

RESEARCH ARTICLE

Effort versus Reward: Preparing Samples for Fungal Community Characterization in High-Throughput Sequencing Surveys of Soils

Zewei Song^{1*}, Dan Schlatter¹, Peter Kennedy², Linda L. Kinkel¹, H. Corby Kistler^{1,3}, Nhu Nguyen², Scott T. Bates¹

1 Department of Plant Pathology, University of Minnesota, Saint Paul, MN 55108, United States of America,

2 Department of Plant Biology, University of Minnesota, Saint Paul, MN 55108, United States of America,

3 USDA ARS Cereal Disease Laboratory, Saint Paul, MN 55108, United States of America

* songx208@umn.edu



OPEN ACCESS

Citation: Song Z, Schlatter D, Kennedy P, Kinkel LL, Kistler HC, Nguyen N, et al. (2015) Effort versus Reward: Preparing Samples for Fungal Community Characterization in High-Throughput Sequencing Surveys of Soils. *PLoS ONE* 10(5): e0127234. doi:10.1371/journal.pone.0127234

Academic Editor: Gabriele Berg, Graz University of Technology (TU Graz), AUSTRIA

Received: December 12, 2014

Accepted: April 12, 2015

Published: May 14, 2015

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. All the data we used in this manuscripts can be calculated based on the OTU table in the supporting materials. All richness data used in statistical analysis can also be found in the supporting materials. The data can be found in University of Minnesota Data Repository through this link: <http://conservancy.umn.edu/handle/11299/170817>

Funding: Funding was received from United States Department of Agriculture (2011-67019-30200 <http://nifa.usda.gov/grants>) grant to LLK & HCK; University

Abstract

Next generation fungal amplicon sequencing is being used with increasing frequency to study fungal diversity in various ecosystems; however, the influence of sample preparation on the characterization of fungal community is poorly understood. We investigated the effects of four procedural modifications to library preparation for high-throughput sequencing (HTS). The following treatments were considered: 1) the amount of soil used in DNA extraction, 2) the inclusion of additional steps (freeze/thaw cycles, sonication, or hot water bath incubation) in the extraction procedure, 3) the amount of DNA template used in PCR, and 4) the effect of sample pooling, either physically or computationally. Soils from two different ecosystems in Minnesota, USA, one prairie and one forest site, were used to assess the generality of our results. The first three treatments did not significantly influence observed fungal OTU richness or community structure at either site. Physical pooling captured more OTU richness compared to individual samples, but total OTU richness at each site was highest when individual samples were computationally combined. We conclude that standard extraction kit protocols are well optimized for fungal HTS surveys, but because sample pooling can significantly influence OTU richness estimates, it is important to carefully consider the study aims when planning sampling procedures.

Introduction

Terrestrial ecosystems are inhabited by an astonishing diversity of soil microorganisms [1], which play critical roles in global geochemical cycles and ecosystem functioning [2,3]. A large portion of this diversity was essentially hidden from microbial ecologists until DNA-based techniques eliminated the constraints of culture-dependent community surveys [4]. Since then, the technology for sequencing DNA has advanced at a staggering rate, with progress currently allowing for the generation of millions of microbial community sequences from environmental samples in very short periods of time [5].

of Minnesota MnDrive (<http://mndrive.umn.edu/>) grant to LLK & STB; and National Science Foundation (Grants EF 12-41895 grant to LLK and Long-Term Ecological Research Network 0620652). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Soil fungi are an important component of terrestrial microbial diversity due to their considerable influence on aboveground biodiversity and primary productivity, symbiotic relationships with most land plants, and as drivers of the global soil carbon cycle [3,6–8]. The number of publications using high-throughput sequencing (HTS) to characterize fungal communities in soil is growing exponentially (S1 Fig), and these studies reveal a high degree of diversity and endemism as well as functional redundancy among these fungi [9–11]. Although general outlines for the fungal community HTS workflow have been published [12], few studies have examined the protocol used in preparing samples for HTS, with some notable exceptions [13–15]. Such considerations are important, as standardization of HTS procedures will allow for cross-study comparisons of microbial HTS datasets, which will ultimately benefit many emerging frontiers in microbial ecology [16].

Modification of sample preparation procedures has been previously shown to influence fungal community characterization in culture-independent non-HTS studies. For example, increasing the amount of soil used in DNA extraction [17–19] or the use of sample pooling methods [20] can enhance capture of rare species. Additional steps in extraction protocol have also been shown to increase DNA yield [21–23] as well as the number of species recovered [24–26]. Conversely, species richness gains can also be achieved through regulating DNA template amount to improve PCR efficiency [12,27]. The influence of these different procedural modifications on our ability to characterize the richness of soil fungal communities, however, has not been sufficiently examined for the deep sequencing levels now possible using HTS.

To assess the role of modified procedures on species richness capture and address the need for standardization, we examined sample preparation protocols used in HTS surveys of soil fungi. Samples were collected at two sites in the Midwestern United States representing prairie and forest soils. Specifically, we examined the effect of extracting DNA from different volumes of soil (10 g, 1 g, 0.25 g) and using various extraction protocol modifications (freeze/thaw, sonication, 65°C water bath) on recovered fungal richness. Along with these treatments, we also examined fungal richness outcomes using varying PCR DNA template amounts (10 ng, 20 ng, 40 ng) and compared different sample pooling methods (physical vs. computational) to assess their ability to capture fungal community richness and composition at both sites.

Methods

Sample collection in Cedar Creek grant through Cedar Creek Ecosystem Science Reserve of the University of Minnesota. Sample collection in Cloquet forest grant through Cloquet Forest Center of the University of Minnesota.

Soil sampling

In fall 2013, soils were collected from two sites: the University of Minnesota Cedar Creek Ecosystem Science Reserve (CCR, 45°24'13" N, 93°11'20" W) and the University of Minnesota Cloquet Forestry Center (CFC, 46°40'45" N, 92°31'08" W). These sites represent distinct biome types, prairie and forest, respectively. CCR has mean monthly temperature range from -10°C to 22°C with an annual mean temperature at 6.7°C and an annual precipitation of 800 mm. The soil is highly sandy (93%) of the Zimmerman series [28]. CFC has mean monthly temperatures range from -14°C to 19°C with an annual mean temperature at 4°C and an annual precipitation of 761 mm. The soil is a loamy sand of the Omega and Cloquet series [29]. At CCR, soils were collected from the base of *Andropogon gerardii* plants growing in monoculture (n = 2 samples) and polyculture (n = 2 samples). For each sample, three 10 cm deep × 2.5 cm wide cores were taken from the base of an individual *A. gerardii* plant and combined in the field. At CFC, soils were collected under four replicate plots containing six tree species (*Pinus*

strobis, *Larix laricina*, *Picea glauca*, *Quercus rubra*, *Betula papyifera*, and *Acer saccharum*). For each sample, four 10 cm deep \times 2.5 cm wide cores were taken one meter apart and combined in the field. At each site, four individual plot sample replicates were gathered. All soil samples were transported in coolers to the lab and stored at -20°C until sieving at 4 mm size to remove fine roots and any part of small woody debris prior to DNA extraction.

