineral nutrients from soils. In addition to existing in both roots and soils, they also occur in different forms (Smith and Read, 2008). Specifically, EM fungi may occur as hyphae colonizing roots, extra-radical hyphae exploring soils, or in various resting stages (e.g., spores or sclerotia). Historically, there was emphasis on characterizing EM fungi in roots owing to interest in their function in host nutrition. Today, we recognize the functional importance of EM fungi not only in plant nutrition but also in underlying ecosystem processes such as soil carbon cycling (Johnson et al., 2017). The advent of culture-independent methods has allowed unprecedented access to fungal diversity and provided the means to identify hyphae extending from roots, tissues likely to form long-term stores of carbon (Clemmensen et al., 2013). Because EM fungi colonizing roots recruit from the surrounding soils, we expect fungal communities associated with roots to be less diverse than those in soils (Goldmann et al., 2016). Taken together, not only could the choice of sequencing method affect the estimation of fungal community composition, but whether roots or soils are sampled may also influence the species of EM fungi observed in a community.

Here, we analyzed fungi on tree roots and in the surrounding forest soil to understand how common methodological choices in fungal community research could affect the estimation of community composition. Two specific methods were investigated. First, we investigated how the choice of sequencing method affects assemblages by comparing fungal communities identified using Sanger or Illumina MiSeq (high throughput; herein "Illumina") methods. Second, we investigated how the sampled habitat affects fungal community composition by comparing Illumina-identified fungal communities on roots and in soils from forests. As part of this latter study, we also characterized the response of root-associated and soil-associated fungi to

forest clearing, to test whether disturbance decouples the composition of root-associated fungal communities from that of soils. Although we can target EM fungi by sampling EM roots, roots are a habitat for fungi belonging to other trophic guilds that will also be inadvertently sequenced. We therefore consider results for EM fungi and the total fungal community, including dark septate endophytic, EM, saprotrophic, pathogenic, and unidentified fungi. Informed by our findings, we provide methodological considerations for future research on EM fungal community ecology.

METHODS

Comparing sequencing methods

Site description, sample collection, and sample preparation

We collected roots from a pure stand of mature jack pine (*Pinus banksiana* Lamb.) located in the University of Alberta Botanical Garden (Devon, Alberta, Canada [53°24′ 28.86″N, 113°45′37.39″W]). In October 2016, we cored (2.5 cm diameter, 20 cm depth) soils at the base of 30 randomly selected trees and that same day transported cores on ice to the laboratory, where they were stored at −20°C until further processing.

Cores were thawed at 2°C, and each soil sample was spread in separate large aluminum trays. Samples were thoroughly inspected for roots, which were removed with forceps and gently rinsed with deionized water over a 1.2 mm sieve to remove adhering soil. Roots from each sample were mixed and then divided into two portions: one for Sanger sequencing and another for Illumina MiSeq. For Sanger sequencing, the cleaned roots were placed in a sterile Petri dish with deionized water. Using a stereoscope, mycorrhizal root tips were identified by the absence of root hairs, the presence of hyphae and a mantle, and the color and texture of root tips (Goodman et al., 1996–2009). For Illumina MiSeq, fine (<2 mm diameter) and higher-order roots were washed and lyophilized at −45°C for 24 h using a VirTis Freezemobile (FM25XL; SP Scientific, Warminster, Pennsylvania, USA). The dried samples were then ground and homogenized to a fine powder using a mixer mill for 1 min at 25 Hz.

Sanger sequencing and bioinformatics

To identify fungi colonizing the EM roots, DNA was extracted from individual root tips using Sigma Extraction Solution and Neutralization Solution B following manufacturer protocols (Sigma Aldrich, St. Louis, Missouri, USA). The internal transcribed spacer (ITS) region (ITS1 and ITS2, 260–1800 bp; Yang et al., 2018) of the extracted nuclear rDNA was amplified using PCR amplification in 25 μL reactions using primers ITS1-F (5′-CTTGGTCATTTAGAGGAAGTAA-3′) for forward and ITS4

(5′-TCCTCCGCTTATTGATATGC-3′) for reverse directions (Gardes and Bruns, 1993) at 1.0 µL of DNA extract, 6.5 µL of autoclaved deionized water, 12.5 µL of EconoTaq PLUS 2X Master Mix (Lucigen, Middleton, Wisconsin, USA), 2.5 µL of 10 µM ITS1-F and 2.5 µL of 10 µM ITS4. Thermal cycling conditions were as follows: initial denaturation of 95℃ for 5 min followed by 40 cycles of denaturation (95℃ for 90 s, annealing at 55℃ for 1 min, and extension at 72℃ for 90 s), and a final extension at 72℃ for 10 min. Amplification was confirmed by visualizing PCR products on a 1.7% agarose gel at 100 V for 30–35 min. Only samples that produced clear single bands were used for further analysis. Amplified products were cleaned and purified enzymatically using Exo-SAP IT (New England Biolabs, Ipswich, Massachusetts, USA). After purification, Sanger bidirectional sequencing was performed with Big-Dye Terminator v3.1 (Applied Biosystems, Foster City, California, USA) using the ITS1-F/ITS4 primers. Sequence reactions were cleaned using EDTA and ethanol and run on an ABI Prism 3730 Genetic Analyzer (Applied Biosystems). Cycle sequencing was performed in 10 µL reactions containing either the forward primer ITS1-F or the reverse primer ITS4 at 0.5 µM concentrations, 1 µL of cleaned PCR product, and 0.5 µL of 5X Sequencing Buffer (Applied Biosystems). Thermal cycling conditions for cycle sequencing reactions were as follows: initial denaturation at 96℃ for 1 min, and 35 cycles of denaturation (96℃ for 20 s), annealing (50℃ for 15 s), and extension (60℃ for 2 min). Sixty-one percent (432 of 709) of submitted EM root tips produced sequences.

Sanger sequences were analyzed in bulk, and initial editing, quality filtering, clustering, and taxonomic identification were processed using the Quantitative Insights into Microbial Ecology (QIIME; v1.9) bioinformatics pipeline (Caporaso et al., 2010). In brief, sequences were converted from .ab1 to .fasta and .qual file formats using Geneious v10 (Biomatters Ltd., Auckland, New Zealand). Sequences were then edited using the *add_qiime_labels.py* to modify the sample ID for all .fasta sequences and ensure uniqueness in .fasta labels. Files were preprocessed using the *convert_fastaqual_fastq.py*, and paired-end sequences per data set were merged using *join_paired_ends.py* with a minimum overlap of 10 bp. Joined sequences were processed separately per data set using *split_libraries_fastq.py* with a minimum Phred quality score of 25. Merged sequences were clustered into operational taxonomic units (OTUs) using *pick_open_reference_otus.py* (Rideout et al., 2014) and a 97% sequence similarity threshold with the *suppress_taxonomy_assignment* flag. Any resulting singleton OTUs were included, as these were representative of fungal species based on identified EM root tips. Taxonomic affiliations were assigned by searching representative sequences from each OTU against GenBank and UNITE+INSD databases using the BLAST option in *assign_taxonomy.py*. OTUs were cross-referenced by morphotype and verified as EM using the FUNGuild database (Nguyen et al., 2016). Sequences of all EM fungal OTUs were submitted to the GenBank database under accession numbers MN733425–MN733437.

Illumina sequencing and bioinformatics

DNA was extracted from a 0.25 g subsample of roots from each sample using the cetyltrimethylammonium bromide (CTAB) method following the procedure by Pec et al. (2017). Briefly, CTAB buffer (700 µL) was added to each sample, followed by 10 µL of proteinase K (600 mAU mL-1; QIAGEN, Mississauga, Ontario, Canada). Samples were incubated at 65℃ for 1 h, cooled to 21℃, and 600 µL of chloroform–isoamyl alcohol (24 : 1) was added to the sample tubes. The sample tubes were centrifuged for 5 min (17,000 g at 21℃). Aqueous supernatant was mixed with 600 µL of isopropanol and chilled at −20℃ for 2 h. The samples were centrifuged for 15 min and the supernatant was discarded. Five hundred microliters of 95% ethanol (v/v) was added to the pellet, vortexed, and centrifuged for 3 min; the same sequence was then repeated using 500 µL of 70% ethanol (v/v). Pellets were resuspended in 50 µL of nuclease-free water (Life Technologies, Carlsbad, California, USA). Following extraction, 1 ng of sample DNA was processed using the Illumina Nextera XT sample preparation kit (Illumina, San Diego, California, USA) following the manufacturer's protocol. This approach uses tagmentation to enzymatically fragment and tag the sample DNA with adapters in random positions. This step was followed by the ITS1 region (~10–1200 bp; Yang et al., 2018) of the extracted DNA being amplified using PCR and ITS1-F and ITS2 primers. Successful amplification was confirmed using gel electrophoresis (1.5% agarose gel). Amplification products were cleaned using the AxyPrep Mag PCR clean-up kit (Corning Life Sciences, Tewksbury, Massachusetts, USA), quantified fluorescently on a Qubit fluorometer (ThermoFisher Scientific, Grand Island, New York, USA), and pooled into equimolar concentrations. Amplicon library sequencing was performed on an Illumina MiSeq v3 (2 × 300 bp) at the Molecular Biological Sciences Facility, University of Alberta. Raw sequences are available at the National Center for Biotechnology Information (NCBI) database under BioProject number PRJNA591371.

