SHORT COMMENTARY

A Unique Signal Distorts the Perception of Species Richness and Composition in High-Throughput Sequencing Surveys of Microbial Communities: a Case Study of Fungi in Indoor Dust

Rachel I. Adams · Anthony S. Amend · John W. Taylor · Thomas D. Bruns

Abstract Sequence-based surveys of microorganisms in varied environments have found extremely diverse assemblages. A standard practice in current high-throughput sequence (HTS) approaches in microbial ecology is to sequence the composition of many environmental samples at once by pooling amplicon libraries at a common concentration before processing on one run of a sequencing platform. Biomass of the target taxa, however, is not typically determined prior to HTS, and here, we show that when abundances of the samples differ to a large degree, this standard practice can lead to a perceived bias in community richness and composition. Fungal signal in settled dust of five university teaching laboratory classrooms, one of which was used for a mycology course, was surveyed. The fungal richness and composition in the dust of the nonmycology classrooms were remarkably similar to each other, while the mycology classroom was dominated by abundantly sporulating specimen fungi, particularly puffballs, and appeared to have a lower overall richness based on rarefaction curves and richness estimators. The fungal biomass was three to five times higher in the mycology classroom than the other classrooms, indicating that fungi added to the mycology classroom swamped the background fungi present in indoor air. Thus, the high abundance of a few taxa can skew the perception of richness and composition when samples are sequenced to an even depth. Next, we used in silico manipulations of the observed data to confirm that a unique signature can be identified with HTS approaches when the source is abundant, whether or not the taxon identity is distinct. Lastly, aerobiology of indoor fungi is discussed.

R. I. Adams (✉) · J. W. Taylor · T. D. Bruns
Department of Plant & Microbial Biology, University of California, Berkeley, CA 94720, USA
e-mail: adamsri@berkeley.edu

A. S. Amend
Department of Botany, University of Hawaii at Manoa, Honolulu, HI 96822, USA

Introduction

Use of high-throughput sequencing (HTS) has transformed the field of microbial ecology. Traditionally, sampling depth, defined both in terms of the number of samples and intensity of analysis within those samples, has been limited due to resources required to assess microbial composition. However, with modern HTS technologies, a comparable effort can result in hundreds of thousands and even millions of sequence reads in a single run [1–3]. These tools have been directed toward understanding the presence and role of microbes in terrestrial [4] and marine [5] environments, including where the microbes associate with "macrobes" such as plants [6], insects [7], and humans [8].

One of the advantages of HTS is that a high number of different environmental samples can be processed simultaneously through "multiplexing," where individual samples are PCR-amplified with barcoded primers that are later used to computationally separate them. The accepted strategy is to pool these amplicon libraries at equimolar concentrations before sequencing together. Despite common concentrations, sequence counts are uneven across the samples, so in the analysis stage, the sequence reads are rarefied to a common number across all samples. The resulting assumption is that an even sampling depth, determined by both the concentration of amplified DNA and subsequent rarefaction, will give a representative glimpse of the true biological community within.

ⓐ Springer

One environment that is increasingly a target for microbial surveys is that of the built environment [9]. Not only a unique ecological habitat, the indoor environment is one that can have important implications for human health [10], and identifying exposure often involves detecting a unique contaminant or an elevated level of a common source. Recent studies using culture-independent techniques show that the outdoors is the dominant origin for fungi detected indoors, and this pattern occurs at both global and local scales and independent of building function or human activity [11–13]. In a recent study of residences, no statistical differences in fungal assemblages could be found between rooms such as kitchens, bathrooms, bedrooms, and living rooms, even though they have obvious differences in uses, water sources, and potential substrates for fungi; instead, each room appeared to be a random subset of the outdoor fungal community [12]. This latter result may simply show again that outdoor sources dominate the process of fungal community assembly. Alternatively, it may show a limitation of sampling dust with a HTS approach, in which the accumulated contribution of extremely diverse outdoor sources (composed of hundreds to thousands of taxa) masks the few active residents of individual rooms, a particular concern if trying to use HTS methods to identify "sick buildings."