Experimental design and DNA extraction

We designed our experiments to examine the effects of soil volume, DNA extraction modifications, PCR DNA template amount, and soil sample pooling on observed fungal richness and community composition. To assess soil volume, DNA was extracted from 10 g, 1 g, and 0.25 g of soil from each plot (12 samples sequenced per site, 24 samples sequenced in total). The PowerMax Soil DNA kit (MO BIO Laboratories, Carlsbad, CA, USA) was used for 10 g extractions. For the 1 g extractions, the PowerSoil kit (MO BIO Laboratories, Carlsbad, CA, USA) was used to extract DNA from four replicate 0.25 g portions of the same plot soil sample (0.25 g is the standard quantity recommended for this kit by the manufacturer). DNA concentration for each of the four replicate extractions were then quantified using a BioTek Synergy H1 plate reader (BioTek, Winooski, VT, USA) and mixed at equal amount to prepare a single '1 g' sample. An additional 0.25 g portion of the same plot soil sample was used in the PowerSoil kit to represent the kit 'standard' extraction. All extractions were carried out according to the manufacturer's instruction. Homogenizing was done, as specified in the kit protocol, by vortexing 10 min on the MO BIO vortex adapters (MO BIO Laboratories, Carlsbad, CA, USA, 13000-V1-24 for PowerSoil kit, and 13000-V1-50 for PowerMax Soil kit).

In the DNA extraction kit modification experiment, 0.25 g samples of soil were added to the PowerSoil 'bead-beater' extraction tubes containing the kit buffer. Samples were then subjected to three different modification procedures, freeze/thaw, sonication, or heating via water bath incubation (12 treatment samples sequenced per site, 24 treatment samples sequenced in total). The freeze/thaw treatment consisted of three rounds of freezing and thawing by submersing tubes in liquid N_2 for 5–10 seconds followed by an incubation step in a 65°C water bath for 10 min. The sonication treatment consisted of sonicating tubes for 10 min in a Branson 8200 sonicator (Thomas Scientific, Swedesboro, NJ, USA). Finally, the heating treatment consisted of incubating tubes in a 65°C water bath for 10 min. After each modification, DNA extraction proceeded according to the kit protocol instructions. The standard 0.25 g soil extraction from the soil volume experiment above served as an unmodified control for comparison.

To assess the effects of soil sample pooling and DNA template amount, four individual plot samples were combined in equal amounts (3 g from each sample) and thoroughly mixed to form a 'bulked' sample (hereafter referred to as the 'physical' pool) for each site. From these physical pool samples, DNA was extracted from 10 g, 1 g, and 0.25 g soil as describe above. DNA extracted from each of these three different soil volumes was then used as template in individual PCR reactions containing different template DNA amounts (10 ng, 20 ng, and 40 ng). Aliquots were prepared from a standardized $10\text{ ng }\mu\text{L}^{-1}$ stock concentration and added to the PCR reaction at 1 μL , 2 μL , and 4 μL volumes to achieve the desired quantity of DNA template (9 samples sequenced per site, 18 samples sequenced in total). Along with all of the aforementioned samples, for each extraction kit that was used (PowerSoil and PowerMax Soil), a soil-free extraction was conducted using the manufacturer's protocol to serve as negative controls. These negative controls allowed us to identify contaminants introduced during the extraction process and remove them from the final analyses. Finally, a 'mock' fungal community consisting of DNA extracted from 25 known fungal taxa was also included in the sequencing run to allow for internal optimization of the OTU clustering method [14].

Fungal ITS amplification and Illumina sequencing

PCR amplification of the ITS1 region of the nuclear ribosomal RNA gene was conducted using the general fungal primers ITS1-F and ITS2 adapted for Illumina sequencing [13]. Using other primer combinations may improve amplification of certain fungal groups [30]; however, testing differences among primer pairs was not the aim of this study. Each 20 μ L PCR reaction consisted of 10 μ L 2x Roche FastStart PCR Master (Roche, Indianapolis, IN, USA), 0.35 μ L of each forward and reverse primers (10 μ M), 1.2 μ L MgCl₂ (25 mM), and 2 μ L template DNA (20 ng, except in the DNA template amount experiment where 10 ng and 40 ng were also used, see above). Triplicate PCR reactions were performed for each sample using three separate annealing temperatures (50°C, 53°C, and 55°C)[14]. Thermocycling conditions included an initial denaturation step of 95°C for 10 min, 30 cycles of 95°C for 30s, annealing at 50°C, 53°C, or 55°C for 20 s, and extension at 72°C for 30 s, and a single final extension at 72°C for 8 minutes. Triplicate PCR reactions at each annealing temperature were pooled, purified using the Agencourt AMPure XP PCR purification kit (Beckman Coulter Inc., Brea, CA, USA). Purified PCR products were quantified using the PicoGreen dsDNA assay (Life Technologies, Grand Island, NY, USA) following the manufacturer's instructions, and 30 ng of DNA from each sample was combined into a single pool for Illumina MiSeq sequencing at the University of Minnesota Biomedical Genomics Center.

Data processing

Low quality ends of reads ($< q20$), primer sites, and adapter sites were trimmed from forward MiSeq Illumina reads (match error rate of $e = 0.2$) using Cutadapt [31]. Trimmomatic [32] was then used to trim from both ends (quality threshold set at 20) and filter out reads less than 125 bp long. Sequences with ambiguous bases and homopolymers greater than 9 bp were removed in Mothur [33]. All sequences files were then combined into a single fasta file with valid QIIME labels [34]. To generate a *de novo* database of operational taxonomic units (OTUs), sequences were dereplicated and singletons were removed in USEARCH [35]. Sequences were then clustered at 97% similarity using the UPARSE algorithm implemented in USEARCH, which included chimera detection and filtering. Subsequently, to reduce the number of spurious OTUs, a second round of clustering was conducted at 95% similarity using the UCLUST algorithm in QIIME, as Nguyen *et al.* (2014) found that 'chaining' OTU clustering algorithms generated more accurate OTU bins in a similarly processed dataset [14]. The representative OTU list after the 'chaining' clustering were filtered against a customized UNITE database [36] with UCLUST, but at a 75% similarity level to remove non-fungal sequences. Any OTU that failed in the third clustering round were removed from the *de novo* OTU list and blasted against another custom database that includes fungal voucher sequences and the UNITE database [14]. OTUs that hit the custom database and passed a quality filter (match length > 0.85 , similarity level $> 0.75\%$, and kingdom 'Fungi' identity in the taxonomic string) were added back to the *de novo* OTU list. Mapping the original sequence file to the *de novo* OTU sequence list generated an OTU table. Finally, the abundances of each OTU found in negative controls were subtracted from all samples prior to rarefaction following Nguyen *et al.* (2014)[14].