Sequences were demultiplexed and quality filtered using Trimmomatic v0.36 (Bolger et al., 2014), removing Illumina adapters, sequences with <100 bp, and bases with minimum quality scores <2.0. All remaining forward and reverse reads were merged with the fastq-join method using the *join_paired_ends.py* script in QIIME (v1.9) (Caporaso et al., 2010) followed by ITS1 region extraction using ITSx v1.0.11 (Bengtsson-Palme et al., 2013). Sequences were clustered into OTUs at 97% sequence similarity while simultaneously removing chimeric sequences using the UPARSE-OTU algorithm (Edgar, 2013) with the *cluster_otus* command in USEARCH (v9.2.64) (Edgar, 2010). Singletons were excluded to reduce artificially inflating OTU richness due to sequencing error.

OTUs were taxonomically identified by searching representative sequences from each OTU against the UNITE fungal ITS database using the BLAST option in the *assign_taxonomy.py* script in QIIME. The OTUs were initially

assigned to three groups (pathogenic, saprotrophic, or EM) based on their genus affiliation, trophic mode, and functional guild as described in Branco et al. (2013), Tedersoo and Smith (2013), and Tedersoo et al. (2014) and using FUNGuild (Nguyen et al., 2016). The OTUs were placed into groupings (i.e., dark septate endophytic, EM, saprotrophic, and pathogenic) only if the assignments were deemed as highly probable or probable based on default parameters in the FUNGuild database (https://github.com/UMNFuN/FUNGuild). Fungal OTUs with an uncertain functional grouping (i.e., guild) that were assigned at least a genus affiliation were evaluated as "unresolved." Non-fungal OTUs were excluded from further analyses while fungal OTUs that either had no known identification in the available databases or were identified only to kingdom level (i.e., fungi) were evaluated as "unidentified."

Comparing fungi between soils and roots

Site description, sample collection, and preparation

Roots and soils were collected from pure jack pine forests north of Fort McMurray, Alberta, Canada (57°21′ 49.1″N, 111°25′45.6″W). Three sites (~1 ha each) were sampled for each of two disturbances: intact jack pine ("Control") and clearcut harvested ("Disturbed"). The Controls were mature stands growing on Eutric Brunisol soils with an understory dominated by *Arctostaphylos uva-ursi* (L.) Spreng. and *Vaccinium vitis-ideae* L. The Disturbed stands had been harvested approximately 17 years prior to sampling and currently had some natural jack pine regeneration. The soils, forest floor material, and vegetation of the Disturbed stands closely resembled those of the Control stands. Each site contained nine 2.5×2.5 m plots that were separated by at least 2 m. In August of 2016, all plots were sampled; one soil core (2.5 cm diameter, 20 cm depth) was taken at the corners and center of each plot for a total of five cores. This resulted in a total of 135 cores per disturbance (5 cores $\times$ 9 plots $\times$ 3 sites = 135). Soil cores were transported to the laboratory on ice, and roots were removed from cores and washed as described above. The bulk soil of the cores was sieved (<2 mm), and any remaining roots (>2 mm diameter), rocks, and coarse woody debris were removed. All subsamples were stored at −20°C until analysis for fungal community characterization.

Sequencing and bioinformatics

DNA was extracted separately from 0.25 g subsamples of roots and soils, and then sequenced using Illumina following the methods described above. Illumina was chosen for its ability to sequence mixed DNA template.

Data analysis

To address whether fungal community composition differed between the sequencing methods (i.e., Sanger and Illumina MiSeq), the presence or absence of fungal OTUs in root samples associated with each of the 30 trees was used to construct communities. One-way permutational multivariate analysis of variance (PERMANOVA) tests were conducted separately for EM fungal communities and total (dark septate endophytes [DSE], pathotrophic, saprotrophic, and EM) fungal communities. These analyses were run with 9999 permutations. Principal coordinate analysis (PCoA) was used to visualize differences between communities from each sequencing method and their within-group variation. To test the statistical significance of differences in the number of fungal OTUs (community richness) between the sequencing methods, generalized linear mixed models using Poisson error structures were calculated separately for the total and EM fungal communities. For these models, the sequencing method was used as a fixed factor, whereas the tree from which the samples originated was included as a random factor to account for potential within-sample autocorrelation.

To determine whether Illumina-sequenced fungal communities differed between roots and soil, and by forest disturbance (i.e., Control and Disturbed), the relative abundance of fungal OTUs was analyzed for two sets of sequence read data: raw (non-rarefied) and rarefied. The latter data set was rarefied to a depth of 4209 reads, which was the minimum total reads among samples. Communities were then constructed using the relative abundance of fungal OTUs calculated from OTU reads by pooling samples by site, such that there were three communities for each forest disturbance type. For each data set, two-way PERMANOVAs were conducted separately for total and EM fungal communities to test the statistical significance of habitat (roots versus soils) and disturbance main effects and habitat × disturbance interaction. Differences in community compositions between roots and soils and their within-habitat variation were visualized using PCoA. To test the statistical significance of differences in the number of fungal OTUs between roots and soils, and disturbances for each data set, generalized linear mixed models with Poisson error structures were conducted separately for the total and EM fungal communities. These models used habitat, disturbance, and habitat × disturbance interaction terms as fixed effects and sampling plot nested in disturbance as a random factor. Given a significant interaction term, pairwise simple effects were tested to determine statistical significance between groups. The classification of fungal OTUs as preferentially associated with either roots ("root abundant") or soils ("soil abundant"), common to both habitats ("common"), or too rare to be associated with either habitat ("rare") was determined using a multinomial species classification method (CLAM) with "super majority" rule

and all other parameters set as default (Chazdon et al., 2011). The CLAM analysis was conducted for each data set, rarefied and non-rarefied.

All analyses were conducted in the R software environment v3.6.0 (R Core Team, 2016). The PERMANOVA, PCoA, and CLAM analyses were performed using functions available in the R package "vegan" v2.5-5 (Oksanen et al., 2018). All generalized linear models were constructed and tested for significance using functions available in R packages "lme4" v1.1-21 (Bates et al., 2015) and "car" v3.0-2 (Fox and Weisberg, 2011).

RESULTS

Comparing sequencing methods—Sanger versus Illumina

Overall, 340 OTUs were identified between the two sequencing methods for the total root-associated fungal community (Appendix S1). Of those, 28 OTUs were identified using Sanger sequencing and 331 OTUs were identified using Illumina. Nineteen OTUs (5.6% of the overall total) were identified by both sequencing methods, leaving nine unique OTUs identified by Sanger sequencing and 313 unique OTUs identified by Illumina sequencing. Sixteen of the 28 OTUs identified by Sanger sequencing were also of the 28 most abundant OTUs identified by Illumina; however, the most abundant OTU identified by Illumina was not detected by Sanger sequencing (Appendix S2).

Sanger and Illumina sequencing generated total fungal communities (OTUs assigned as DSE, pathotrophic, saprotrophic, and EM guilds, and unidentified and unresolved groups) that significantly differed (PERMANOVA, $F_{(1,58)} = 14.30$, $P < 0.001$) in composition from each other (Figure 1A). The Illumina-sequenced communities were more similar in composition to each other than to the Sanger-sequenced communities, as indicated by less within-group/inter-sample variability (i.e., tighter clustering) for the Illumina group in the PCoA analysis (Figure 1A). The sequencing methods produced different numbers of fungal OTUs as the richness of the total fungal communities significantly differed (Wald $X^2_{(1)} = 994.09$, $P < 0.001$), with estimates of richness of the Illumina-sequenced communities being on average 3518% greater than those of the Sanger-sequenced communities (Figure 1B).

