Chapter 4

Protocols for Investigating the Leaf Mycobiome Using High-Throughput DNA Sequencing

Shawn P. Brown, Devin R. Leopold, and Posy E. Busby

Abstract

High-throughput sequencing of taxon-specific loci, or DNA metabarcoding, has become an invaluable tool for investigating the composition of plant-associated fungal communities and for elucidating plant–fungal interactions. While sequencing fungal communities has become routine, there remain numerous potential sources of systematic error that can introduce biases and compromise metabarcoding data. This chapter presents a protocol for DNA metabarcoding of the leaf mycobiome based on current best practices to minimize errors through careful laboratory practices and validation.

Key words Fungal leaf endophytes, Next-generation sequencing, Metabarcoding, Mycobiome, Amplicons, Phyllosphere, Phytobiome, Sample bias, Primer bias, Mock community

1 Introduction

Fungal ecology has taken a monumental step forward with the advent of high-throughput DNA sequencing, which allows the simultaneous sequencing of millions of PCR-amplified gene fragments (amplicons). Using this technology to sequence amplicons from many samples in parallel while targeting taxonomically informative loci, a process also known as DNA metabarcoding, it is now possible to directly query unculturable and novel fungal communities to gain new insights into plant–fungal interactions. Dramatic reductions in cost and widespread availability of this technology have facilitated the democratization of DNA sequencing-based ecology [1] and many researchers now view high-throughput sequencing as an essential tool in their methodological toolbox. However, despite the seemingly straightforward process of obtaining sequences from environmental samples [2–4], there remain many opportunities to introduce errors and bias that can affect the balance of data across samples and among fungal taxa. Every step of the process, beginning with sample collection, has the potential

Wenbo Ma and Thomas Wolpert (eds.), *Plant Pathogenic Fungi and Oomycetes: Methods and Protocols*, Methods in Molecular Biology, vol. 1848, https://doi.org/10.1007/978-1-4939-8724-5_4,

to introduce biases that can then compound at each step. Although some biases can be identified and controlled for, for relatively unknown systems and taxa it can be difficult to identify problems simply because there is often a dearth of information about the communities under investigation.

In this chapter, we provide a protocol for sample preparation and amplicon library generation for investigating the leaf mycobiome, or the communities of fungal symbionts that live on and in plant leaves. The leaf mycobiome has been described as "hyperdiverse" [5] and often includes many novel taxa [6] whose ecologies and functions are unknown, making them challenging to study. Our protocol begins with sample collection and follows a flexible two-step PCR process for amplicon library generation [7], in which an initial PCR reaction amplifies template DNA with gene primers and a second PCR reaction adds sequencing adapters and sample-specific molecular identifiers. Notes are included to provide additional details, alternative approaches, and to emphasize techniques for minimizing error and user-generated biases.

2 Materials

2.1 Sample Collection and Processing

1. Individually labeled zip-top bags for leaf collection.
2. Cooler containing ice.
3. 70% EtOH (ethanol).
4. 1 L beakers.
5. Triton X-100.
6. Sterile water.
7. Hole punch or cork borer.

2.2 Genomic DNA Extraction

1. Plant or soil DNA extraction kits and required equipment and supplies (*see* kit protocols for these requirements).
2. Bead mill homogenizer (*see* **Note 1**).
3. DNA quantification equipment (e.g., Qubit or NanoDrop).

2.3 Amplification of Target Genes: PCR, and Sequencing

1. Stage 1 and 2 PCR primers specific to the targeted locus (*see* **Note 2**) and sequencing platform (i.e., Illumina MiSeq, *see* **Note 3**).
2. PCR reagents (*see* **Note 4**).
3. PCR product cleanup kits (96-well plate based cleanup and normalization kits such as SequalPrep; *see* **Note 5**).
4. Sequencing reagent kits—most sequencing facilities will provide this encompassed in the user cost.