Statistical analysis

All samples were rarefied to an equal sampling depth of 18,778 sequences, and observed fungal OTU richness, Chao1, Simpson as well as Shannon indexes for each sample were calculated using the QIIME alpha_diversity.py command. To assess differences among treatments, a series of one-way ANOVAs were applied on OTU richness with Tukey's multiple comparisons in the R statistical software (www.r-project.org). Diversity estimates, such as Chao 1 was not

used in our analyses due to the potential biases associated with them for next generation sequencing studies [37]; however, we have provided them in the supplementary materials (S1 File.) for general comparison. A Bray-Curtis dissimilarity matrix among samples was calculated, and community distance was visualized in non-metric multidimensional scaling (NMDS) plots using the *vegan* [38] and *ggplot2* [39] packages in R. The *vegan* package was also used to assess potential community composition differences among treatments by testing with ADO-NIS. For each site, the physical pooled samples sequenced from the standard 0.25 g soil sample (using a 20 ng DNA template amount for PCR) were also compared to a corresponding pool of the individual plot samples (0.25 g soil and 20 ng DNA template) achieved through computation. For this 'computational' pool, the number of each OTU recovered from the four individual plot samples were summed, and the resultant composite pool was rarefied at the 18,778-sequence level. We also compared the average occurrence frequency of OTUs between physical and individual pools. The occurrence frequency of each OTU is defined as the occurrence (sequence number ≥ 1) in either 9 physical pools or 24 individual pools in the form of percentage in both site. The average occurrence frequency was calculated as the average of 20 OTUs around each individual OTU starting from the 10th (i.e., 1st to 20th OTU, 2nd to 21st OTU, . . ., 100th to 119th . . .). Paired T-test was applied between each pair of the 20-OTUs subset ($\alpha = 0.05$). All T-tests were adjusted with Bonferroni correction with hypotheses $n = 20$.

Results and Discussion

High-throughput sequencing of soil fungal communities

After quality filtering and sequence processing, we obtained 2,958,381 sequences, with a range of 18,778 to 78,509 sequences per sample. Rarefying at the lowest per sample recovery threshold (18,778 sequences), we recovered a total of 2,009 unique fungal OTUs from our 66 soil samples. These OTUs represented five known phyla (S2 Fig). We found a total of 1,273 and 1,436 OTUs in CCR (Cedar Creek Ecosystem Science Reserve) and CFC (Cloquet Forestry Center) soils (33 samples at each site), respectively. On average, individual samples contained ~330 OTUs for CCR and ~390 OTUs for CFC. Due to substantial community turnover among samples at each site, soils at CCR and CFC harbored a very high degree of total fungal richness, with rarefaction curves failing to achieve an asymptote (S3 Fig). These results agree with other high-throughput sequencing studies [11,40–42] that consistently reveal extremely high levels of fungal richness in soils of various ecosystems. In fact, only one study to date has successfully saturated the rarefaction curve for soil fungi, recovering a total of 1,002 OTUs from a single biome type [43], a richness level that is considerably smaller than what we detected at our sites.

Treatment effects on fungal alpha diversity

Extracting DNA from larger volumes of soil did not result in significantly higher OTU richness for either site (Fig 1A and 1B for richness and S1 Table for other indexes), although average richness did increase slightly with more soil used in the extraction. Extracting DNA from more soil may be desirable if detecting low abundance taxa is critical. For example, methods for early detection of soil-borne plant pathogens have been developed that require using large volumes of soil [17–19]. In support of this approach, we were able to detect potential plant pathogens, such as *Cochliobolus sativus*, in our 10 g and 1 g soil extraction samples that were not present when only 0.25 g of soil were used. However, our data suggest that extracting DNA from small soil volumes (e.g., the standard 0.25 g recommended in widely-used commercial extraction kits) is sufficient for most HTS applications in which a general characterization of soil fungal richness is required.

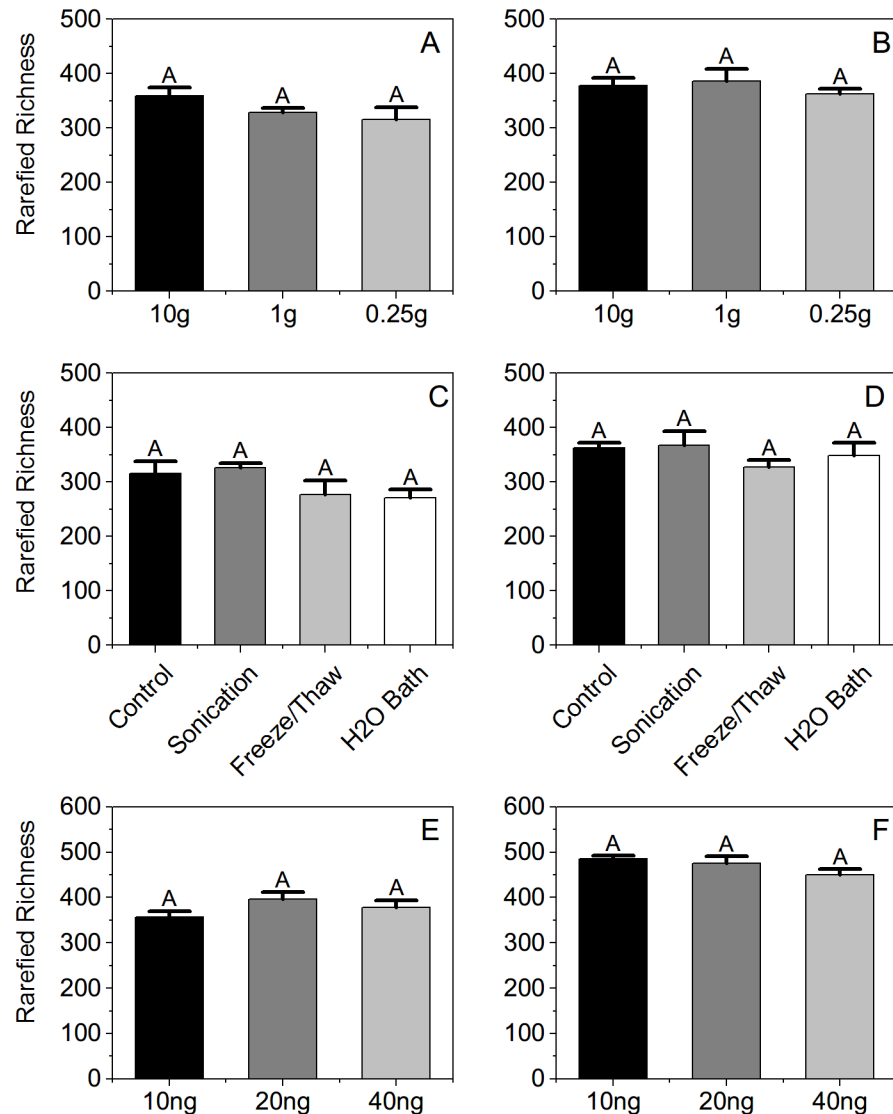


Fig 1. Rarefied OTU richness estimates of soil fungi by treatment. A-B. Amount of soil extracted for CCR and CFC, respectively (n = 5 for each treatment). C-D. Extraction modification for CCR and CFC, respectively (n = 4 for each treatment). E-F. Amount of DNA template in each PCR reactions for CCR and CFC, respectively (n = 3 for each treatment). The same letter above each bar indicates no significant difference ($P \leq 0.01$) in ANOVA using Tukey's multiple comparison tests ($\alpha = 0.05$).