The sequencing methods identified fungal communities that differed in the relative abundance of functional guilds. Sanger sequencing identified EM taxa nearly exclusively, representing 93% of all OTUs produced by this method (Figure 2). However, Illumina sequencing identified taxa from a broader array of guilds, with DSE, pathotrophic, saprotrophic, and EM fungi present (Figure 2). Among these guilds, EM OTUs were most abundant (19.0% of all Illumina OTUs) followed by saprotrophic OTUs (11.8%) (Figure 2). However, OTUs representing unidentified or unresolved taxa

were numerous, with each representing 33.4% of all Illumina OTUs (Figure 2).

We observed similar patterns in communities of EM fungal OTUs as those observed in the total fungal community. Overall, a total of 71 EM OTUs were identified between the sequencing methods, with Illumina identifying the most OTUs (23 for Sanger and 64 for Illumina). Sixteen OTUs (23% of the overall total) were identified by both of the sequencing methods, leaving seven OTUs unique to Sanger and 48 unique to Illumina. Significant differences (PERMANOVA $F_{(1,56)} = 13.25$, $P < 0.001$) were detected between Sanger-sequenced and Illumina-sequenced EM fungal communities, with less within-group variability observed for the Illumina group (Figure 1C). The average richness of EM fungal communities identified by Illumina significantly differed (Wald $X^2_{(1)} = 211.05$, $P < 0.001$) from that of communities identified by Sanger sequencing, with Illumina communities being 518% richer (Figure 1D). The similarity between the average richness of EM fungal OTUs identified by Sanger sequencing ($\bar{x} = 2.52$ [±0.33 standard error (SE)]; Figure 1D) and that of the total fungal community ($\bar{x} = 2.63$ [±0.32 SE]; Figure 1B) was likely because OTUs for EM fungal taxa represented 93% of all fungal OTUs identified by this method.

Comparing fungi associated with roots and soils in intact and disturbed forests—Illumina sequencing

For the total fungal community (DSE, pathotrophic, saprotrophic, and EM guilds, and unidentified and unresolved groups), a total of 3966 OTUs (8,812,733 total raw reads) were detected between roots and soils and between the disturbance types, with the largest inter-sample variation (coefficient of variation [CV]) in OTUs and reads from the raw data set occurring from soils (Table 1). Rarefaction reduced the total number of OTUs by 68–75% and the total number of reads by 99%, depending on whether roots and soils were pooled or separated (Table 1). The large decline in total reads was due to rarefying to a minimum sample total (4209 reads) that was 87% smaller than the next largest total (33,401 reads). This minimum read total was from a soil sample; this is likely the cause of the large CV for that habitat alone and a contributing factor to the variation when roots and soils were pooled (Table 1).

The composition of the total fungal community significantly differed between roots and soils for the raw (PERMANOVA $F_{(1,8)} = 3.64$, $P = 0.003$; Figure 3A) and rarefied (PERMANOVA $F_{(1,8)} = 3.48$, $P = 0.005$; Figure 3C) data sets. The difference in composition between roots and soils was similar for each disturbance type as significant habitat × disturbance interactions were not detected for either data set. For both data sets, soil-based fungal communities showed smaller within-group variation in composition than root-based communities, as the PCoA analysis showed the tightest clustering of points for the former
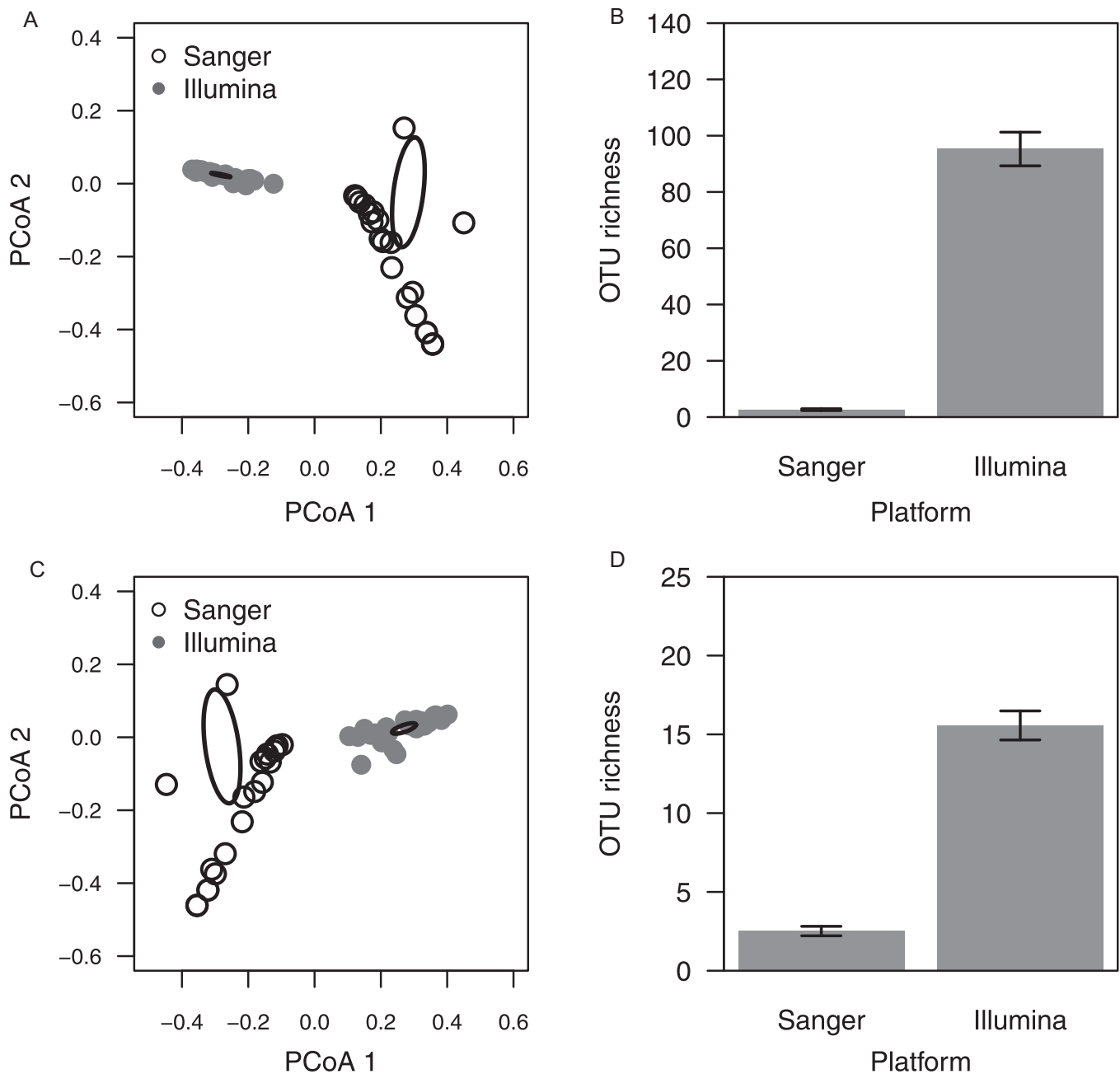
FIGURE 1 Variation in the composition (A, C) and richness (B, D) of operational taxonomic units (OTUs) in communities of all fungal taxa (total community; A, B) and ectomycorrhizal (EM) taxa (EM community; C, D) determined by Sanger (open circles) or Illumina MiSeq (high-throughput; closed circles) DNA sequencing methods. Black ellipses indicate the 95% confidence limits around cluster centroids determined by principal coordinate analysis (PCoA) in panels A and C. Bars in panels B and D represent group means (±standard errors)

group (Figures 3A, 3C). The difference between root and soil community composition in both data sets may have been due to a dominance of habitat-abundant taxa, as five or six of the 10 most abundant OTUs in each habitat were classified as either root- or soil-abundant while the other OTUs occurred in both habitats at largely similar amounts (Table 2).