2.4 Additional Supplies and Equipment

Sterile 96-well PCR plates, reagent reservoirs, a variety of sterile DNase/RNase-free low-retention filter micropipette tips, sterile microcentrifuge tubes, thermal cyclers with 96-well capability, gel electrophoresis equipment and supplies, reagent reservoirs, pipettes (multichannel pipettes preferred), sterile disposable gloves, and any additional equipment/supplies required for kit-based extraction and PCR cleanup kits (*see* kit protocols).

3 Methods

3.1 Sample Collection

1. Sample a minimum of three randomly (or haphazardly) selected leaves at a fixed height, at equal distances (e.g., 120°) around the plant to control for within-plant variation [8–10]. In some cases, depending on the experimental question, leaves with particular phenotypic traits (e.g., herbivore damage, necrotic tissue) may be avoided or targeted (*see* **Note 6**). Sampling leaves of the same age to prevent the introduction of a confounding phenological effect [11] is recommended. Monitor budburst and tag target leaves for later collection, or standardize sampling by the Leaf Plastochron Index (LPI) [12, 13], a measure of leaf age based on morphological development.
2. Place leaves into clean, individually labeled zip-top plastic bags and store immediately on ice or at 4 °C. Process leaves within 24 h to prevent fungi with saprobic capabilities from proliferating (*see* **Note 7**).

3.2 Sample Processing

1. In the laboratory, manually agitate leaves for 1 min in sterile ddH_2O containing Triton X-100 (1 mL Triton X-100 in 1 L ddH_2O), a surfactant that will remove spores and hyphal fragments on the leaf surface [14]. Next, rinse leaves by manually agitating for 30 s in sterile ddH_2O, in each of three consecutive washes (*see* **Note 8**).
2. Sub-sample a standardized mass of leaf material across sample units. The mass needed for DNA extraction varies with the extraction kit, but generally ranges from 0.25 g to 0.5 g (*see* extraction kit protocols). Excise an appropriate number of leaf discs (equal to desired mass) from the sampled leaves with a handheld hole-punch or a cork borer on a rubber mat (clean tool between samples by rinsing in 70% EtOH). If leaf tissue will not be used for DNA extraction immediately, store frozen at −20 °C for ~6 months or at −80 °C for ~3 years.

3.3 DNA Extraction

1. Genomic DNA extraction should follow kit protocols and include manual disruption on a ball mill homogenizer (*see* **Note 1**).

2. Examine DNA quantity and quality for all samples using fluorometic (e.g., Qubit) or spectrophotometric (e.g., NanoDrop) methods. For samples that were poorly extracted (consisting of low DNA concentrations, coextraction of PCR inhibitory compounds, or protein coextraction), reextract DNA or clean using a DNA clean up kit or DNA precipitation protocols (e.g., precipitation using isopropanol for low concentrations).
3. Store DNA in labeled 96-well 0.2 mL PCR plates or racks of 8-strip PCR tubes. For every 96 samples, a minimum of one well/tube should be reserved for sterile molecular grade water to be used as a negative control [15] with no fewer than four negative controls per sequencing reaction. Additionally, a "mock community" containing a known quantity of DNA from a mixture of known fungal isolates should be included as a positive control (*see* **Note 9**). Template DNA plates can be stored in the freezer (−20 °C) for up to 2 years or archived frozen (−80 °C) for long-term storage.