In preparation for the global study of settled dust [11] and as part of a class study, we sampled vacuumed dust in the mycology teaching lab at University of California, Berkeley, analyzed it via pyrosequencing, and found an unexpectedly high representation of puffball taxa that were clearly derived from material previously used in teaching. This observation provided us with a room with a unique fungal signature in airborne dust. In the current study, we followed up on this observation by passively sampling circulating dust over the course of one semester in this lab and in four adjacent labs housed on the same floor of a single building. We then used multiple in silico manipulates of the observed data to test the sensitivity of the method to identify community differences between rooms. Specifically, we reassigned puffball reads to taxa that were already dominant in the background community, and we lowered the read abundance of the abundant puffball sequence to match that of common indoor fungi. Together, this study allowed us to explore two questions as follows: (1) How sensitive are estimates of richness to differences in abundance? (2) What is required for an assemblage to appear statistically different from other assemblages when the differences are layered on top of a diverse background community?

Materials and Methods

Sample Collection

Samples were taken from five laboratory-type classrooms located on the second story of a two-story building with no reported mold problems (Supplementary Fig. 1). The passive collection of settled dust on a sterile sampling surface follows that of the "dustfall collector" developed by Würtz et al. [14]. Here, we used a base of an empty 9-cm petri dish, a method previously used to sample bioaerosols in residences [12], to collect dust for the duration of the fall teaching semester, from late August 2011 through mid-December 2011. Samplers were placed on top of shelves at a height of 2.1 m and at a minimum of 1 m from a supply air outlet. The volume of the classrooms ranged from 340 to 377 m^3. Four pairs of samplers were distributed in each lab (Supplementary Fig. 1). Among the classes held in classroom E was a mycology course in which both instructors and students brought local specimens into the room for study. Other rooms held laboratory courses on various topics related to microbiology and plant biology but not mycology. All of the labs have identical custodial regimes and ventilation treatments, in that fume hood air vents directly outside of the building and return vent air is circulated through the ventilation system (Supplementary Fig. 1).

DNA Extraction, Amplification, and Sequencing

Sample processing of dust followed the methods of Adams et al. [12]. Briefly, dust was first collected from the dish on a sterile, cotton-tipped, wooden-handled swab by moistening the swab in nucleic acid-free water, rubbing across the dish for 5 s, twisting the dish 90°, and rubbing for another 5 s. Genomic DNA was extracted from the swab tip using a chloroform–phenol procedure in conjunction with a soil extraction kit. The extraction efficiency of this method was not quantified. We targeted the internal transcribed spacer (ITS) region 1 of the nuclear ribosomal coding cistron [15], using the ITS1F and ITS2 primers [16, 17], as these primers are predicted to amplify a diverse set of fungal lineages [18], the ITS1 region is taxonomically informative [19–21], and shorter amplicon regions tend to increase the number of reads per sequencing effort (compared to, for example, sequencing the entire ITS region, including ITS2). Genomic DNAs from the 40 environmental samples were PCR-amplified in triplicate at a 1:10 dilution, purified, quantified, and pooled at equimolar concentrations for pyrosequencing at the W.M. Keck Center for Comparative and Functional Genomics of the University of Illinois at Urbana-Champaign on 1/8th of a 454 FLX+ picotiter plate. Raw pyrosequences were submitted to the Sequence Read Archive of the National Center for Biotechnology Information under accession number SRP018151. To assess the total fungal biomass in each of the samples, we employed quantitative PCR using the FF2/FR1 primers developed by Zhou et al. [22], which target a locus of nearly invariant length of the fungal 18S ribosomal genes. Standard curves were generated using diluted concentrations of *Penicillium purpurogenum* spores

(see [12] for more details). Results are reported as *P. purpurogenum* spore equivalents and compared across classrooms as relative differences in total fungal biomass.