The number of fungal OTUs (i.e., richness) of the total fungal community associated with soils and roots varied with disturbance type for both data sets, as indicated by a significant habitat × disturbance interaction for the raw (Wald $X^2_{(1)} = 1108.40$, $P < 0.001$; Figure 3B) and rarefied (Wald $X^2_{(1)} = 31.58$, $P < 0.001$; Figure 3D) data sets. Considering both data sets, the total fungal richness of roots and soils consistently differed from each other (Figures 3B, 3D). The richness of all fungal OTUs of roots ranged from 52% lower (rarefied data set, disturbed forest) to 93% greater (raw data set, intact forest) than that of soils (Figures 3B, 3D). Forest disturbance changed the pattern and magnitude of differences in richness of all fungal OTUs between roots

and soils. For the raw data set, disturbance was associated with a shift from roots having richer fungal communities than soils to having less-rich communities than soils (Figure 3B). However, for the rarefied data set, disturbance was associated with a greater difference between the richness of fungal communities from roots and soils than that observed in the intact forests (Figure 3D).
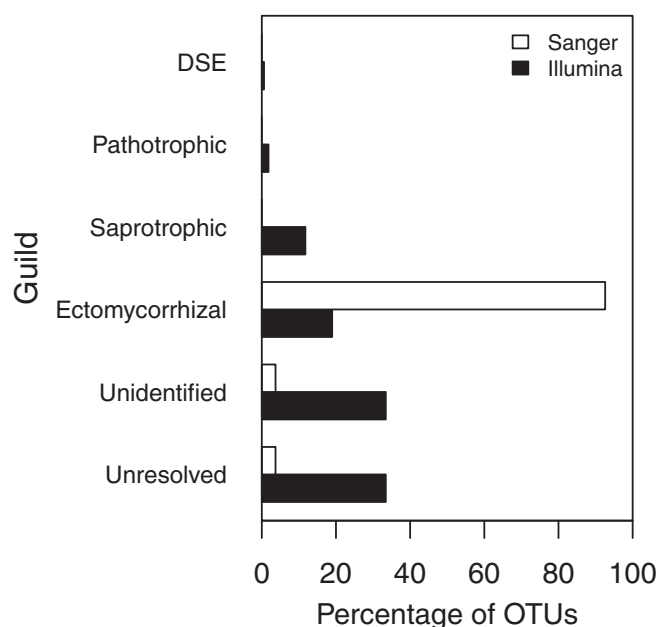

FIGURE 2 Percentages of operational taxonomic units (OTUs) out of all fungal OTUs produced by either Sanger (white bars) or Illumina MiSeq (black bars) DNA sequencing methods for six fungal guilds: dark septate endophytes (DSE), pathotrophic, saprotrophic, ectomycorrhizal, unidentified (taxonomic assignment limited to kingdom Fungi), and unresolved (taxonomic assignment limited to phylum). Guild assignments were made using the FUNGuild database (Nguyen et al., 2016)

The percentages of all fungal OTUs (out of the total number of OTUs for roots and soils separately) classified by FUNGuild (Nguyen et al., 2016) into the DSE, pathotrophic, saprotrophic, and EM guilds were similar between roots and soils for the raw (Figure 4A) and rarefied data sets (Figure 4B), with the EM and saprotrophic guilds having the highest percentages (Figure 4). The CLAM analysis indicated that the EM and saprotrophic guilds contained similar percentages of OTUs frequently identified from roots or soils (Table 3). Furthermore, these guilds were dominated by OTUs with low relative abundance (i.e., classified as "rare"). Fungal communities associated with both roots and soils were dominated by saprotrophic taxa, as indicated by the relative abundance of OTUs of the 10 most abundant taxa in each habitat (Table 2). Roots and soils had similar percentages of OTUs from the DSE and pathotrophic guilds (Figure 4A), both of which contained some root- or soil-abundant taxa (Table 3).

For both the raw and rarefied data sets, the majority of fungal OTUs associated with roots and soils either could not be assigned to a taxonomic group within kingdom Fungi (i.e., unidentified) or represented taxa whose taxonomy could not be resolved past their phylum (i.e., unresolved). The unidentified and unresolved groups each occurred in similar percentages in roots and soils (Figure 4). Each of these groups had similar amounts of root- and soil-abundant OTUs, with the majority of OTUs classified as rare taxa (Table 3).

Overall, a total of 605 EM OTUs (843,037 raw reads) were detected across the entire raw data set, with the soils having the largest CV. Rarefaction reduced the number of EM OTUs by 70–77% and reduced the total number of reads for EM taxa by 99% (Table 1), depending on whether roots and soils were pooled or separated (Table 1). EM fungal community composition significantly differed between soils and roots in the raw

TABLE 1 The total number of operational taxonomic units (OTUs) and total DNA sequence reads for fungal taxa produced by Illumina MiSeq for two fungal communities (total [all fungal OTUs] and ectomycorrhizal) and from two data sets (raw [non-rarefied] reads and rarefied reads). Total OTUs and reads were determined for two fungal habitats (roots and soil) both separately and pooled (total). Inter-sample variation is shown as coefficients of variation (CV)

| Community	Data set	Habitat	Total OTUs (CV)	Total reads (CV)
Total (all fungi)	Raw	Roots	3460 (18%)	5,537,023 (51%)
		Soils	3071 (51%)	3,275,710 (157%)
		Total	3966 (36%)	8,812,733 (94%)
	Rarefied	Roots	866 (22%)	25,254 (0%)
		Soils	932 (6%)	25,254 (0%)
		Total	1254 (30%)	50,508 (0%)
Ectomycorrhizal	Raw	Roots	537 (18%)	487,928 (97%)
		Soils	459 (52%)	355,109 (160%)
		Total	605 (36%)	843,037 (119%)
	Rarefied	Roots	123 (27%)	2231 (64%)
		Soils	133 (13%)	2656 (16%)
		Total	183 (33%)	4887 (42%)
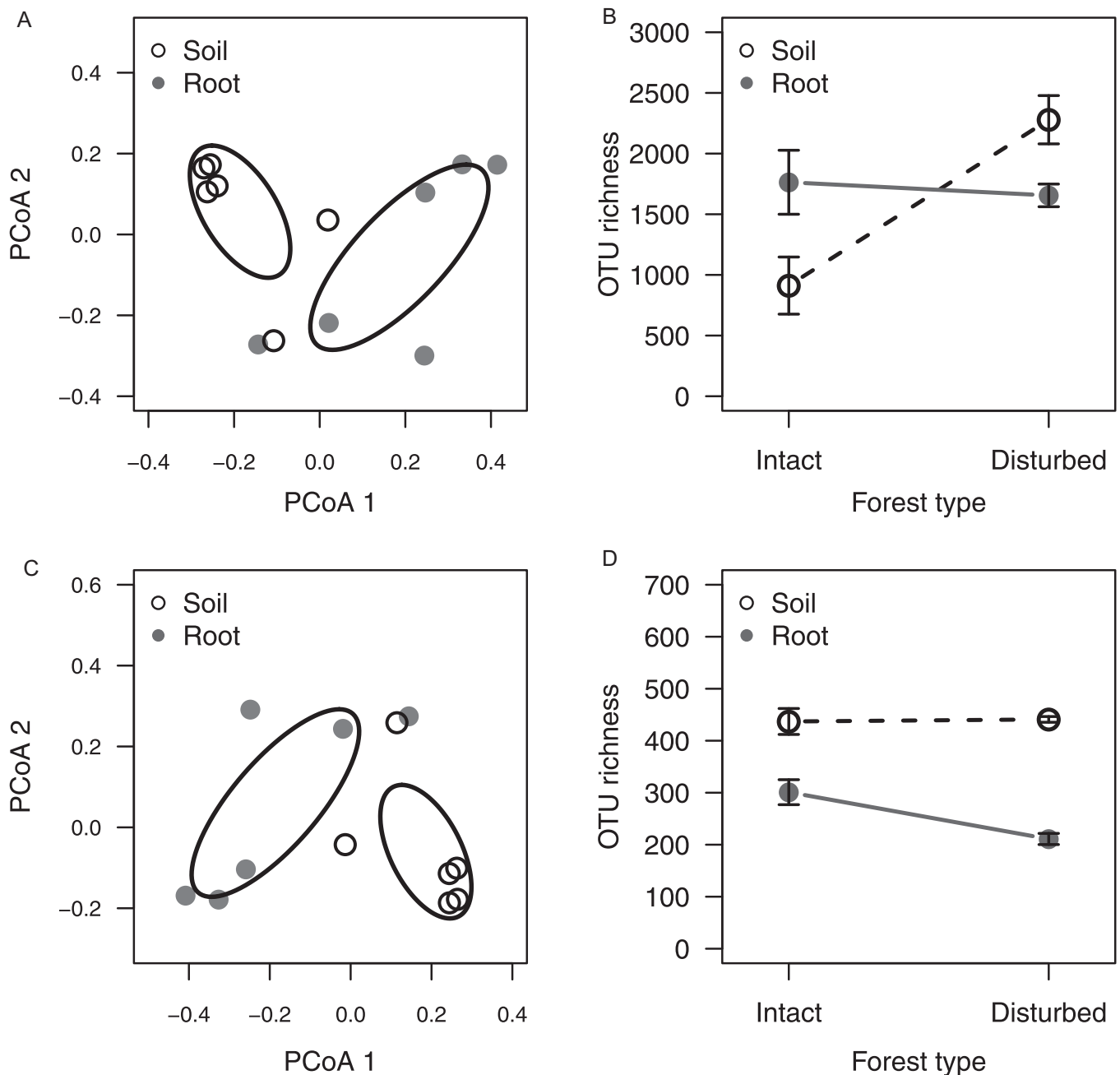
FIGURE 3 Variation in the composition (A, C) and richness (B, D) of operational taxonomic units (OTUs) in communities of all fungal (total community) OTUs produced by the Illumina MiSeq method for root (closed circles) and soil (open circles) habitats from intact and disturbed forests. Raw (A, B) and rarefied (C, D) sequence read data were analyzed separately. Black ellipses indicate the 95% confidence limits around cluster centroids determined by principal coordinate analysis (PCoA) in panels A and C. Panels B and D represent interaction plots showing group means (±standard errors)