3.4 Amplicon Library Generation

1. Optimize PCR parameters: Use a random subset of at least eight samples to test parameters for stage 1 and 2 PCR (below), checking for robust amplification and nonspecific amplification by gel electrophoresis. To limit bias and incorporation of PCR errors, the number of cycles should be minimized. In addition, temperature ramp rate should be limited to 1 °C/s and the PCR enzyme manufacturer's recommended extension time should be increased by 15 s to limit formation of chimeras during PCR [16]. Other parameters that can be varied include annealing temperature, template concentration, and reaction volume (*see* **Note 10**). Other parameters, such as denaturation and extension temperature are determined by the polymerase and buffers used (*see* manufacturer's documentation).
2. First-stage PCR (~25–30 cycles): If practical, given the numbers of samples being processed, check for amplification in all stage 1 PCR reactions by gel electrophoresis. Otherwise, test at least one sample per column for each 96-well plate plus positive and negative control wells.
3. Optional: Cleanup after first-stage PCR (removes PCR buffer, remaining nucleotides, primer dimers, and polymerase). In some cases, primary PCR product includes an abundance of unused primers, which can increase the generation of primer dimers and amplification of short nonfungal amplicons during secondary PCR. To prevent these dimers from proliferating in the stage 2 PCR, primary PCR products should be cleaned prior to use as secondary PCR template (*see* **Note 5**).
4. Second-stage PCR (~5 cycles, *see* **Note 3**): Use gel electrophoresis to check for amplification of one sample per column for

each 96-well plate plus positive and negative control wells. Note that the second-stage PCR product will contain primary PCR product, so these primary amplicons may be visible via gel electrophoresis even if the second-stage PCR fails. When visualizing primary and secondary PCR products simultaneously, the secondary PCR product should be noticeably larger (i.e., the expected size increase should equal the length of overhanging sequencing adapters and molecular identifiers from the secondary PCR).

5. Cleanup, normalization, and pooling: Use a plate-based cleanup and normalization kit (e.g., SequalPrep) to bind a fixed quantity of PCR product within each well and to remove excess DNA, primer dimer, and remaining PCR reagents. After elution, combine equal volumes of the normalized PCR product from each well into a microcentrifuge tube to create the final amplicon library.
6. Library quality control: Before sequencing, ensure that the library contains amplicons of the expected length and is free of residual primer dimers or other short DNA fragments (<100 bp). Precise measurement of the amplicon size distribution can be achieved with a BioAnalyzer or TapeStation. If short fragments are detected (indicating a problem with earlier cleanup steps), an additional magnetic-bead-based size-selection will be needed. The molarity of the library is then determined using qPCR or a BioAnalyzer and adjusted to the concentration required by the sequencing facility. In most cases, final quantification, along with verification of amplicon size distribution, will be performed by the sequencing facility. However, at a minimum, the library concentration should be fluorometrically quantified (e.g., Qubit) and adjusted to the concentration required by the sequencing facility, usually ~10 nM for Illumina MiSeq.

3.5 Sequence Processing and Bioinformatics

A complete discussion of the computational methods used to convert high-throughput sequencing data into fungal community data for downstream ecological analyses is beyond the scope of this chapter. Fortunately, there are many computational tools and bioinformatics pipelines designed explicitly for this purpose (e.g., DADA2 [17], mothur [18], PipeCraft [19], PIPITS [20], QIIME [21], and UPARSE [22]), each of which is accompanied by tutorials and documentation. Irrespective of the choice of computational tools, any high-throughput sequencing pipeline will involve steps to filter low quality sequences and methodological artifacts (e.g., chimeras) and ultimately to assign sequences into operational taxonomic units (OTUs) based on sequence similarity. While robust patterns of community composition should be

recovered regardless of OTU clustering methods chosen, it is important to recognize that the assignment of sequences to OTUs will vary among clustering algorithms (e.g., agglomerative vs. greedy clustering [23]). In addition, because many tools were initially designed for bacterial 16S amplicon sequencing, default parameters in a sequence processing pipeline should be modified for fungal data (e.g., ITS amplicons). Most importantly, no default pipeline should be used without substantial data exploration and parameter optimization.