Pyrosequencing Analysis

Pyrosequence data were processed using the Quantitative Insights Into Microbial Ecology (QIIME) pipeline [23]. Amplicons had to meet a mean quality score of 25 and were limited to those with a minimum length of 200 nucleotide base pairs (bp) and maximum length of 412 bp. Efforts to reduce sequencing errors that are an artifact of pyrosequencing relied on the "denoising" algorithm [24]. The ITS1 spacer region was extracted from the sequences [25], and the subsequent reads checked for chimeras using UCHIME [26] and clustered into operational taxonomic units (OTUs) at 97 % similarity using the USEARCH method [27]. OTUs present in negative controls (nine of the >1,100 taxa identified—Supplementary Table 1) were removed from the samples.

Community Analyses and Taxonomy Assignment

For observed richness rarefaction curves in EstimateS [28], samples were pooled within rooms. Within the R statistical language [29], communities were rarefied to an even sampling depth of 407 sequence reads per sample, as that was the lowest read abundance of a sample after eliminating samples with <100 sequence reads per sample. The presence of indicator taxa (OTUs associated with a category of sample) was tested in QIIME [23] and was included if significant after Bonferroni corrections. For taxonomy assignment, a representative sequence of each OTU was aligned by the BLAST algorithm against the UNITE fungal database [30], version April 13, 2012.

In Silico Manipulations

One of the advantages of the laboratory classroom setting is that it can inform the process of detecting fungal contamination in more typical indoor environments. However, the mycology classroom provides an example of taxonomically unique sources (i.e., puffballs) that are also highly abundant, whereas the more likely situation in buildings is an increase in the abundance of a common fungal taxon. Thus, we manipulated results in silico to disentangle the effects of taxonomic identity and read abundance in order to investigate the process of identifying a unique source. First, we reassigned the nine puffball taxa (Table 1) to fungi that are commonly found growing in buildings and that were already present in the sample. For example, the amplicon reads of the most abundant puffball, *Battarrea*, were added to those of the most abundant nonpuffball taxon, *Epicoccum*, and so on through the nine puffball taxa (Supplementary Table 2). This first manipulation created a situation in which differences between classroom E and other rooms were solely the abundances of the dominant taxa. Second, one puffball taxon had a read abundance of nearly 5,748 reads and the remainder ranged from 15 to 1,498 (mean 127.6, median 19.5), while the next most abundant nonpuffball taxon (an *Epicoccum*) was 1,906. Thus, we reduced the read abundance per sample of *Battarrea* taxon to be on par with the read abundance per sample of that of *Epicoccum* (Supplementary Table 2) by generating random values around the mean and standard deviation read abundance of *Epicoccum*. This, in effect, equalized the read abundance of the taxa, and thus, differences between rooms were solely qualitative (i.e., the identities of the dominant taxa across room were different, not their abundances or frequency). Thus, with the two manipulations, we explore the quantitative and qualitative differences under which environments appear distinct in HTS data. Statistical analysis for community analysis relied on ADONIS (a permutational multivariate ANOVA [31]) using the Morisita–Horn (abundance based), Sorensen (presence–absence based), and Canberra (abundance based and emphasizes rare taxa) indices. We visualized the effect of these manipulations using nonmetric multidimensional scaling based on the Morisita–Horn index, although results are consistent across the three indices.

Results and Discussion

The Link Between Diversity, Richness, and Evenness

Results are based on 34 samples: five samples in each of rooms A and C and eight samples in each of rooms B, D, and E because PCR amplification failed in four samples and produced too few high-quality reads (<407) in two other samples. Observed richness, which was based on pooled samples of between 2,889 (classroom C) and 8,944 (classroom E) sequences per room, did not approach saturation (Fig. 1(I)). Observed richness (and Chao's estimated richness—Supplementary Fig. 2) was lower in classroom E than in all the other rooms (ANOVA: $p<0.001$, Fig. 1(I)), despite it having the highest number of sequences. In contrast to richness, the total fungal biomass seen in classroom E was three to five times higher than the other classrooms, a similar number seen in all the other rooms (ANOVA: $p<0.001$, Fig. 1(II)).