(PERMANOVA, $F_{(1,8)} = 3.26$, $P = 0.002$; Figure 5A) and rarefied (PERMANOVA $F_{(1,8)} = 3.38$, $P = 0.003$; Figure 5C) data sets. In both data sets, soil-based communities showed less within-group variability than those from roots (Figures 5A, 5C). These differences were similar for both the Control and Disturbed groups as no significant habitat × disturbance effects were detected for either the raw or rarefied data set. The differences in community composition between roots and soils may have been due to the dominance (i.e., relative

abundance) of a few habitat-abundant OTUs, as two to five of the 10 most abundant OTUs in each habitat were classified as root- or soil-abundant (Table 4). Five or six of the other OTUs included among the 10 most abundant were common to both habitats and/or data sets and occurred at similar abundances between roots and soils for a given data set (Table 4).

The average number of EM fungal OTUs associated with roots and soils varied between the Control and Disturbed groups for the raw (Wald $X^2_{(1)} = 173.81$, $P < 0.001$;

TABLE 2 The 10 taxa with the highest relative abundance of sequence reads (for a given habitat) from communities of all fungal operational taxonomic units (OTUs) from root and soil habitats for raw (non-rarefied) and rarefied read data sets

Data set	Habitat	OTU	Relative abundance (%)[a,b]	Guild[c]	Classification[d]
Raw	Root	*Archaeorhizomyces* 1	15.14	Saprotrophic	Common
		Lachum 1	5.21	Saprotrophic	Root abundant
		Fungi 49	4.72	Unidentified	Root abundant
		Fungi 259	4.72	Unidentified	Root abundant
		Ascomycota 59	4.72	Unresolved	Root abundant
		Ascomycota 6	4.72	Unresolved	Root abundant
		Vibrisseaceae	3.45	Unresolved	Common
		Archaeorhizomyces	3.03	Saprotrophic	Common
		Fungi 597	3.17	Unidentified	Common
		Lachnum pygmaeum	1.69	Saprotrophic	Soil abundant
	Soil	*Archaeorhizomyces* 1	12.12	Saprotrophic	Common
		Lachnum pygmaeum	4.27	Saprotrophic	Soil abundant
		Vibrisseaceae	2.95	Unresolved	Common
		Fungi 597	2.56	Unidentified	Common
		Archaeorhizomyces	2.56	Saprotrophic	Common
		Helotiales 19	2.48	Unresolved	Soil abundant
		Helotiales 54	2.48	Unresolved	Soil abundant
		Helotiales 246	2.48	Unresolved	Soil abundant
		Agaricomycetes 210	2.48	Unresolved	Soil abundant
		Lachnum 1	2.06	Saprotrophic	Root abundant
Rarefied	Root	*Archaeorhizomyces* 1	14.78	Saprotrophic	Common
		Lachnum 1	4.91	Saprotrophic	Root abundant
		Ascomycota 59	4.33	Unresolved	Root abundant
		Fungi 49	4.19	Unidentified	Root abundant
		Ascomycota 6	4.07	Unresolved	Root abundant
		Fungi 259	3.95	Unidentified	Root abundant
		Vibrisseaceae	3.23	Unresolved	Common
		Archaeorhizomyces	3.18	Saprotrophic	Common
		Fungi 597	3.07	Unidentified	Common
		Amphinema 2	1.36	Ectomycorrhizal	Common
	Soil	*Archaeorhizomyces* 1	13.21	Saprotrophic	Common
		Lachnum pygmaeum	4.47	Saprotrophic	Soil abundant
		Vibrisseaceae	3.25	Unresolved	Common
		Fungi 597	3.05	Unidentified	Common
		Archaeorhizomyces	2.87	Saprotrophic	Common
		Lachnum 1	2.11	Saprotrophic	Soil abundant
		Helotiales 246	1.72	Unresolved	Soil abundant

(Continues)

TABLE 2 (Continued)

Data set	Habitat	OTU	Relative abundance (%)[a,b]	Guild[c]	Classification[d]
		Helotiales 54	1.68	Unresolved	Soil abundant
		Agaricomycetes 210	1.68	Unresolved	Soil abundant
		Helotiales 19	1.67	Unresolved	Soil abundant

[a]Sequence reads were produced by the Illumina MiSeq method.

[b]Percentage of total reads from root (5,537,023 raw reads, 16,878 rarefied reads) or soil (3,275,710 raw reads, 16,878 rarefied reads) habitats.

[c]Taxa guilds were determined using the FUNGuild database (Nguyen et al., 2016).

[d]Classification as determined by multinomial species classification performed separately for the raw and rarefied data.
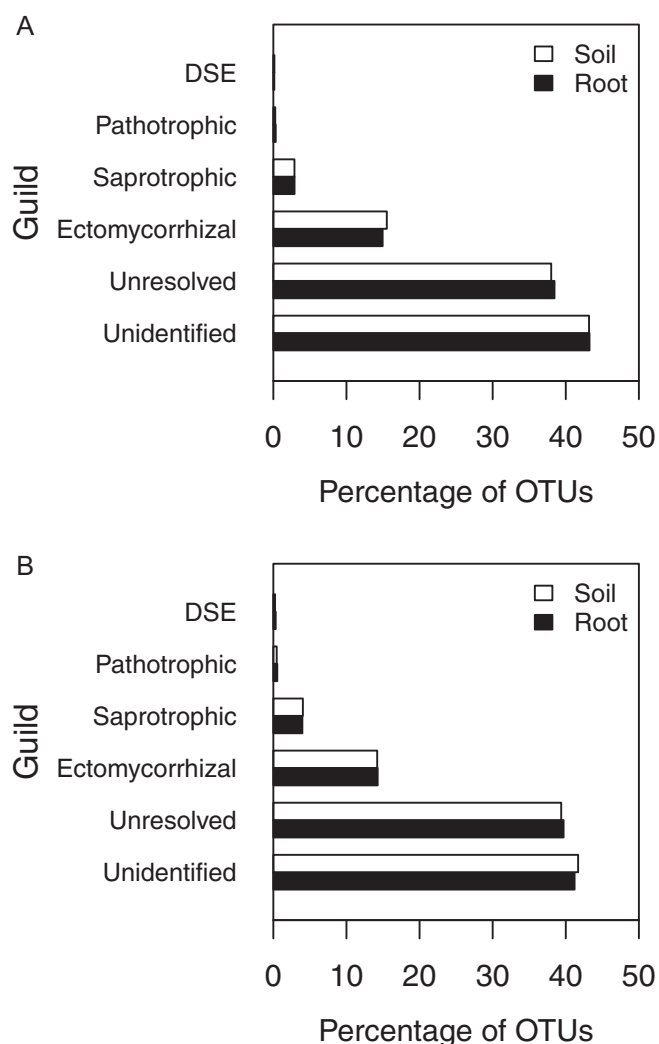

FIGURE 4 The percentage of fungal operational taxonomic units (OTUs) out of all fungal OTUs from either root (white bars) or soil (black bars) habitats for six functional guilds: dark septate endophytes (DSE), pathotrophic, saprotrophic, ectomycorrhizal, unidentified (taxonomic assignment limited to kingdom Fungi), and unresolved (taxonomic assignment limited to phylum). Percentages were calculated for data sets of raw (non-rarefied; A) and rarefied (B) sequence reads produced by the Illumina MiSeq method. Guild assignments were made using the FUNGuild database (Nguyen et al., 2016)

Figure 5B) and rarefied (Wald $X_{(1)}^2$ = 5.75, $P = 0.017$; Figure 5D) data sets, as indicated by significant habitat × disturbance interactions for OTU richness. Considering both data sets, EM fungal richness of roots and soils consistently differed from each other in the comparisons shown in Figures 5B and 5D. These comparisons show that the richness of EM fungal OTUs associated with roots ranged from 57% lower (rarefied data set, disturbed forest) to 104% greater (raw data set, intact forest) than that of the soils (Figures 5B, 5D). Forest disturbance changed the pattern and magnitude of differences of EM fungal richness between habitats. For the raw data set, disturbance was associated with a shift from the roots having richer EM fungal communities than soils to having less rich communities than soils (Figure 5B). However, for the rarefied data set, disturbance was associated with a greater difference between the richness of fungal communities from root and soils (Figure 5D).