doi:10.1371/journal.pone.0127234.g001

Including additional steps in the extraction protocol, such as sonication treatments, freeze/thaw regimes, or period of incubation in a hot water bath, all of which are thought to increase DNA yields and recovered richness, likewise did not significantly impact observed fungal richness (Fig 1C and 1D for richness & S1 Table for other diversity indexes). Although previous studies have found significant effects of DNA extraction method on recovered fungal and bacterial richness [21,24–26,44–46], those studies used older molecular techniques (e.g., RFLP and DGGE) that are likely more sensitive to biases due to the limited depth at which they survey the communities [47–49]. While our results suggest unmodified protocols from standard DNA extraction kits is sufficient for general characterizing fungal richness of soil samples in HTS surveys, we reiterate that study aims must be considered. For example, more stringent DNA

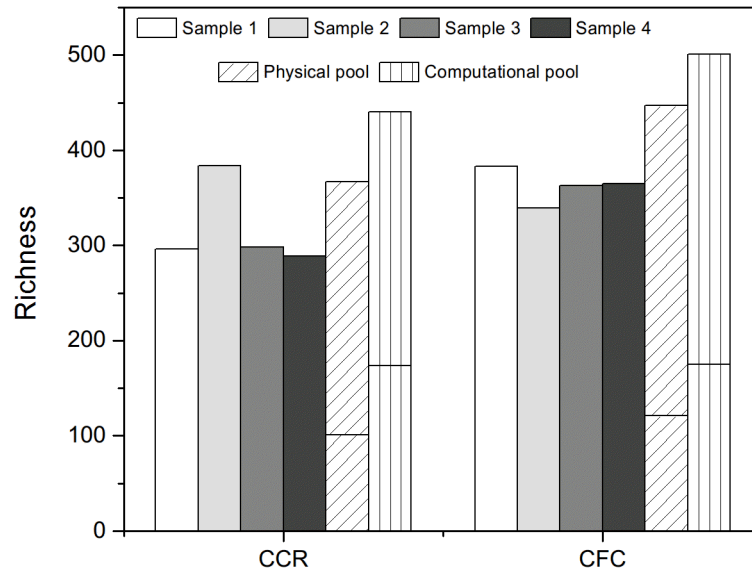


Fig 2. Soil fungal OTU richness of individual samples, physical pool, and computational pool. The shorter bars inside of physical and computational pools indicated the number of unique OTUs that were not shared between the two pools in each site. All samples used 0.25 g soil with standard DNA extraction and PCR procedures.

doi:10.1371/journal.pone.0127234.g002

extraction treatments may be necessary to detect fungi exhibiting resistant structures [50–52]; however, facilitating extraction of DNA from such structures (e.g., resistant spores) is often undesirable as the associated fungal taxa may represent transient, rather than living, components of the community.

As with the other preparation treatments, the amount of DNA template used in PCR reactions (10 ng, 20 ng, or 40 ng) did not significantly influence recovered fungal richness in our samples either (Fig 1E and 1F for richness & S1 Table for other indexes). While significant effects on fungal richness were not detected overall, opposing trends were noted between sites, with slightly lower richness observed for the highest template amount at CCR and slightly higher richness recorded for higher template amounts at CFC. It is possible that these results are related to the fact that increasing DNA template amount can adversely affect PCR by also increasing the concentrations of inhibitors [27,53,54] and Lindahl *et al.* (2013) noted that diluting DNA template for PCR steps in HTS library preparation is often required [12]. While we did not directly measure potential inhibitors in our samples, our results do highlight the often inconsistent results of increasing DNA templates amounts among different sample types as well as the need for template optimization in preparing HTS runs.

Sample pooling and treatment effects on fungal community composition

At both sites, physical pools generally had higher richness (Fig 2) and more unique OTUs (S4 Fig) when compared with each individual plot sample. Abundant OTUs (> 1% of the total sequence abundance, consisted of ~70% sequences in total for both sites) appearing in each individual plot sample were also more fully captured in the physical pools at both sites (Fig 3 and S2 File.). Between the two pooling methods, computational pools of each site had higher richness measures than physical pools (Fig 2), due mostly to the contribution of unique OTUs from all four individual plot samples (S4 Fig). A number of OTUs were also recovered in the computation pool that were not present in the physical pool, while a smaller portion were