In comparing fungal assemblages across environments (in this case, classrooms), we utilized the common strategy of sequencing equimolar concentrations of each sample. Without prior understanding of the setting, one would conclude that the dust community in classroom E is less rich than that in the other classrooms. In actuality, that environment is likely more rich than the others, in that it contains the "background" fungal assemblage that characterizes the other

Table 1 Identities of the 20 most abundant OTUs along with additional taxa that are significant indicator taxa of a particular environment

Abundance	Taxa	Class	Description	Commonly indoor	Indicator taxa—classroom
1	*Battarrea* sp.	Agaricomycetes	Puffball		+ E
2	*Pisolithus* sp.	Agaricomycetes	Puffball		+ E
3	*Epicoccum nigrum*	Dothideomycetes	Mold	Yes	− E
4	*Cladosporium* sp.	Dothideomycetes	Mold	Yes	− E
5	*Alternaria* sp.	Dothideomycetes	Saprobe	Yes	− E
6	*Penicillium* sp.	Eurotiomycetes	Mold	Yes	
7	*Cladosporium* sp.	Dothideomycetes	Mold	Yes	− E
8	*Nigrospora* sp.	Sordariomycetes	Mold	Yes	
9	*Phoma glomerata*	Dothideomycetes	Plant pathogen	Yes	
10	*Aureobasidium pullulans*	Dothideomycetes	Mold	Yes	
11	*Cryptococcus carnescens*	Tremellomycetes	Yeast		
12	*Cryptococcus victoriae*	Tremellomycetes	Yeast		
13	*Cryptococcus* sp.	Tremellomycetes	Yeast		
14	*Pithomyces chartarum*	Dothideomycetes	Saprobe	Yes	
15	*Stemphylium* sp.	Dothideomycetes	Saprobe	Yes	
16	*Cladosporium* sp.	Dothideomycetes	Mold	Yes	
17	*Battarrea* sp.	Agaricomycetes	Puffball		+ E
18	*Cryptococcus magnus*	Tremellomycetes	Yeast		
19	*Battarrea* sp.	Agaricomycetes	Puffball		+ E
20	*Rhodotorula fujisanensis*	Microbotryomycetes	Mold	Yes	
n/a	*Calvatia* sp.	Agaricomycetes	Puffball		+ E
n/a	*Lycoperdon perlatum*	Agaricomycetes	Puffball		+ E
n/a	*Pisolithus* sp.	Agaricomycetes	Puffball		+ E
n/a	*Handkea subcretacea*	Agaricomycetes	Puffball		+ E
n/a	*Pisolithus* sp.	Agaricomycetes	Puffball		+ E
n/a	*Pisolithus* sp.	Agaricomycetes	Puffball		+ E
n/a	*Teratosphaeria microspora*	Dothideomycetes	Plant pathogen		− E

Taxa appear in the same abundance order (1–20) that they appear in Fig. 1(III). All significant indicator taxa were either positively associated with classroom E (the puffballs) or negatively with classroom E (common fungi in low abundance)

classrooms plus the intentional fungal introductions. Thus, the real fungal community is less diverse in that it is less even but not less rich [32], a distinction lost when sampling biological communities of different abundances to the same depth and analyzed with common richness estimators (e.g., Supplementary Fig. 2). Similarly, the concept of "species density," most often used for plant communities, has been used to describe the number of species in a unit area [33]. In the case of HTS, area is equivalent to concentration of amplified DNA, so in the sample dominated by puffball spores, the species density is low but it would be inaccurate to conclude that total richness is also low. We predict that an increased sampling and sequencing of classroom E, to a concentration three to five times higher than the other classrooms (as

informed by biomass estimates), would result in the detection of that background assemblage in addition to the fungal introductions. Indeed, the Chao estimator (Supplementary Fig. 2) hints that classroom E might be just as rich as classrooms A and C but would take much more sequencing effort to detect.