DISCUSSION

We first investigated the influence of sequencing method on fungal community composition by comparing root-associated fungal communities identified by Sanger and Illumina (high-throughput) sequencing. Regardless of sequencing method, EM fungi were the dominant functional guild, and similar taxa made up the common OTUs. OTU richness was higher for Illumina sequencing; this was likely due to: (1) Illumina sequencing is not inhibited by species co-occurrence (i.e., mixed DNA template), (2) sequencing depth is much higher than that of Sanger (Sanger: 432 sequences across 30 samples; Illumina: 1,018,520 sequences across the same 30 samples), and (3) DNA from fine roots, including EM root tips, was extracted and likely increased the pool of root-associated fungi. Extrapolating the species accumulation curve of Sanger sequences to the number of samples submitted (i.e., 709 EM root tips), estimates of OTU richness are similar between the two methods (Sanger: mean = 97, SE = 0.3; Illumina: mean = 102, SE = 10.3)

TABLE 3 The percentage of each fungal guild or group determined using a multinomial species classification method (CLAM) as root-abundant (highest abundance on roots), soil-abundant (highest abundance in soil), common (found at a similar abundance on roots and in soil), or rare (abundance too low to classify). The class percentages were calculated separately from the total number of OTUs for each guild. CLAM analyses and percentage calculations were performed separately for the raw (non-rarefied) and rarefied data sets

Data set	Guild	Percentage of total fungal OTUs			
		Root-Abundant	Soil-Abundant	Common	Rare
Raw	Dark septate endophytes (5 OTUs)	20.00	40.00	40.00	0.00
	Pathotrophic (11 OTUs)	18.18	9.09	45.45	27.27
	Saprotrophic (115 OTUs)	18.26	16.52	26.96	38.26
	Ectomycorrhizal (605 OTUs)	15.54	15.37	22.98	46.12
	Unresolved (1512 OTUs)	16.47	17.79	22.22	43.52
	Unidentified (1718 OTUs)	15.19	16.65	22.41	45.75
Rarefied	Dark septate endophytes (3 OTUs)	0.00	33.33	33.33	33.33
	Pathotrophic (5 OTUs)	0.00	0.00	20.00	80.00
	Saprotrophic (45 OTUs)	6.67	4.44	28.89	60.00
	Ectomycorrhizal (183 OTUs)	5.46	3.28	15.30	75.96
	Unresolved (493 OTUs)	5.88	5.48	11.56	77.08
	Unidentified (525 OTUs)	4.76	4.95	13.33	76.95

(Appendix S3). Thus, if the success rate of Sanger sequencing were improved through, for example, quality control procedures at critical stages in construction, sequence assembly, and annotation (Ma and Fedorova, 2010), this method is likely comparable to Illumina in capturing OTU richness. Furthermore, Sanger's shallow sequencing depth caused differences in fungal community composition, and inter-sample variability was higher in Sanger-sequenced versus Illumina-sequenced communities. The distinct pattern of inter-sample variability is the law of large numbers at work: increasing the sample size lowers the sample variance. Because ecological communities have inverse J-shaped species abundance distributions, intensely sampling the common species (i.e., what is effectively done through Illumina sequencing) will result in community convergence. If detecting shifts in common EM fungal species colonizing roots is the study goal, Sanger sequencing may be adequate. However, given the inter-sample variability, a higher number of Sanger-sequenced samples would be required to detect these shifts compared to those generated by Illumina. If capturing fungal diversity is the study goal, especially across functional guilds, Illumina is the method of choice. However, most of the diversity in OTUs generated by Illumina had "unidentified" or "unresolved" assignments. Without taxonomic assignment, the function or biology of species comprising fungal communities cannot be inferred or interpreted. Importantly, richness and composition of Sanger- and Illumina-sequenced communities should not be compared without adjusting for sampling intensity.

The next choice in methods we explored was how sampling roots or soils may affect fungal community composition. We found that roots and soils yielded different fungal communities, and this remained consistent regardless of forest disturbance, rarefaction of sequences, and fungal community (total versus EM fungi). Furthermore, we found that fungal communities in soils were more clustered in composition than those associated with roots, suggesting that soil conditions may act as stronger environmental filters than roots. Fungi occupying roots may be buffered from soil heterogeneity (Beck et al., 2015; Goldmann et al., 2016). The deterministic processes underlying community assembly emerging from environmental variation in soils may be weakened by the presence of living roots, for which colonization is mostly governed by stochastic priority effects (Kennedy et al., 2009). Alternatively, when regional species pools are much greater in richness than local communities, the influence of stochasticity can be greater in local communities (Chase, 2003; Chase and Myers, 2011). However, given the inconsistent patterns in fungal richness between roots and soils in our study, it is difficult to assess where the higher diversity truly lies. As part of investigating the differences in fungal community composition between roots and soils, we also characterized the response of root-based and soil-based fungi to forest clearing, to test whether this disturbance decouples the two community types. We found that for both total and EM fungi, community composition did not increasingly diverge with disturbance, rather the distinctness of fungal communities between roots and soils persisted and represented two "views" of belowground fungi.