4 Notes

1. Disrupting leaf tissue with a bead mill homogenizer (i.e., bead beating) is critical for ensuring lysis of fungal cells within leaves, yet is not always included in plant and soil DNA extraction kit protocols. Bead beating can be incorporated into the first step of most DNA extraction kit protocols by adding the sample, initial lysis buffer, and 3–5 small (~2 mm) beads (stainless steel or zirconium oxide) into a reinforced screw cap tube. Samples should be homogenized at high speed in the bead mill for at least 1 min. Alternatively, lyophilized leaf tissue can be ground using a bead mill homogenizer prior to the addition of lysis buffer, though additional homogenization in lysis buffer will further improve yields. When working with new plant species, extra samples should be collected to optimize bead beating. Improved homogenization can be achieved through the use of heavier beads (e.g., stainless steel) and by increasing the oscillation rate or run time of the ball mill. In addition, greater homogenization efficiency can often be achieved by minimizing the volume of the bead tube by using 0.5–1 mL skirted, screw-cap tubes or 96-well plates. If a bead mill homogenizer is unavailable, tissue can be homogenized prior to DNA extraction under liquid nitrogen, using a micropestle and a 1.5 mL microcentrifuge tube rated for cryogenic temperatures
2. The most common gene targets used in fungal metabarcoding studies include variable regions of the small and large subunits (SSU and LSU) and Internal Transcribed Spacer regions (ITS1 and ITS2) of the rRNA gene repeat. While there is no single best locus, the ideal gene marker is conserved across fungi, has nonbiased priming sites, and has high interspecies sequence variation but low intraspecies variation [4]. The ITS regions fit these criteria and have been designated as the official fungal barcode for systematics [24]. Prior to the widespread use of high-throughput sequencing, most fungal sequencing libraries (either cloned or directly sequenced via Sanger sequencing) were sequenced with the robust primer pair ITS1F [25] and

ITS4 [26], which will amplify the entire ITS1 and ITS2 regions, including the 5.8S ribosomal subunit. Because the ITS1F primer is highly selective for fungi (i.e., excluding most plants) and amplifies a wide range of fungi with efficiency (with the exception of Glomeromycota), many studies have focused on the ITS1 region by pairing ITS1F with a reverse primer in the 5.8S region [26, 27]. Related degenerate primers, or mixes of primers with varying bases in some positions, have now been introduced to increase the taxonomic coverage of ITS amplification [28, 29]. However, due to the presence of introns near the ITS1F priming site in some fungi and significant length heterogeneity in this region, ITS2 is increasingly being chosen as a target locus [30, 31], though priming sites for this region are often less selective. As a result, PNA clamps [32] or blocking oligos [33] may be needed to prevent host ITS2 amplification.

The LSU region of the nrDNA gene is more phylogenetically conserved than the ITS region and is a better target if analyses will include questions on evolutionary ecology [34, 35]. However, the number of primers designed that target LSU specifically for short read Illumina chemistry are fewer compared to ITS regions (*but see,* [36]). Consequently, most studies utilize larger targets designed for 454-pyrosequencing, and only use the forward Illumina sequence read [37]. Because this region is highly conserved, host amplification blockers may be required for successful fungal community analyses in the leaf mycobiome.

While gene target selection should be optimized for the specific research questions, several studies have shown that ecological patterns inferred from fungal metabarcode data are generally consistent across gene markers [34, 38, 39]. Nonetheless, selection of gene primers to amplify target loci has potential to introduce taxon-specific amplification bias and skew the relationship between species abundance and the number of sequences obtained [30, 40, 41]. Additionally, the long primer overhangs used for amplicon library generation can result in disproportionate amplification due to interactions between primers and DNA template adjacent to priming sites [42]. To minimize the chances of large, unexpected primer bias, it is best to select primers based on previously reported error rates and community representation from systems similar to the one being studied.