This same bias, of a high abundance of a particular taxon shrinking the read abundance of other taxa and skewing the detected community structure, has been observed in other studies. In one, individual dust samples inoculated with a high number of spores prior to DNA extraction lost the environmental signal entirely [19]; in another, a community sample that contained an individual taxon with low DNA concentration but high PCR affinity resulted in distorted read abundances for the entire community [34]. In a recent
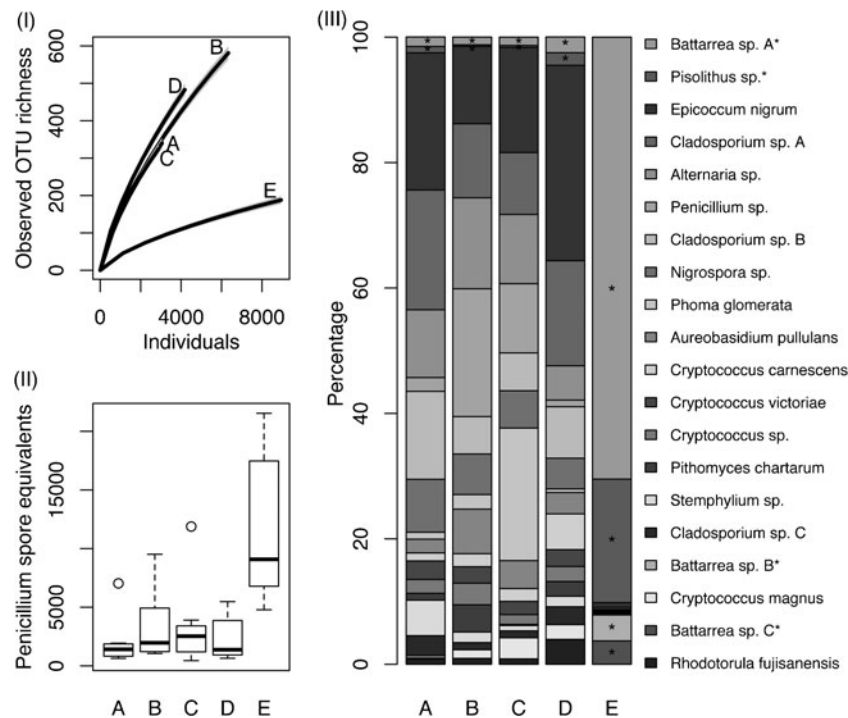
Fig. 1 Observed richness of fungal OTUs in each of the five classrooms sampled. (*I*) Observed fungal OTUs are lower in classroom E where mycology class was conducted than in the other classrooms. *Shading around the lines* represents the standard deviation. (*II*) Fungal biomass in each of the sampled rooms. Fungal biomass is significantly higher in classroom E than in the other classrooms. Box and whisker plot showing the median (*thick black line*), the quartiles (*boxed*), and the extreme values (as *whiskers*). Outliers are shown as *circles*. Biomass is represented as *P. purpurogenum* spore equivalents as measured by qPCR of SSU rRNA gene. (*III*) Proportion of sequence reads for each of the 20 most abundant taxa in each of the classrooms, where each *shade* represents a different OTU. Note that classrooms A–D have a similar composition, while classroom E is distinct. Puffball taxa are indicated with an *asterisk*. The taxa are the same as those listed in Table 1

sampling campaign of residences [12], more singleton species were detected indoors than outdoors. We hypothesized that this pattern emerged not because these rare taxa were not present outdoors but because the outdoor samples, with higher biomass, were, in effect, sequenced to a lower depth when the amplicon libraries were adjusted to equal concentrations prior to sequencing. Consequently, rarer taxa were detected less frequently in the outdoor (higher biomass) samples. Potential biases of observed community structure can thus emerge in several steps of molecular processing, and there are many aspects of the sequencing pipeline that can disrupt the links between true abundance and estimates of richness, evenness, and diversity. If research questions are centered on community membership, it may be more appropriate to pool amplicon libraries based on the concentration of genomic material in the original sample rather than to a common value.