The final choice in methods we explored was the effect of rarefaction on fungal communities. In our case, the sample on which rarefaction was based (4209 reads) had 13% of the reads than the next smallest sample (33,401
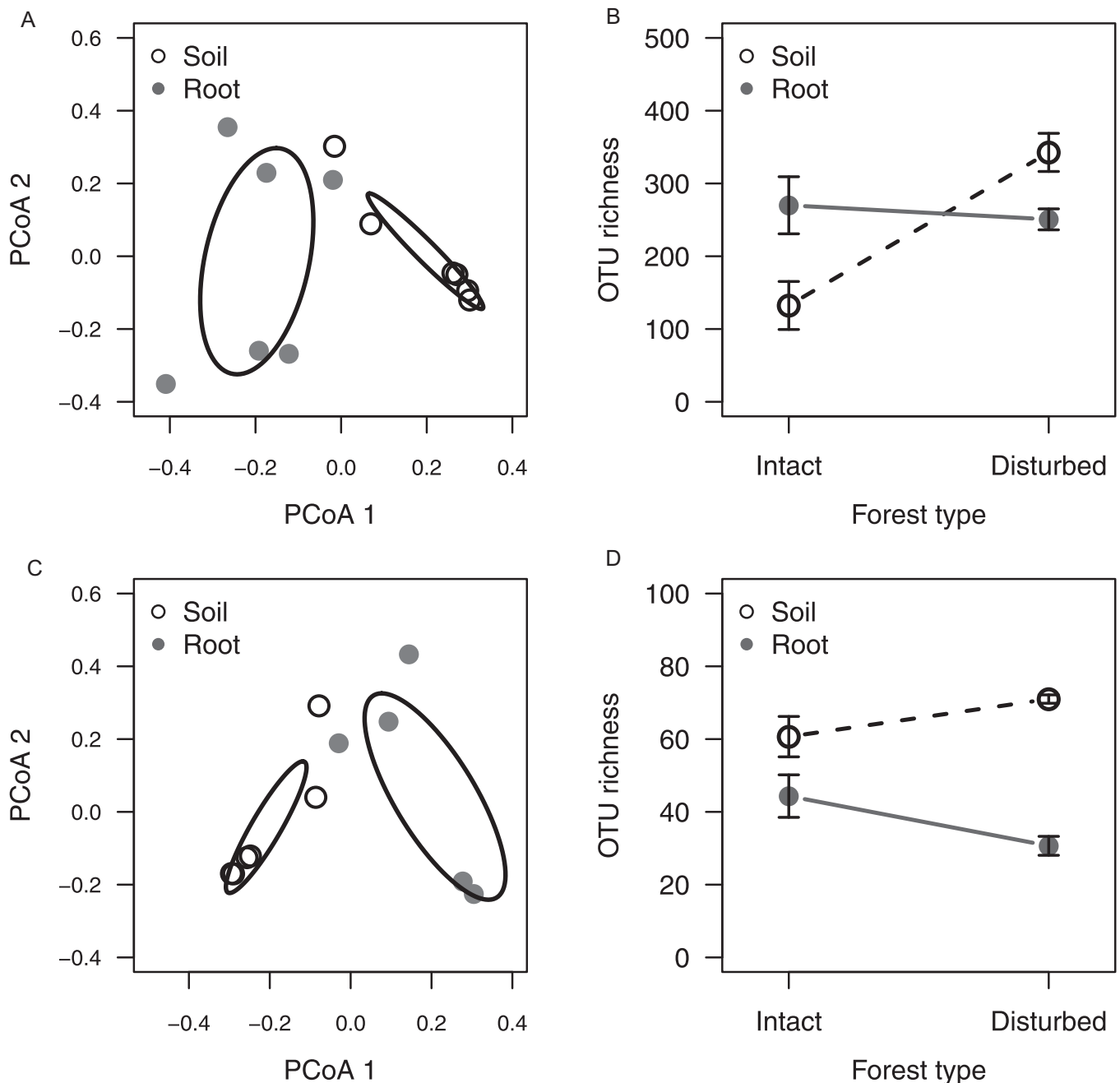
FIGURE 5   Variation in the composition (A, C) and richness (B, D) of operational taxonomic units (OTUs) in communities of ectomycorrhizal (EM) fungal (EM community) OTUs produced by the Illumina MiSeq method for root (closed circles) and soil (open circles) habitats from intact and disturbed forests. Raw (A, B) and rarefied (C, D) sequence read data were analyzed separately. Black ellipses indicate the 95% confidence limits around cluster centroids determined by principal coordinate analysis (PCoA) in panels A and C. Panels B and D represent interaction plots showing group means (±standard errors)

reads). Although rarefying had the desirable effect of normalizing read counts across samples, it had a sizeable influence on trends in OTU richness to the extent that patterns may be distorted. Rarefying sequence reads that have large inter-sample variability may affect the among/between-group patterns observed (McMurdie and Holmes, 2014). For total and EM fungal communities, rarefaction resulted in changes to between-group patterns in

OTU community richness. In our study, where sequencing depth varied substantially among soil samples, rarefaction showed a completely different response of fungal communities to disturbance. Without rarefaction, soil fungal communities increased in richness with disturbance, whereas with rarefaction, richness remained stable. We suggest that researchers should carefully evaluate inter-sample variability before deciding whether to rarefy data.

TABLE 4 The 10 taxa with the highest relative abundance of sequence reads (for a given habitat) from communities of ectomycorrhizal fungal operational taxonomic units (OTUs) from root and soil habitats for raw (non-rarefied) and rarefied read data sets

Data set	Habitat	OTU	Relative abundance (%)[a,b]	Classification[c]
Raw	Root	*Amphinema* 2	1.36	Common
		Russulaceae 123	1.24	Root abundant
		Russulaceae 278	1.24	Root abundant
		Tricholoma	0.97	Common
		Amphinema 1	0.36	Common
		Cortinarius 4	0.36	Root abundant
		Russulaceae 207	0.20	Root abundant
		Hygrophorus	0.19	Common
		Cortinarius 74	0.19	Common
		Russulaceae 240	0.19	Common
	Soil	*Acephala applanata*	1.66	Soil abundant
		Cortinarius 145	1.54	Soil abundant
		Amphinema 2	0.76	Common
		Amphinema 1	0.70	Common
		Tricholoma	0.62	Common
		Russulaceae 17	0.48	Soil abundant
		Russulaceae 123	0.45	Root abundant
		Russulaceae 278	0.45	Root abundant
		Hygrophorus	0.21	Common
		Cortinarius 74	0.21	Common
Rarefied	Root	*Amphinema* 2	1.36	Common
		Russulaceae 278	1.06	Common
		Tricholoma	1.01	Common
		Russulaceae 123	0.98	Common
		Gymnomyces 1	0.48	Root abundant
		Amphinema 1	0.36	Common
		Cortinarius 4	0.28	Root abundant
		Russulaceae 240	0.19	Common
		Russulaceae 291	0.19	Common
		Acephala applanata	0.18	Soil abundant
	Soil	*Acephala applanata*	1.20	Soil abundant
		Cortinarius 145	1.14	Soil abundant
		Amphinema 2	1.03	Common
		Russulaceae 123	0.81	Common
		Russulaceae 278	0.78	Common

(Continues)

TABLE 4 (Continued)

Data set	Habitat	OTU	Relative abundance (%)[a,b]	Classification[c]
		Tricholoma	0.58	Common
		Amphinema 1	0.44	Common
		Russulaceae 17	0.36	Soil abundant
		Russulaceae 299	0.29	Soil abundant
		Russulaceae 291	0.25	Common

[a]Percentage of total reads from root (5,537,023 raw reads, 25,254 rarefied reads) or soil (3,275,710 raw reads, 25,254 rarefied reads) habitats.
[b]Sequence reads were produced by the Illumina MiSeq method.
[c]Classification as determined by a multinomial species classification method performed separately for the raw and rarefied data.

For data sets with large inter-sample variability such as ours, researchers may want to consider analyzing both raw and rarefied data, then presenting both results if the data sets differ or showing only the rarefied results if no differences are observed. While rarefied and non-rarefied data sets generated different patterns for community richness, patterns in the composition of fungal communities and functional guilds between intact and disturbed forests were similar. Although the majority of taxa in fungal communities are "rare," the observed patterns are primarily driven by the most dominant taxa. This suggests that when choosing between numbers of samples and sequencing depth, investing in the former will provide more power to detect changes in community composition or turnover.

CONCLUSIONS

We reviewed outcomes of choices on sampling, sequencing, and analyzing DNA from plant root symbionts (i.e., mycorrhizal fungi) on fungal community composition and OTU richness. As roots are habitat for fungi belonging to other trophic guilds (e.g., Unuk et al., 2019), we also considered effects on endophytic, saprotrophic, pathogenic, and unidentified fungi. We first investigated the influence of sequencing method on fungal community composition by comparing root-associated fungal communities identified by Sanger and Illumina (high-throughput) sequencing. We found that by selecting EM roots and Sanger sequencing the fungi colonizing those roots, we were able to target EM fungi and avoid those belonging to other trophic guilds. Illumina sequencing captured a wide range of trophic groups; however, many sequences were uninformative because the sequences could not be matched to an identified species of fungi. When accounting for differences in sequencing depth, the two sequencing methods were comparable. To detect shifts in common fungi comprising communities, more

samples would be required for Sanger versus Illumina sequencing. Next, we investigated how sampling roots versus soils might impact community composition and found that the two habitats capture relatively distinct fungal communities. Higher inter-sample variability in community composition among root versus soil samples indicates that more sampling would be necessary to capture community shifts of root-associated fungi. Dissimilarity between root- and soil-associated fungal communities was insensitive to forest disturbance (harvesting), suggesting the two communities change in composition to the same extent with recovery. In other words, root-associated fungal communities do not seem more sensitive than soil-associated fungal communities to forest disturbance, despite the death of trees (Pec and Cahill, 2019). Finally, we found that samples highly variable in sequencing depth are particularly sensitive to rarefaction with respect to estimating richness. Researchers may want to carefully weigh rarefying (Haegeman et al., 2013), excluding samples, or incorporating other normalization and transformation methods (McMurdie and Holmes, 2013; Love et al., 2014) when working with samples that are extremely unequal in sequencing depth.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

S.M.L. and J.K. planned and designed the research; N.S. and G.J.P. conducted field and laboratory work; N.S., G.J.P., and J.A.C. analyzed the data; J.A.C. wrote the manuscript. All authors contributed equally to editing the text and approved the final version of the manuscript.