3. While numerous approaches to amplicon library generation are available, a dual-indexed, two-stage PCR approach is recommended as a highly flexible and cost-effective option. This is because the stage 2 primers, which include sample specific molecular identifiers (MIDs), can be purchased once and used

for generating amplicon libraries targeting various loci by using different gene primers in stage 1. In addition, dual-indexed library constructs include two MIDs, allowing each unique pair of forward and reverse stage 2 primers to be used as a sample identifier for multiplexing. In this approach, primers for the first stage of amplicon library generation will include the gene target primer at the 3′ end with the addition of a 5′ overhang consisting of ~25 bp of the platform specific sequencing adapter. In order to increase base diversity during sequencing, stage 1 primers should include a variable length (3–6 bp) degenerate spacer between the partial sequencing adapter and the gene primer (5′-partial sequencing adapter–[3–6 bp Ns]–gene primer-3′) [32]. Separate stage 1 primers must be ordered for each length variant (i.e., 3 Ns, 4 Ns, etc.) and mixed in equal concentrations prior to use. Stage 2 primers complement the partial sequencing adapter added in stage 1 and include a 5′ overhang consisting of the sample-specific MID and the Illumina sequencing linker, which binds the amplicon to the Illumina flowcell in the correct orientation during sequencing (e.g., 5′-linker–MID–partial sequencing adapter-3′). For an example of two-stage PCR amplicon library preparation in practice, see [43].

4. We recommend use of a high-fidelity polymerase to limit PCR errors, which can lead to inflated species richness estimates [44, 45] and increased rates of chimera formation [46]. In addition, high-fidelity polymerase can facilitate the amplification of taxa with primer mismatches [45], resulting in broader taxonomic coverage. However, be aware that the ability of high-fidelity polymerase to "edit" primers to match template sequences [45] can have unintended consequences, such as host plant DNA amplification.
5. Magnetic bead-based cleanup and plate-based cleanup kits with normalization are common methods. Bead-based kits (e.g., AmPure) require an additional DNA quantification step, manual pooling of samples, and additional equipment (magnetic plate).
6. Specific methods for selecting leaves will depend on the particular study question. For example, leaves with visible pathogen damage should be avoided if the goal is to query endophytes. Inclusion of diseased tissues can flood the sample with pathogen DNA, potentially occluding amplification and sequencing of DNA from endophytes.
7. If working in areas where samples cannot be kept in ice-filled coolers after collection, it may be prudent to field-wash leaves (following the same laboratory method of washing in dilute Triton-X 100 solution followed by rinsing in sterilized water)

and punch leaf discs directly into microcentrifuge tubes. Tubes can then be placed into a small cryoshipper that can be attached to a backpack and keeps samples frozen for up to 7 days. Alternatively, leaf discs can be placed directly into tubes containing a lysis buffer (e.g., CTAB—Cetrimonium bromide), which will halt fungal growth and can be used directly for sample homogenization during the first step of DNA extraction.

8. Specific methods for washing leaves will depend on the particular study question. If targeting foliar endophytes (i.e., non-pathogenic fungi that reside within the leaf), epiphytes (i.e., fungi that reside on the surface of the leaf) and spore and hyphal fragments on the surface of the leaf should be removed using a strong surfactant. Surfactant solution and rinse water should be changed between each sample, or as often as possible. This method can be modified for high-throughput processing to expedite processing [14]. Alternatively, perturbation of surface dwelling fungi may be a concern if the study will also examine epiphytes. Exclusive sampling of epiphytes can be accomplished by washing leaves and collecting the liquid, which can be filtered or pelleted via centrifugation. Finally, note that many published studies on endophyte communities include a surface sterilization step using bleach. This works well for culture-based investigations; however, bleach ***will not*** destroy DNA on the leaf surface or remove fungal particles embedded in microcrevasses on the leaf surface and therefore will not eliminate epiphytes and surface adhered fungi from the results. If surface sterilization is desired, because samples will be used for both culturing and sequencing fungal endophytes, a mixture of bleach and a surfactant should be used.
9. In addition to serving as a positive control in PCR reactions, mock community data are useful for downstream bioinformatics and clustering optimization [15, 47] as well as calculating library-wide error rates [2]. Mock communities are generally constructed from pure culture isolates of 10–20 fungal taxa, representing a wide phylogenetic range. The choice of species to include will depend on the expected composition of the fungal communities being studied and the specific research questions, though they do not necessarily need to be members of the sampled communities [48]. Genomic DNA for each isolate is first extracted and used for Sanger sequencing of the target gene region to create a database of expected sequences. DNA extracts from individual isolates are then quantified and pooled by either normalizing to a consistent concentration or by staggering concentrations over several orders of magnitude. Normalized gDNA concentrations can be used to determine the amount of taxon-specific bias

stemming from nrDNA copy number, GC content, or other sources, while a staggered concentration can be used to determine the sensitivity for detecting species with low relative abundance.