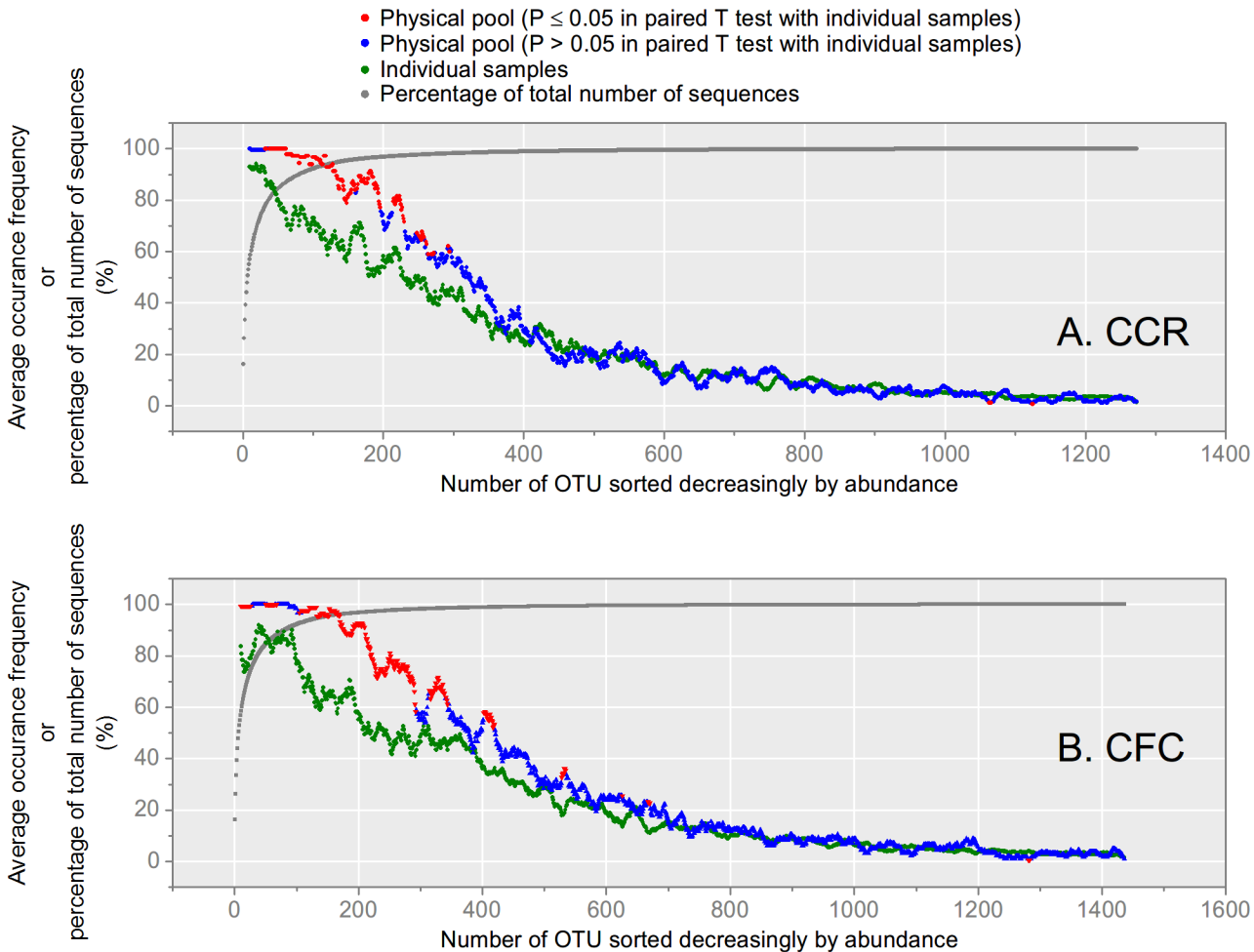


Fig 3. Average occurrence frequency of OTUs in physical pools versus individual samples, and percent of total number of sequence, moving from abundant (left) to rare (right) taxa. Physical pools are in red and blue points, individual samples are in green points, total number of sequence is in gray points. The occurrence frequency of each OTU is defined as the occurrence (sequence number ≥ 1) in either 9 physical pools or 24 individual pools in the form of percentage in both site. The average occurrence frequency shown in the plot was calculated as the average of 20 OTUs around each individual OTU. In the physical pool, red points indicate significant difference ($P \leq 0.05$) in paired T-test between physical and individual pools ($\alpha = 0.05$), while blue points were not significantly different. All T-tests were adjusted with Bonferroni correction with hypotheses $n = 20$. The percent of total number of sequence was defined as the sum of abundance percentage from 1st to current OTU in descending order of abundance.

doi:10.1371/journal.pone.0127234.g003

unique to the physical pool alone (Fig 2). Since soil from each of the study sites originated from a very similar source (see methods), it is unlikely that the differences on unique OTUs could be due to data process procedures or PCR amplifications. Our findings are similar to Branco *et al.* (2013), who found that computational pooling yielded higher fungal OTU richness than physical sample pool and Dickie *et al.* (2010), who found different rare taxa in pooled and individual samples [20,55]. Importantly, the high level of fungal community turnover demonstrated in this study and others [11,40–42] indicates the difficulty in achieving comprehensive characterization of soil fungal communities, even using the deep level of survey possible with HTS.

In addition to differences in richness, physically pooled and individual plot samples for each site differed significantly in their fungal community structure (ADONIS based on Bray Curtis and presence/absence distance matrixes respectively, $F_{4, 28} = 17.71$ and 6.10 for CCR, $F_{4,28} = 81.06$ and 6.71 for CFC, $P = 0.001$ for all tests). Our results are again similar to those of Branco *et al.* (2013), who showed significant differences in fungal community structure between

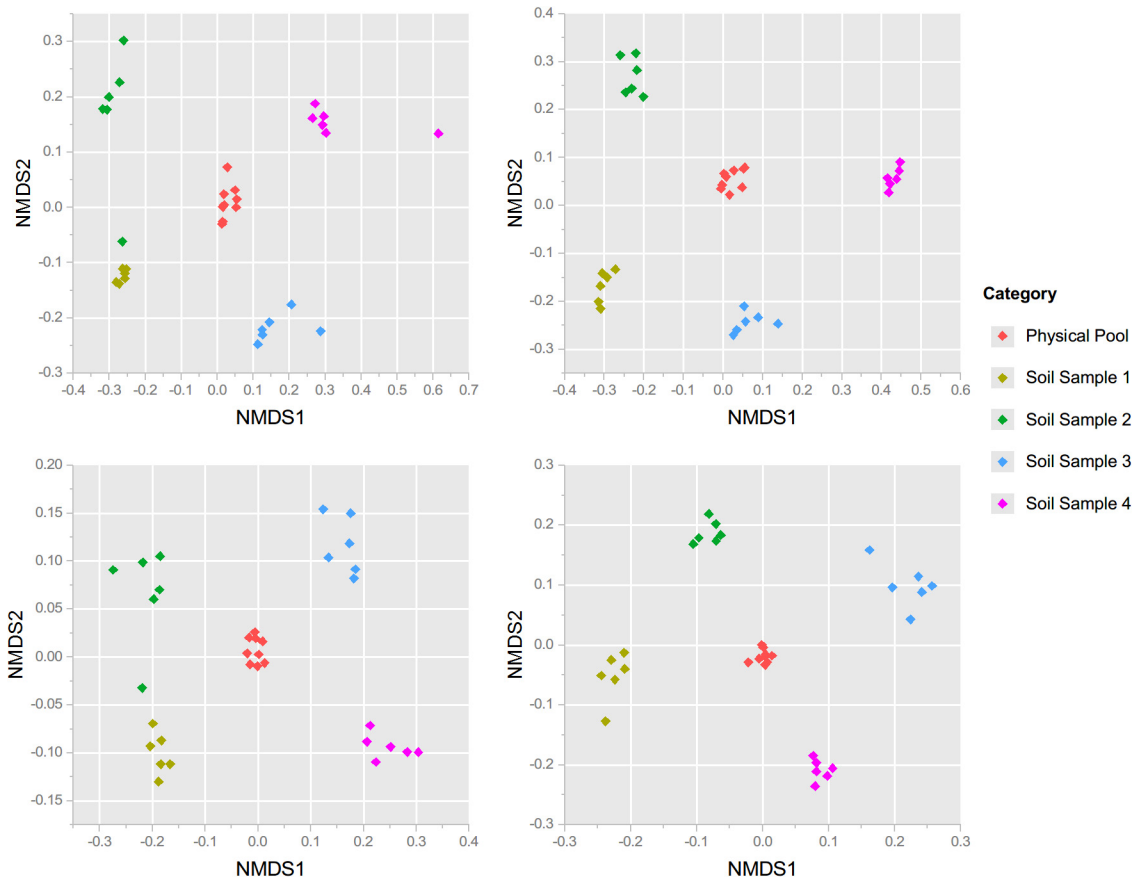


Fig 4. Nonmetric Multidimensional Scaling Plot of physical pools and individual plot samples from each site. Plots based on either Bray-Curtis distance matrices generated from rarefied taxon abundances of A. CCR and B. CFC or from presence/absence matrix of C. CCR and D. CFC.

doi:10.1371/journal.pone.0127234.g004

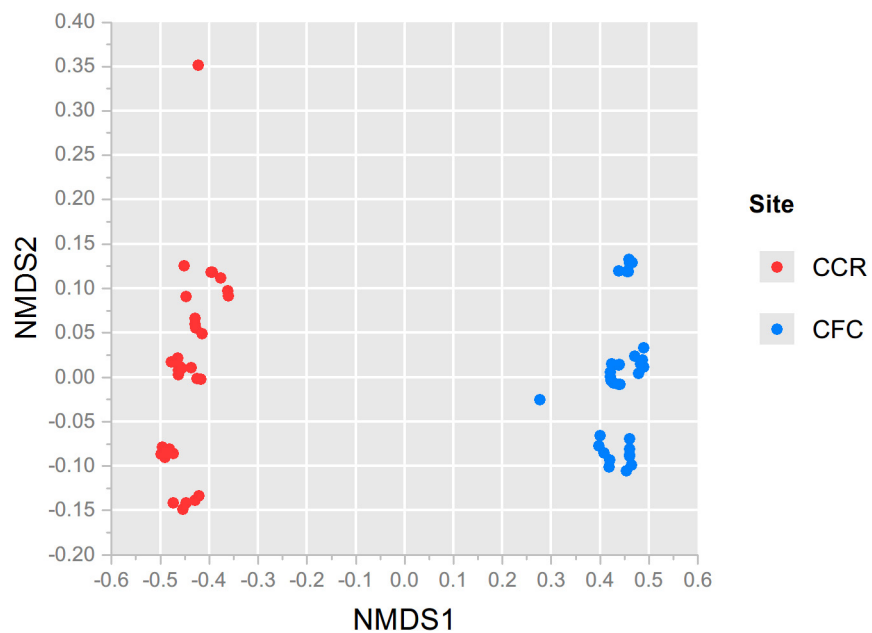


Fig 5. Nonmetric Multidimensional Scaling Plot of samples from CCR and CFC. Plot is based on a Bray-Curtis distance matrix generated from rarefied taxon abundances.