DATA AVAILABILITY STATEMENT

Sequences of all EM fungal OTUs were submitted to the GenBank database under accession numbers MN733425–MN733437. Raw sequences yielded from Illumina sequencing are available at the National Center for Biotechnology Information (NCBI) database under BioProject number PRJNA591371.

REFERENCES

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.

Beck, S., J. R. Powell, B. Drigo, J. W. G. Cairney, and I. C. Anderson. 2015. The role of stochasticity differs in the assembly of soil- and root-associated fungal communities. *Soil Biology and Biochemistry* 80: 18–25.

Bellemain, E., T. Carlsen, C. Brochmann, E. Coissac, P. Taberlet, and H. Kauserud. 2010. ITS as an environmental DNA barcode for fungi: An in silico approach reveals potential PCR biases. *BMC Microbiology* 10: 189.

Bengtsson-Palme, J., M. Ryberg, M. Hartmann, S. Branco, Z. Wang, A. Godhe, P. D. Wit, et al. 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution* 4: 914–919.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Branco, S., T. D. Bruns, and I. Singleton. 2013. Fungi at a small scale: Spatial zonation of fungal assemblages around single trees. *PLoS ONE* 8: e78295–e78295.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, et al. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7: 335–336.

Chase, J. M. 2003. Community assembly: When should history matter? *Oecologia* 136: 489–498.

Chase, J. M., and J. A. Myers. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society of London B, Biological Sciences* 366: 2351–2363.

Chazdon, R. L., A. Chao, R. K. Colwell, S.-Y. Lin, N. Norden, S. G. Letcher, D. B. Clark, et al. 2011. A novel statistical method for classifying habitat generalists and specialists. *Ecology* 92: 1332–1343.

Clasen, L. A., A. P. Detheridge, J. Scullion, and G. W. Griffith. 2020. Soil stabilisation for DNA metabarcoding of plants and fungi: Implications for sampling at remote locations or via third-parties. *Metabarcoding and Metagenomics* 4: e58365.

Clemmensen, K. E., A. Bahr, O. Ovaskainen, A. Dahlberg, A. Ekblad, H. Wallander, J. Stenlid, et al. 2013. Roots and associated fungi drive long-term carbon sequestration in boreal forest. *Science* 339: 1615–1618.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460–2461.

Edgar, R. C. 2013. UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* 10: 996–998.

Fox, J., and S. Weisberg. 2011. An {R} companion to applied regression, 2nd ed. Sage Publishing, Thousand Oaks, California, USA.

Gardes, M., and T. D. Bruns. 1993. ITS primers with enhanced specificity for basidiomycetes: Application to the identification of mycorrhizae and rusts. *Molecular Ecology* 2: 113–118.

Goldmann, K., K. Schröter, R. Pena, I. Schöning, M. Schrumpf, F. Buscot, A. Polle, and T. Wubet. 2016. Divergent habitat filtering of root and soil fungal communities in temperate beech forests. *Scientific Reports* 6: 31439.

Goodman, D. M., D. M. Durall, J. A. Trofymow, and S. M. Berch. 1996–2009. A manual of concise descriptions of North American ectomycorrhizae including microscopic and molecular characterization. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, British Columbia, Canada, and British Columbia Ministry of Forests, Victoria, British Columbia, Canada.

Guerrieri, A., A. Bonin, T. Münkemüller, L. Gielly, W. Thuiller, and G. F. Ficetola. 2021. Effects of soil preservation for biodiversity monitoring using environmental DNA. *Molecular Ecology* 30: 3313–3325.

Haegeman, B., J. Hamelin, J. Moriarty, P. Neal, J. Dushoff, and J. S. Weitz. 2013. Robust estimation of microbial diversity in theory and in practice. *The ISME Journal* 7: 1092–1101.

Johnson, N. C., C. Gehring, and J. Jansa. 2017. Mycorrhizal mediation of soil: Fertility, structure, and carbon storage. Elsevier, Amsterdam, The Netherlands.

Kennedy, P. G., K. G. Peay, and T. D. Bruns. 2009. Root tip competition among ectomycorrhizal fungi: Are priority effects a rule or an exception? *Ecology* 90: 2098–2107.

Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15: 550.

Ma, L.-J., and N. D. Fedorova. 2010. A practical guide to fungal genome projects: Strategy, technology, cost and completion. *Mycology* 1: 9–24.

McMurdie, P. J., and S. Holmes. 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8: e61217.

McMurdie, P. J., and S. Holmes. 2014. Waste not, want not: Why rarefying microbiome data is inadmissible. *PLoS Computational Biology* 10: e1003531.

Nguyen, N. H., Z. Song, S. T. Bates, S. Branco, L. Tedersoo, J. Menke, J. S. Schilling, and P. G. Kennedy. 2016. FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecology* 20: 241–248.

Nilsson, R. H., S. Anslan, M. Bahram, C. Wurzbacher, P. Baldrian, and L. Tedersoo. 2019. Mycobiome diversity: High-throughput sequencing and identification of fungi. *Nature Reviews Microbiology* 17: 95–109.

Oksanen, J., G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. J. McGlinn, P. Minchin, et al. 2018. vegan: Community ecology package. R package version 2.5-2.

Pauvert, C., M. Buée, V. Laval, V. Edel-Hermann, L. Fauchery, A. Gautier, I. Lesur, et al. 2019. Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal Ecology* 41: 23–33.

Pec, G. J., and J. James F. Cahill. 2019. Large-scale insect outbreak homogenizes the spatial structure of ectomycorrhizal fungal communities. *PeerJ* 7: e6895.

Pec, G. J., J. Karst, D. L. Taylor, P. W. Cigan, N. Erbilgin, J. E. K. Cooke, S. W. Simard, and J. F. Cahill Jr. 2017. Change in soil fungal community structure driven by a decline in ectomycorrhizal fungi following a mountain pine beetle (*Dendroctonus ponderosae*) outbreak. *New Phytologist* 213: 864–873.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rideout, J. R., Y. He, J. A. Navas-Molina, W. A. Walters, L. K. Ursell, S. M. Gibbons, J. Chase, et al. 2014. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ* 2: e545.

Shemesh, H., B. E. Boaz, C. I. Millar, and T. D. Bruns. 2020. Symbiotic interactions above treeline of long-lived pines: Mycorrhizal advantage of limber pine (*Pinus flexilis*) over Great Basin bristlecone pine (*Pinus longaeva*) at the seedling stage. *Journal of Ecology* 108: 908–916.

Slatko, B. E., A. F. Gardner, and F. M. Ausubel. 2018. Overview of next-generation sequencing technologies. *Current Protocols in Molecular Biology* 122: e59.

Smith, S. E., and D. J. Read. 2008. Mycorrhizal symbiosis, 3rd ed. Academic Press, London, United Kingdom.

Tedersoo, L., and M. E. Smith. 2013. Lineages of ectomycorrhizal fungi revisited: Foraging strategies and novel lineages revealed by sequences from belowground. *Fungal Biology Reviews* 27: 83–99.

Tedersoo, L., M. Bahram, S. Polme, U. Koljalg, N. S. Yorou, R. Wijesundera, L. V. Ruiz, et al. 2014. Global diversity and geography of soil fungi. *Science* 346: 1078.

Unuk, T., T. Martinović, D. Finžgar, N. Šibanc, T. Grebenc, and H. Kraigher. 2019. Root-associated fungal communities from two phenologically contrasting silver fir (*Abies alba* Mill.) groups of trees. *Frontiers in Plant Science* 10: 214.

Yang, R.-H., J.-H. Su, J.-J. Shang, Y.-Y. Wu, Y. Li, D.-P. Bao, and Y.-J. Yao. 2018. Evaluation of the ribosomal DNA internal transcribed spacer (ITS), specifically ITS1 and ITS2, for the analysis of fungal diversity by deep sequencing. *PLoS ONE* 13: e0206428.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Appendix S1. Grouping assignments for fungi identified between Sanger and Illumina MiSeq sequencing platforms.

Appendix S2. Rank abundance curves of (A) all fungi (28 total fungal operational taxonomic units [OTUs]) identified by Sanger sequencing and (B) the 28 most abundant fungi identified by Illumina MiSeq sequencing.

Appendix S3. An extrapolated species accumulation curve of Sanger sequences depicting the relationship between the number of samples collected (generated over 709 sampling iterations) and the number of fungal operational taxonomic units (OTUs) estimated using the R package "vegan" (Oksanen et al., 2018). Gray lines represent 95% confidence interval, while the solid black line represents extrapolated richness.