10. Annealing temperature is dictated predominantly by the melting temperature of the primers, which is determined by their specific nucleotide composition and the chemical composition of the PCR reaction buffer. Most manufacturers of PCR reagents provide specific guidelines or online tools to calculate an appropriate annealing temperature. When using these tools, only the portion of the primers that complement the priming site should be considered, excluding any overhanging sequence (e.g., Illumina adapters, MIDs, etc.). Also, with mixed or degenerate primers it is important to consider the range of possible melting temperatures for the primer variants. Whenever possible, the annealing temperature should be set at least a few degrees lower than recommendations to facilitate amplification of mixed templates and reduce amplification biases. When visualizing PCR product during PCR optimization, length variation of the target loci in different taxa will result in a somewhat diffuse band or a few bands clustered within the expected range. If unexpected bands or excessive smearing are seen, gradient PCR can be used to help select an annealing temperature that will limit nonspecific amplification.

    Template concentration should be standardized as much as possible across samples to ensure consistent sampling effort and PCR efficiency. When DNA extraction efficiency is highly variable, normalization can be achieved by quantifying the amount of DNA recovered in each extraction and diluting to a consistent concentration across samples (generally 5 ng/μL–25 ng/μL). However, optimizing DNA extraction (beat beating, etc.) to minimize variation in extraction efficiency and using a consistent amount of plant tissue is often sufficient. Samples should also be tested for the presence of PCR inhibitors by running replicate PCR reactions with various dilutions of the same samples. If more dilute samples result in more robust amplification, inhibitors are present and all samples should be diluted by a fixed amount (i.e., the minimum amount necessary to achieve consistent amplification). When the concentration of fungal DNA is expected to be very low and diluting templates is not desirable, increasing the PCR reaction volume without increasing the volume of template used can often facilitate robust amplification, at the expense of using more PCR reagents. Additional modifications, such as the inclusion of DMSO or betaine can facilitate amplification when dealing with difficult templates and specific recommendations are often provided by polymerase manufacturers.

Acknowledgments

The authors are grateful to Maggie Wagner, Edward Barge, Kyle Gervers, Thomas Wolpert, and Wenbo Ma for helpful feedback in the development of this protocol. This research was supported by the Department of Energy Feedstock Genomics Award 219086 (PEB) and the National Science Foundation Science Engineering and Education for Sustainability Award 1314095 (PEB).