doi:10.1371/journal.pone.0127234.g005

individual versus physically pooled samples [20]; however, the differences among individual samples taken at their site were not significant. In our study, significant compositional differences among individual samples are apparent in the distinct clustering patterns of the NMDS plots (Fig 4). The importance of individual taxa, rather than their respective abundances, in driving these clustering patterns (Fig 4) is demonstrated by the consistency of the results between the two different matrix types (Bray Curtis vs. presence/absence). In addition, the tight clusters formed by the physical pool samples in the central 'ordination' spaces of Fig 4, suggests the averaging effect that physical pooling had on the fungal communities recovered. Overall, site differences in fungal community structure were much stronger than any other effect observed within the sites (Fig 5).

Conclusions

When considering the effort applied in the sample preparation phase of this survey, we found that the effort generally outweighed any associated rewards. Specifically, we found that sonication, freeze/thaw cycling, and heating during DNA extraction did not result in notable gains of richness or structural characterization of soil fungal communities. Likewise, community-level results were very similar regardless of the amount of soil used in the DNA extraction protocol or the amount of DNA template used in the PCR reactions. We did, however, find that pooling individual soil samples, either physically or computationally, resulted in higher richness. Although computational pooling of samples provides a more comprehensive assessment of local OTU richness than physical pooling, there are additional associated effort of preparation and sequencing. We therefore suggest that for complete biotic inventories, collecting a high number of individual samples at a given site would be preferable, while a single pooled sample from many sites would be preferable for studies aimed at larger spatial scales. Overall, we conclude that following standard protocols of widely used DNA extraction kits is sufficient to successfully assess soil fungal community richness and composition in HTS surveys.

Supporting Information

S1 Fig. Number of publications related to fungal HTS studies from 2000 to 2013 found by searching key words—fungi OR fungal, "high throughput sequencing"—in Google Scholar (<http://scholar.google.com>). The points plotted fit an exponential curve ($\alpha = 0.05$; $P \leq 0.01$). (TIF)

S2 Fig. Relative abundance of fungal phyla in Cedar Creek Reserve (CCR) and Cloquet Forest Center (CFC) soils after rarefaction (n = 619,674 sequences/site). (TIF)

S3 Fig. Taxon (OTU) accumulation curves in this study. Taxon (OTU) accumulation curves depicting the rarefied richness* of fungal taxa as a function of increased sample sequencing effort at CCR (brown) and CFC (green) in the Midwestern United States (*sample depth = 18,778 sequences per sample). (TIF)

S4 Fig. Number of unique and shared OTUs between individual samples or computational pool compared with the physical pool. Samples using 0.25 g soil with standard extraction and PCR procedure were used. Computational pool was generated with the four individual samples and rarefied at the same level of richness (18,778 sequences). (TIF)

S1 Table. One-way ANOVA of alpha diversity indexes and ADONIS in the three treatments on sequencing sample preparations.

(DOCX)

S1 File. Diversity indexes in different treatments.

(XLSX)

S2 File. OTU tables for CCR and CFC.

(XLSX)

Acknowledgments

We thank the University of Minnesota Supercomputing Institute (MSI) for providing computing support.

Author Contributions

Conceived and designed the experiments: DS PK LK HCK NN STB. Performed the experiments: ZS DS PK NN STB. Analyzed the data: ZS DS PK NN STB. Contributed reagents/materials/analysis tools: PK NN STB. Wrote the paper: ZS DS PK LK HCK NN STB.

References

1. Fierer N, Bradford MA, Jackson RB. Toward an ecological classification of soil bacteria. *Ecology*. 2007; 88: 1354–64. PMID: [17601128](#)
2. Falkowski PG, Fenchel T, Delong EF. The microbial engines that drive Earth's biogeochemical cycles. *Science*. 2008; 320: 1034–1039. doi: [10.1126/science.1153213](#) PMID: [18497287](#)
3. Wagg C, Bender SF, Widmer F, van der Heijden MGA. Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proc Natl Acad Sci*. 2014; 111: 5266–70. doi: [10.1073/pnas.1320054111](#) PMID: [24639507](#)
4. Curtis TP, Head IM, Lunn M, Woodcock S, Schloss PD, Sloan WT. What is the extent of prokaryotic diversity? *Phil Trans R Soc Lond A*. 2006; 361: 2023–37. doi: [10.1098/rsta.2006.1921](#) PMID: [17028084](#)
5. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*. 2012; 6: 1621–4. doi: [10.1038/ismej.2012.8](#) PMID: [22402401](#)
6. Van der Heijden MGA, Klironomos J. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature*. 1998; 74: 69–72.
7. Van der Heijden MGA, Bardgett RD, van Straalen NM. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett*. 2008; 11: 296–310. doi: [10.1111/j.1461-0248.2007.01139.x](#) PMID: [18047587](#)
8. De Vries FT, Manning P, Tallwin JRB, Mortimer SR, Pilgrim ES, Harrison K a, et al. Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities. *Ecol Lett*. 2012; 15: 1230–9. doi: [10.1111/j.1461-0248.2012.01844.x](#) PMID: [22882451](#)
9. McGuire KL, Payne SG, Palmer MI, Gillikin CM, Keefe D, Kim SJ, et al. Digging the New York City Skyline: soil fungal communities in green roofs and city parks. *PLoS One*. 2013; 8: e58020. doi: [10.1371/journal.pone.0058020](#) PMID: [23469260](#)
10. Kemler M, Garnas J, Wingfield MJ, Gryzenhout M, Pillay KA, Slippers B. Ion Torrent PGM as tool for fungal community analysis: a case study of endophytes in *Eucalyptus grandis* reveals high taxonomic diversity. *PLoS One*. 2013; 8: e81718. doi: [10.1371/journal.pone.0081718](#) PMID: [24358124](#)
11. Talbot JM, Bruns TD, Taylor JW, Smith DP, Branco S, Glassman SI, et al. Endemism and functional convergence across the North American soil mycobiome. *Proc Natl Acad Sci*. 2014; 111: 6341–6. doi: [10.1073/pnas.1402584111](#) PMID: [24733885](#)
12. Lindahl BD, Nilsson RH, Tedersoo L, Abarenkov K, Carlsen T, Pennanen T, et al. Fungal community analysis by high throughput sequencing of amplified markers—a user's guide. *New Phytol*. 2013; 199: 288–299. doi: [10.1111/nph.12243](#) PMID: [23534863](#)