Differences in Community Composition

The community composition in classroom E was different from the other classrooms (ADONIS: $p < 0.001$ for all indices, Supplementary Table 4). The read abundances of the 20 most abundant taxa in all samples show that the compositions of the nonmycology classrooms are similar in their richness and composition to each other (Fig. 1(III)) and to other indoor samples from the area [12] and are more diverse than classroom E (Fig. 1(III)). Taxonomic identity of these 20 OTUs shows that the abundant taxa in the nonmycology classes are largely Dothideomycetes molds, cryptococcal yeasts, and other taxa commonly found indoors (Table 1, Supplementary Table 3). Classroom E, on the other hand, is dominated by a handful of taxa, particularly puffballs that are characterized by spores with a spheroid shape that are released in large quantity when the fruiting body is disturbed. Indicator taxa analysis highlights puffballs as positively associated with classroom E and identifies some common indoor molds as negatively associated with classroom E due to the overrepresentation of puffballs in the sequence reads (Table 1).

Our results showed the typical "long tail" of a taxon appearance with a handful of abundant taxa and the majority of taxa appear only once (Supplementary Fig. 3). The legitimacy of the "rare biosphere" has received a great deal of attention, whether it is a biological reality or an exaggerated artifact of current sequencing techniques (e.g., [35]). Many common beta-diversity metrics were designed for plant community data, in which case, observed abundance matches true

abundance [36]. With HTS approaches, observed and true abundance can be unrelated due to amplification noise (e.g., [34]), so the use of these metrics for sequence data introduces a potential for high-abundance bias. One approach to lessen the impact of these overabundant taxa is by using indices that emphasize rare taxa. In our case, we found that classroom E was distinct from all other classrooms regardless of beta-diversity metric ($p < 0.001$) and, consequently, that the choice of metric did not affect whether samples were different. Rather, we found that the scale of difference, or how distinct samples appear, depended on the choice of beta-diversity metric used to compare them. The percent variation that classroom explained varied greatly by the model used, such that with the abundance-based Morisita–Horn, classroom explained 60 % of the variation; with the binary Sorensen, classroom explained 24 %; and with Canberra, it was down to 12 % (Supplementary Table 4). Thus, our data support the idea that "judicious" use of multiple metrics can be an important approach for uncovering diversity patterns [37].

When we used in silico manipulations to reassign puffball sequence reads in classroom E to other common fungal taxa or to lower the abundance of an overabundant puffball sequence to the same levels as other common fungi, we found that the fungal assemblage in room E remained distinct from those of other classrooms (ADONIS: $p < 0.001$; all dissimilarity metrics, Supplementary Table 4 and Supplementary Fig. 4). While still distinct, the overall community distances between classroom E and the other classrooms decreased under the manipulated scenarios (Supplementary Table 4 and Supplementary Fig. 4). These manipulations suggest that a unique source can be identified in HTS surveys if it is abundant or taxonomically distinct and that it need not be both.

Indoor Air and Fungal Biology

The fact that only puffball sequences were increased in abundance in classroom E is interesting because many other fungi were equally common as class material. While many different types of fungi, including charismatic mushrooms, were brought into classroom E by the students, the puffballs specifically dominate the assemblage. For example, a teaching assistant for the class noted the popularity of several mushroom genera: *Suillus*, *Lactarius*, *Russula*, *Ganoderma*, *Armillaria*, and *Amanita* (Nhu Nguyen, personal communication). However, these taxa were not abundant, in either raw read abundance or frequency of samples, and, for the most part, were no more likely to be detected in classroom E than the other classrooms. It is likely that deeper sequencing of the classroom E samples would find a signature of these additions to the room as well as puffballs. Nevertheless, the preferential detection of puffballs over other mushrooms speaks to the effective strategy for spore dispersal of puffballs, in terms of the raw numbers and/or the ease of airborne transmission of their spores.