References

1. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N et al (2012) Ultra-hight-throughput microbial analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6:1621–1624
2. Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One 6:e27310
3. Lee CK, Herbold CW, Polson SW, Wommack KE, Williamson SJ, McDonald IR et al (2012) Groundtruthing next-gen sequencing for microbial ecology-biases and errors in community structure estimates from PCR amplicon pyrosequencing. PLoS One 7:e44224
4. Lindahl BD, Nilsson RH, Tedersoo L, Abarenkov K, Carlsen T, Kjøller R et al (2013) Fungal community analysis by high-throughput sequencing of amplified markers—a user's guide. New Phytol 199:288–299
5. Arnold AE, Maynard Z, Gilbert GS, Coley PD, Kusar TA (2000) Are tropical endophytes hyperdiverse? Ecol Lett 3:267–274
6. Hardoim PR, van Overbeek LS, Berg G, Pirttilä AM, Compant S, Campisano A et al (2015) The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. Microbiol Mol Biol Rev 79:293–320
7. Herbold CW, Pelikan C, Kuzyk O, Hausmann B, Angel R, Berry D et al (2015) A flexible and economical barcoding approach for highly multiplexed amplicon sequencing of diverse target genes. Front Microbiol 6:1–8
8. Harrison JG, Forister ML, Parchman TL, Kock GW (2016) Vertical stratification of the foliar fungal community in the world's tallest trees. Am J Bot 103:2087–2095
9. Bahnweg G, Heller W, Stich S, Knappe C, Betz G, Heerdt C et al (2005) Beech leaf colonization by the endophyte *Apiognomonia errabunda* dramatically depends on light exposure and climatic conditions. Plant Biol 7:659–669
10. Rai M, Agarkar G (2016) Plant-fungal interactions: what triggers the fungi to switch among lifestyles? Crit Rev Microbiol 42:428–438
11. López-González RC, Gómez-Cornelio S, De la Rosa-García SC, Garrido E, Oropeza-Mariano O et al (2017) The age of lima bean leaves influences the richness and diversity of the endophytic fungal community, but not the antagonistic effect of endophytes against *Colletotrichum lindemuthianum*. Fungal Ecol 26:1–10
12. Erickson RO, Michelini FJ (1957) The plastochron index. Am J Bot 44:297–305
13. Meicenheimer RD (2014) The plastochron index: still useful after nearly six decades. Am J Bot 101:1821–1835
14. Greenfield M, Pareja R, Ortiz V, Gómez-Jiménez MI, Vega FE, Parsa S (2015) A novel method to scale up fungal endophyte isolations. Biocontrol Sci Tech 25(10):1208–1212
15. Nguyen NH, Smith D, Peay K, Kennedy P (2015) Parsing ecological signal from noise in next generation amplicon sequencing. New Phytol 205:1289–1393
16. Stevens JL, Jackson RL, Olson JB (2013) Slowing PCR ramp speed reduces chimera formation from environmental samples. J Microbiol Methods 93:203–205
17. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inferences from Illumina amplicon data. Nat Methods 13:581–583
18. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB et al (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541
19. Anslan S, Bahram M, Hiiesalu I, Tedersoo L (2017) PipeCraft: flexible open-source toolkit for bioinformatics analysis of custom high-throughput amplicon sequencing data. Mol