13. Smith DP, Peay KG. Sequence Depth, Not PCR Replication, Improves Ecological Inference from Next Generation DNA Sequencing. *PLoS One*. 2014; 9: e90234. doi: [10.1371/journal.pone.0090234](https://doi.org/10.1371/journal.pone.0090234) PMID: [24587293](https://pubmed.ncbi.nlm.nih.gov/24587293/)
14. Nguyen N, Smith D, Peay K, Kennedy P. Parsing ecological signal from noise in next generation amplicon sequencing. *New Phytol*. 2014; doi: [10.1111/nph.12923](https://doi.org/10.1111/nph.12923)
15. Vishnivetskaya TA, Layton AC, Lau MCY, Chauhan A, Cheng KR, Meyers AJ, et al. Commercial DNA extraction kits impact observed microbial community composition in permafrost samples. *FEMS Microbiol Ecol*. 2014; 87: 217–30. doi: [10.1111/1574-6941.12219](https://doi.org/10.1111/1574-6941.12219) PMID: [24102625](https://pubmed.ncbi.nlm.nih.gov/24102625/)
16. Knight R, Jansson J, Field D, Fierer N, Desai N, Fuhrman J a, et al. Unlocking the potential of metagenomics through replicated experimental design. *Nat Biotechnol*. 2012; 30: 513–20. doi: [10.1038/nbt.2235](https://doi.org/10.1038/nbt.2235) PMID: [22678395](https://pubmed.ncbi.nlm.nih.gov/22678395/)
17. Woodhall JW, Webb KM, Giltrap PM, Adams IP, Peters JC, Budge GE, et al. A new large scale soil DNA extraction procedure and real-time PCR assay for the detection of *Sclerotium cepivorum* in soil. *Eur J Plant Pathol*. 2012; 134: 467–473. doi: [10.1007/s10658-012-0025-2](https://doi.org/10.1007/s10658-012-0025-2)
18. Than DJ, Hughes KJD, Boonhan N, Tomlinson JA, Woodhall JW, Bellgard SE. A TaqMan real-time PCR assay for the detection of *Phytophthora* “taxon Agathis” in soil, pathogen of Kauri in New Zealand. Andrea V, editor. *For Pathol*. 2013; 43: 324–330. doi: [10.1111/efp.12034](https://doi.org/10.1111/efp.12034)
19. Boine B, Renner AC, Zellner M, Nechwatal J. Quantitative methods for assessment of the impact of different crops on the inoculum density of *Rhizoctonia solani* AG2-2IIIB in soil. *Eur J Plant Pathol*. 2014; doi: [10.1007/s10658-014-0506-6](https://doi.org/10.1007/s10658-014-0506-6)
20. Branco S, Bruns TD, Singleton I. Fungi at a Small Scale: Spatial Zonation of Fungal Assemblages around Single Trees. *PLoS One*. 2013; 8: 1–10. doi: [10.1371/journal.pone.0078295](https://doi.org/10.1371/journal.pone.0078295)
21. Fredricks D, Smith C, Meier A. Comparison of six DNA extraction methods for recovery of fungal DNA as assessed by quantitative PCR. *J Clin Microbiol*. 2005; 43: 5122–5128. doi: [10.1128/JCM.43.10.5122](https://doi.org/10.1128/JCM.43.10.5122) PMID: [16207973](https://pubmed.ncbi.nlm.nih.gov/16207973/)
22. Karakousis A, Tan L, Ellis D, Alexiou H, Wormald PJ. An assessment of the efficiency of fungal DNA extraction methods for maximizing the detection of medically important fungi using PCR. *J Microbiol Methods*. 2006; 65: 38–48. doi: [10.1016/j.mimet.2005.06.008](https://doi.org/10.1016/j.mimet.2005.06.008) PMID: [16099520](https://pubmed.ncbi.nlm.nih.gov/16099520/)
23. Francesconi A, Kasai M, Harrington SM, Beveridge MG, Petraitiene R, Petraitis V, et al. Automated and manual methods of DNA extraction for *Aspergillus fumigatus* and *Rhizopus oryzae* analyzed by quantitative real-time PCR. *J Clin Microbiol*. 2008; 46: 1978–84. doi: [10.1128/JCM.02246-07](https://doi.org/10.1128/JCM.02246-07) PMID: [18353931](https://pubmed.ncbi.nlm.nih.gov/18353931/)
24. Martin-Laurent F, Philippot L. DNA extraction from soils: old bias for new microbial diversity analysis methods. *Appl Environ Microbiol*. 2001; 67: 2354–2359. doi: [10.1128/AEM.67.5.2354](https://doi.org/10.1128/AEM.67.5.2354) PMID: [11319122](https://pubmed.ncbi.nlm.nih.gov/11319122/)
25. Roh C, Villatte F, Kim B, Schmid RD. Comparative study of methods for extraction and purification of environmental DNA from soil and sludge samples. *Appl Biochem Biotechnol*. 2006; 134: 97–112. PMID: [16943632](https://pubmed.ncbi.nlm.nih.gov/16943632/)
26. Sagova-Mareckova M, Cermak L, Novotna J, Plhachova K, Forstova J, Kopecky J. Innovative methods for soil DNA purification tested in soils with widely differing characteristics. *Appl Environ Microbiol*. 2008; 74: 2902–7. doi: [10.1128/AEM.02161-07](https://doi.org/10.1128/AEM.02161-07) PMID: [18344341](https://pubmed.ncbi.nlm.nih.gov/18344341/)
27. Wilson I. Inhibition and facilitation of nucleic acid amplification. *Appl Environ Microbiol*. 1997; 63: 3741–51. PMID: [9327537](https://pubmed.ncbi.nlm.nih.gov/9327537/)
28. Carol Adair E, Reich PB, Trost JJ, Hobbie SE. Elevated CO₂ stimulates grassland soil respiration by increasing carbon inputs rather than by enhancing soil moisture. *Glob Chang Biol*. 2011; 17: 3546–3563. doi: [10.1111/j.1365-2486.2011.02484.x](https://doi.org/10.1111/j.1365-2486.2011.02484.x)
29. Dijkstra F a., West JB, Hobbie SE, Reich PB. Antagonistic effects of species on C respiration and net N mineralization in soils from mixed coniferous plantations. *For Ecol Manage*. 2009; 257: 1112–1118. doi: [10.1016/j.foreco.2008.11.014](https://doi.org/10.1016/j.foreco.2008.11.014)
30. Bellemain E, Carlsen T, Brochmann C, Coissac E, Taberlet P, Kauserud H. ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiol*. 2010; 10: 189. doi: [10.1186/1471-2180-10-189](https://doi.org/10.1186/1471-2180-10-189) PMID: [20618939](https://pubmed.ncbi.nlm.nih.gov/20618939/)
31. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*. 2011; 17: 10–12. <http://dx.doi.org/10.14806/ej.17.1.200>
32. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30: 2114–2120. doi: [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170) PMID: [24695404](https://pubmed.ncbi.nlm.nih.gov/24695404/)
33. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*. 2009; 75: 7537–41. doi: [10.1128/AEM.01541-09](https://doi.org/10.1128/AEM.01541-09) PMID: [19801464](https://pubmed.ncbi.nlm.nih.gov/19801464/)

34. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010; 7: 335–6. doi: [10.1038/nmeth.f.303](https://doi.org/10.1038/nmeth.f.303) PMID: [20383131](https://pubmed.ncbi.nlm.nih.gov/20383131/)
35. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010; 26: 2460–1. doi: [10.1093/bioinformatics/btq461](https://doi.org/10.1093/bioinformatics/btq461) PMID: [20709691](https://pubmed.ncbi.nlm.nih.gov/20709691/)
36. Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M, et al. Towards a unified paradigm for sequence-based identification of fungi. *Mol Ecol*. 2013; 22: 5271–5277. doi: [10.1111/mec.12481](https://doi.org/10.1111/mec.12481) PMID: [24112409](https://pubmed.ncbi.nlm.nih.gov/24112409/)
37. Dickie IA. Insidious effects of sequencing errors on perceived diversity in molecular surveys. *New Phytol*. 2010; 188: 916–918. doi: [10.1111/j.1469-8137.2010.03473.x](https://doi.org/10.1111/j.1469-8137.2010.03473.x) PMID: [20854395](https://pubmed.ncbi.nlm.nih.gov/20854395/)
38. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. *vegan: Community Ecology Package*. R package version 2.0–7. 2013;
39. Wickham H. *ggplot2: elegant graphics for data analysis*. Springer New York; 2009.
40. Jumpponen A, Jones KL, David Mattox J, Yaeger C. Massively parallel 454-sequencing of fungal communities in *Quercus* spp. ectomycorrhizas indicates seasonal dynamics in urban and rural sites. *Mol Ecol*. 2010; 19 Suppl 1: 41–53. doi: [10.1111/j.1365-294X.2009.04483.x](https://doi.org/10.1111/j.1365-294X.2009.04483.x) PMID: [20331769](https://pubmed.ncbi.nlm.nih.gov/20331769/)
41. Tedersoo L, Nilsson RH, Abarenkov K, Jairus T, Sadam A, Saar I. Methods 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytol*. 2010; 188: 291–301. doi: [10.1111/j.1469-8137.2010.03373.x](https://doi.org/10.1111/j.1469-8137.2010.03373.x) PMID: [20636324](https://pubmed.ncbi.nlm.nih.gov/20636324/)
42. Xu L, Ravnskov S, Larsen J, Nilsson RH, Nicolaisen M. Soil fungal community structure along a soil health gradient in pea fields examined using deep amplicon sequencing. *Soil Biol Biochem*. 2012; 46: 26–32. doi: [10.1016/j.soilbio.2011.11.010](https://doi.org/10.1016/j.soilbio.2011.11.010)
43. Taylor D, Hollingsworth T, McFarland J. A first comprehensive census of fungi in soil reveals both hyperdiversity and fine-scale niche partitioning. *Ecology*. 2013; 84: 3–20.
44. He J, Xu Z, Hughes J. Pre-lysis washing improves DNA extraction from a forest soil. *Soil Biol Biochem*. 2005; 37: 2337–2341. doi: [10.1016/j.soilbio.2005.04.016](https://doi.org/10.1016/j.soilbio.2005.04.016)
45. Carrigg C, Rice O, Kavanagh S, Collins G, O'Flaherty V. DNA extraction method affects microbial community profiles from soils and sediment. *Appl Microbiol Biotechnol*. 2007; 77: 955–64. doi: [10.1007/s00253-007-1219-y](https://doi.org/10.1007/s00253-007-1219-y) PMID: [17960375](https://pubmed.ncbi.nlm.nih.gov/17960375/)
46. Lakay FM, Botha A, Prior BA. Comparative analysis of environmental DNA extraction and purification methods from different humic acid-rich soils. *J Appl Microbiol*. 2007; 102: 265–73. doi: [10.1111/j.1365-2672.2006.03052.x](https://doi.org/10.1111/j.1365-2672.2006.03052.x) PMID: [17184343](https://pubmed.ncbi.nlm.nih.gov/17184343/)
47. Muyzer G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol*. 1993; 59: 695–700. PMID: [7683183](https://pubmed.ncbi.nlm.nih.gov/7683183/)
48. Sekiguchi H, Tomioka N, Nakahara T, Uchiyama H. A single band does not always represent single bacterial strains in denaturing gradient gel electrophoresis analysis. *Biotechnol Lett*. 2001; 23: 1205–1208.
49. Pilloni G, Granitsiotis MS, Engel M, Lueders T. Testing the limits of 454 pyrotag sequencing: reproducibility, quantitative assessment and comparison to T-RFLP fingerprinting of aquifer microbes. *PLoS One*. 2012; 7: e40467. doi: [10.1371/journal.pone.0040467](https://doi.org/10.1371/journal.pone.0040467) PMID: [22808168](https://pubmed.ncbi.nlm.nih.gov/22808168/)
50. Sánchez B, Rodríguez M. Development of an efficient fungal DNA extraction method to be used in random amplified polymorphic DNA–PCR analysis to differentiate cyclopiazonic acid mold. *J Food Prot*. 2008; 71: 2497–2503. PMID: [19244904](https://pubmed.ncbi.nlm.nih.gov/19244904/)
51. Jang CS, Lim JH, Seo MW, Song JY, Kim HG. Direct Detection of *Cylindrocarpus destructans*, Root Rot Pathogen of Ginseng by Nested PCR from Soil Samples. *Mycobiology*. 2010; 38: 33–8. doi: [10.4489/MYCO.2010.38.1.033](https://doi.org/10.4489/MYCO.2010.38.1.033) PMID: [23956622](https://pubmed.ncbi.nlm.nih.gov/23956622/)
52. Goldschmidt P, Degorge S, Merabet L, Chaumeil C. Enzymatic treatment of specimens before DNA extraction directly influences molecular detection of infectious agents. *PLoS One*. 2014; 9: e94886. doi: [10.1371/journal.pone.0094886](https://doi.org/10.1371/journal.pone.0094886) PMID: [24936792](https://pubmed.ncbi.nlm.nih.gov/24936792/)
53. Wintzingerode FV, Göbel UB, Stackebrandt E. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol Rev*. 2006; 21: 213–229. doi: [10.1111/j.1574-6976.1997.tb00351.x](https://doi.org/10.1111/j.1574-6976.1997.tb00351.x)
54. Wu J, Jiang X, Jiang Y, Lu S, Zou F, Zhou H. Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. *BMC Microbiol*. BioMed Central Ltd; 2010; 10: 255. doi: [10.1186/1471-2180-10-255](https://doi.org/10.1186/1471-2180-10-255)
55. Dickie IA, Bolstridge N, Cooper JA, Peltzer DA. Co-invasion by Pinus and its mycorrhizal fungi. *New Phytol*. 2010; 187: 475–84. doi: [10.1111/j.1469-8137.2010.03277.x](https://doi.org/10.1111/j.1469-8137.2010.03277.x) PMID: [20456067](https://pubmed.ncbi.nlm.nih.gov/20456067/)