Interestingly, a few puffballs appear in trace amounts in the other nonmycology classrooms (Fig. 1(III)), indicating that dispersal between rooms of a building can occur. Because *Battarrea* puffballs are not known to be present on the campus (Tom Bruns, personal observation), the specimens in classroom E are the likely source. Thus, we see three routes for the movement of fungal propagules between rooms, which are separated by two doors and a hallway. One pathway is through tracking: movements made directly by people themselves or their activities. Particles of 7–9 μm in diameter, the approximate spore size of *Pisolithus*, are a size conducive for deposition onto surfaces, including human surfaces [38]. Dispersal of bioaerosols within and between buildings can occur through clothing [39] or custodial equipment. Under these tracking scenarios, the fungal spores transported to the other rooms would have to be resuspended and, subsequently, settle on the high-situated passive sampling locations (2.1 m). The second pathway is a direct transfer through airflow between the rooms, which could occur during the transition between classes when both doors are opened and there is a rush of students moving around the air on the floor. The last pathway is through the building's ventilation system (Supplementary Fig. 1). Further work is required to elucidate the validity of these pathways, but this shows that a fungal source localized to one room can move between rooms throughout the building, even when closed doors separate those rooms.

Conclusions

A localized source can shift the perceived community compositions in HTS approaches as long as it is either taxonomically unique and/or abundant across samples. The unique assemblage will be detected, in part, because as the abundance of a particular taxon rises, the apparent richness of the sample will drop. New sequence platforms may lessen this latter effect, but will eliminate it only if enough sequence depth is employed to plateau the taxon accumulation curves for all samples. Investigation on the benefits of pooling samples for HTS surveys based on sample biomass, rather than at equimolar concentrations of the amplicons as is the common practice, is a worthwhile research avenue to pursue.

References

1. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380. doi:10.1038/Nature03959
2. Metzker ML (2010) Sequencing technologies—the next generation. Nat Rev Genet 11:31–46. doi:10.1038/nrg2626

3. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG et al (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53–59. doi:10.1038/Nature07517

4. Clemmensen KE, Bahr A, Ovaskainen O, Dahlberg A, Ekblad A, Wallander H et al (2013) Roots and associated fungi drive long-term carbon sequestration in boreal forest. Science 339:1615–1618. doi:10.1126/science.1231923

5. Gibbons SM, Caporaso JG, Pirrung M, Field D, Knight R, Gilbert JA (2013) Evidence for a persistent microbial seed bank throughout the global ocean. Proc Natl Acad Sci USA 110:4651–4655. doi:10.1073/pnas.1217767110

6. Zimmerman NB, Vitousek PM (2012) Fungal endophyte communities reflect environmental structuring across a Hawaiian landscape. Proc Natl Acad Sci USA 109:13022–13027. doi:10.1073/Pnas.1209872109

7. Huang X-F, Bakker M, Judd T, Reardon K, Vivanco J (2013) Variations in diversity and richness of gut bacterial communities of termites (*Reticulitermes flavipes*) fed with grassy and woody plant substrates. Microb Ecol 65:531–536. doi:10.1007/s00248-013-0219-y

8. Segata N, Haake S, Mannon P, Lemon K, Waldron L, Gevers D et al (2012) Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. Genome Biol 13:R42

9. Kelley ST, Gilbert JA (2013) Studying the microbiology of the indoor environment. Genome Biol 14:202. doi:10.1186/gb-2013-14-2-202

10. Meng J, Barnes CS, Rosenwasser LJ, Hlth CMCE (2012) Identity of the fungal species present in the homes of asthmatic children. Clin and Exp Allergy 42:1448–1458. doi:10.1111/J.1365-2222.2012.04001.X

11. Amend AS, Seifert KA, Samson R, Bruns TD (2010) Indoor fungal composition is geographically patterned and more diverse in temperate zones than in the tropics. Proc Natl Acad Sci USA 107:13748–13753. doi:10.1073/Pnas.1000454107

12. Adams RI, Miletto M, Taylor JW, Bruns TD (2013) Dispersal in microbes: fungi in indoor air are dominated by outdoor air and show dispersal limitation at short distances. ISME J 7:1262–1273. doi:10.1038/ismej.2013.28