Ecol Resour 17(6):e234–e240. https://doi.org/10.1111/1755-0998.12692

20. Gweon HS, Oliver A, Taylor J, Booth T, Gibbs M, Read DS et al (2015) PIPITS: an automated pipeline for analysis of fungal internal transcribed spacer sequences from the Illumina sequencing platform. Methods Ecol Evol 6:973–980
21. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336
22. Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 10:996–998
23. Schmidt TSB, Matias Rodrigues JF, von Mering C (2015) Limits to robustness and reproducibility in the demarcation of operational taxonomic units. Environ Microbiol 17:1689–1706
24. Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA et al (2012) Nuclear ribosomal internal transcribed space (ITS) region as a universal DNA barcode marker for fungi. PNAS 109:6241–6246
25. Gardes M, Bruns TD (1993) ITS primers with enhanced specificity for basidiomycetes—application to the identification of mycorrhizae and rusts. Mol Ecol 2:113–118
26. White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky J, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, London, pp 315–322
27. Smith DP, Peay KG (2014) Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. PLoS One 9(2):e90234
28. Tedersoo L, Bahram M, Polme S, Koljalg U, Yorou NS, Wijesundera R et al (2014) Global diversity and geography of soil fungi. Science 346:1256688–1256688
29. Toju H, Tanabe AS, Yamamoto S, Sato H (2012) High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. PLoS One 7(7):e40863
30. Ihrmark K, Bödeker ITM, Cruz-Martinez K, Friberg H, Kubartova A, Schenck J et al (2012) New primers to amplify the fungal ITS2 region—evaluation by 454-sequencing of artificial and natural communities. FEMS Microbiol Ecol 82:666–677
31. Taylor DL, Walters WA, Lennon NJ, Bochicchio J, Krohn A, Caporaso JG et al (2016) Accurate estimation of fungal diversity and abundance through improved lineage-specific primers optimized for Illumina amplicon sequencing. App Environ Microbiol 82:7217–7226
32. Lundberg DS, Yourstone S, Mieczkowski P, Jones CD, Dangl JL (2013) Practical innovations for high-throughput amplicon sequencing. Nat Methods 10:999–1002
33. Vestheim H, Jarman SN (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples—a case study on prey DNA in Antarctic krill stomachs. Front Zool 5:12
34. Brown SP, Rigdon-Huss AR, Jumpponen A (2014) Analyses of ITS and LSU gene regions provide congruent results on fungal community responses. Fungal Ecol 9:65–68
35. Halwachs B, Madhusudhan N, Krause R, Nilsson RH, Moissl-Eichinger C, Högenauer C et al (2017) Critical issues in mycobiota analysis. Front Microbiol 8:180. https://doi.org/10.3389/fmicb.2017.00180
36. Mueller RC, Gallegos LV, Kuske CR (2016) A new fungal large subunit ribosomal RNA primer for high-throughput sequencing surveys. FEMS Microbiol Ecol 92:fiv153
37. Amend AS, Martiny AC, Allison SD, Berlemont R, Goulden ML, Lu Y et al (2016) Microbial response to simulated global change is phylogenetically conserved and linked with functional potential. ISME J 10:109–118
38. Berriti A, Desirò A, Visentin S, Zecca O, Bonfante P (2017) ITS fungal barcoding primers vs 18S AMF-specific primers reveal similar AMF-based diversity patterns in roots and soils of three mountain systems. Environ Microbiol Rep 9(5):658–667. https://doi.org/10.1111/1758-2229.12574
39. Blaalid R, Kumar S, Nilsson RH, Abarenkov K, Kirk PM, Kauserud H (2013) ITS1 versus ITS2 as DNA metabarcodes for fungi. Mol Ecol Resour 13:218–224
40. Bellemain E, Carlsen T, Brochmann C, Coissas E, Taberlet P, Kauserud H (2010) ITS as an environmental DNA barcode for fungi: an *in silico* approach reveals potential PCR biases. BMC Microbiol 10:189
41. Anderson IC, Cairney JWG (2004) Diversity and ecology of soil fungal communities: increased understanding through the application of molecular techniques. Environ Microbiol 6:769–779
42. Berry D, Mahfoudh KB, Wagner M, Loy A (2011) Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. App Environ Microbiol 77:7846–7849
43. Toju H, Tanabe AS, Ishii HA (2016) Ericaceous plant-fungus network in a harsh alpine-subalpine environment. Mol Ecol 25:3242–3257

44. Oliver AK, Brown SP, Callaham MA Jr, Jumpponen A (2015) Polymerase matters: non-proofreading enzymes inflate fungal community richness estimates by up to 15%. Fungal Ecol 15:86–89
45. Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A, Becker A, Gould TJ, Clayton JB, Johnson TJ, Hunter R, Knights D, Beckman KB (2016) Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. Nat Biotechnol 34(9):942–952
46. Lahr DJG, Katz LA (2009) Reducing the impact of PCR-mediated recombination in molecular evolution and environmental studies using a new-generation high-fidelity DNA polymerase. BioTechniques 47:857–863
47. Aas AB, Davey ML, Kauserud H (2017) ITS all right mama: investigating the formation of chimeric sequences in the ITS2 region by DNA metabarcoding analysis of fungal mock communities in different complexities. Mol Ecol Resour 17:730–741
48. Brown SP, Ferrer A, Dalling JW, Heath KD (2016) Don't put all your eggs in one basket: a cost-effective and powerful method to optimize primer choice for rRNA environmental community analysis using the Fluidigm access array. Mol Ecol Resour 16:946–956