13. Pitkaranta M, Meklin T, Hyvarinen A, Nevalainen A, Paulin L, Auvinen P et al (2011) Molecular profiling of fungal communities in moisture damaged buildings before and after remediation—a comparison of culture-dependent and culture-independent methods. BMC Microbiol 11:235

14. Würtz H, Sigsgaard T, Valbjørn O, Doekes G, Meyer HW (2005) The dustfall collector—a simple passive tool for long-term collection of airborne dust: a project under the Danish Mould in Buildings program (DAMIB). Indoor Air 15:33–40. doi:10.1111/j.1600-0668.2005.00342.x

15. Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA et al (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci USA. doi:10.1073/pnas.1117018109

16. Gardes M, Bruns T (1993) ITS primers with enhanced specificity for basidiomycetes—application to the identification of mycorrhizae and rusts. Mol Ecol 2:113–118

17. White T, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis N, Gelfand D, Sninsky J, White T (eds) PCR: protocols and applications—a laboratory manual. Academic, New York, pp 315–322

18. Bellemain E, Carlsen T, Brochmann C, Coissac 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

19. Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: does read abundance count? Mol Ecol 19:5555–5565. doi:10.1111/J.1365-294x.2010.04898.X

20. Mello A, Napoli C, Murat C, Morin E, Marceddu G, Bonfante P (2011) ITS-1 versus ITS-2 pyrosequencing: a comparison of fungal populations in truffle grounds. Mycologia 103:1184–1193. doi:10.3852/11-027

21. Bazzicalupo AL (2013) Comparison of ITS1 and ITS2 rDNA in 454 sequencing of hyperdiverse fungal communities. Fungal Ecol 6:102–109

22. Zhou G, Whong WZ, Ong T, Chen B (2000) Development of a fungus-specific PCR assay for detecting low-level fungi in an indoor environment. Mol Cell Probes 14:339–348

23. 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. doi:10.1038/Nmeth.F.303

24. Reeder J, Knight R (2010) Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. Nat Methods 7:668–669. doi:10.1038/Nmeth0910-668b

25. Nilsson RH, Veldre V, Hartmann M, Unterseher M, Amend A, Bergsten J et al (2010) An open source software package for automated extraction of ITS1 and ITS2 from fungal ITS sequences for use in high-throughput community assays and molecular ecology. Fungal Ecol 3:284–287. doi:10.1016/J.Funeco.2010.05.002

26. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200. doi:10.1093/Bioinformatics/Btr381

27. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461. doi:10.1093/Bioinformatics/Btq461

28. Colwell RK (2006) EstimateS: statistical estimation of species richness and shared species from samples. http://www.purl.oclc.org/estimates. Accessed 10 Sept 2012

29. R Development Core Team (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org. Accessed 3 Mar 2011

30. Abarenkov K, Nilsson RH, Larsson KH, Alexander IJ, Eberhardt U, Erland S et al (2010) The UNITE database for molecular identification of fungi—recent updates and future perspectives. New Phytol 186:281–285

31. Oksanen J, Blanchet FG, Kindt R, Legendre R, Minchin PR, O'Hara RB, et al. (2012) vegan: community ecology package

32. Shannon CE (1948) A mathematical theory of communication. At&T Tech J 27:623–656

33. Gotelli NJ, Ellison AM (2004) A primer of ecological statistics. Sinauer, Sunderland

34. 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. doi:10.1371/journal.pone.0044224

35. Huse SM, Welch DM, Morrison HG, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ Microbiol 12:1889–1898. doi:10.1111/j.1462-2920.2010.02193.x

36. Bray JR, Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr 27:326–349

37. Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL et al (2011) Navigating the multiple meanings of beta diversity: a roadmap for the practicing ecologist. Ecol Lett 14:19–28. doi:10.1111/j.1461-0248.2010.01552.x

38. Galand PE, Casamayor EO, Kirchman DL, Lovejoy C (2009) Ecology of the rare microbial biosphere of the Arctic Ocean. Proc Natl Acad Sci USA 106:22427–22432. doi:10.1073/Pnas.0908284106

39. Nazaroff WW (2004) Indoor particle dynamics. Indoor Air 14:175–183