

Advancing biodiversity assessments with environmental DNA: Long-read technologies help reveal the drivers of Amazonian fungal diversity

Camila Duarte Ritter¹, Micah Dunthorn², Sten Anslan³, Vitor de Lima⁴, Leho Tedersoo²,
R. Henrik Nilsson⁵, and Alexandre Antonelli⁶

¹University of Duisburg-Essen

²Affiliation not available

³Technische Universität Braunschweig

⁴Universidade Federal de Pernambuco

⁵University of Gothenburg

⁶Royal Botanic Gardens Kew

May 5, 2020

Abstract

Fungi are a key component of tropical biodiversity. Due to their inconspicuous and largely subterranean nature, they are however usually neglected in biodiversity inventories. The goal of this study was to identify the key determinants of fungal richness, community composition, and turnover in tropical rainforests. We tested specifically for the effect of soil properties, habitat, and locality in Amazonia. For these analyses, we used high-throughput sequencing data of short and long reads of fungal DNA present in soil and organic litter samples, combining existing and novel genomic data. Habitat type (phytophysionomies) emerges as the strongest factor in explaining fungal community composition. Naturally open areas – campinas – are the richest habitat overall. Soil properties have different effects depending on the soil layer (litter or mineral soil) and the choice of genetic marker. We suggest that campinas could be a neglected hotspot of fungal diversity. An underlying cause for their rich diversity may be the overall low soil fertility, which increases the reliance on biotic interactions essential for nutrient absorption in these environments, notably ectomycorrhizal fungi–plant associations. Our results highlight the advantages of using both short and long DNA reads produced through high-throughput sequencing to characterize fungal diversity. While short-reads can suffice for diversity and community comparison, long-reads add taxonomic precision and have the potential to reveal population diversity.

Introduction

Fungi are inconspicuous organisms, only a proportion of which sporadically reveal their presence through the formation of tangible morphological structures such as fruiting bodies¹. The scientific study of fungi has therefore benefited immensely from the development of molecular (DNA) sequencing tools during the last 30 years. However, even with the use of molecular tools, studies involving the tropics have neglected fungi despite the fact that the majority of the undescribed fungi are thought to occur in the tropics^{2, 3, 4}. Among all tropical biomes, rainforests provide the widest range of ecosystem services through high above- and below-ground biodiversity⁵, including water cycling and carbon storage^{6, 7}. The largest and most diverse of those

forests is Amazonia^{8,9}, which comprises approximately 40% of the area occupied by rainforest habitats around the world. Amazonian ecosystem services can only be maintained through abiotic and biotic processes, many of which are mediated by fungi.

To better characterize fungal communities in Amazonia, short-read High-Throughput Sequencing (HTS) platforms such as Illumina are being increasingly used^{10,11,12,13,14}. These approaches are often used together with PCR techniques to amplify individual markers. In particular, the nuclear ribosomal Internal Transcribed Spacer (ITS) region has been selected as the best DNA region to identify the widest possible range of fungal groups and is therefore commonly used as a universal DNA barcode for fungi¹⁵. However, this region is typically 500–600 bases long, preventing it from being sequenced under some sequencing technologies. The use of partial sequencing (targeting only a sub-region such as ITS1 or ITS2) has at times limited the taxonomic coverage and identification of fungi by not providing enough variation to tell species apart¹⁶. Furthermore, even though HTS approaches produce hundreds of thousands or millions of sequences per sample, the limited length of these sequences can introduce critical biases to the precise taxonomic identification of the underlying lineages^{17, 18}.

Long-read HTS has the potential to overcome some of these limitations, but they have rarely been used in environmental studies^{18,19}. One of the most well-developed platforms is the single-molecule real-time sequencing platform of Pacific Biosciences (PacBio)²⁰. Although the PacBio platform had a high error rate at the time it was launched, the error rate is currently less than 1%²¹. Recent studies have shown that the potential of the PacBio platform for the identification of fungal communities using environmental samples is high^{18, 19}, but so far they have not been widely applied to any ecosystems.

Taken together, the use of short- and long-sequence HTS techniques offers the potential to overcome the challenges of characterizing fungal diversity in species-rich ecosystems, such as Amazonia in northern South America. Amazonia is a heterogeneous biome, and its biodiversity has been shown to vary considerably across geographical ranges. On a large scale, a west (more diverse) to east (less diverse) diversity gradient has been observed in many animal and plant groups^{22, 23, 24} and also in micro-organisms, including fungi^{10, 25}. Another source of heterogeneity in Amazonia is the presence of distinct habitats types. Each phytophysiognomy comprises a largely distinct biota and own soil characteristics, flooding regimes, and nutrient availability^{11, 26}. Four widespread and important habitats, here given in the order of decreasing plant and animal diversity^{25, 26}, are: unflooded tropical forests (terra-firme); forests seasonally flooded by fertile white-water rivers (várzeas); forests seasonally flooded by unfertile black water rivers (igapós); and naturally open areas associated with white-sand soils (campinas). The richness gradient for micro-organisms has been found to differ from this general trend, as campinas harbour the highest microbial richness^{10, 25}.

Soil physicochemical characteristics are often considered crucial for biotic dynamics, vegetation, and diversity patterns at local to regional scales across Amazonia^{14, 27, 28, 29}. Although several studies have reported on the importance of soil characteristics in shaping community structure, no unified pattern has emerged. In a recent study using HTS with short reads from environmental samples in Amazonia, members of our team showed a mixed effect of soil properties on the microorganism richness and community turnover¹¹. In that study, we used general primers to target all eukaryotes, and we did not address specifically these effects on Fungi.

This study seeks to characterize fungal communities across Amazonia using environmental samples of soil and litter. For the first time (to our knowledge) in an Amazonian context, we use a long-read approach to sequence the full fungal ITS region on the PacBio platform. In addition, we combine our novel long-read data with our previously released short-read HTS data of the nuclear ribosomal 18S rRNA small subunit (18S) gene and the mitochondrial cytochrome c oxidase subunit I (COI) gene produced in a Illumina sequencing platform. We discuss the patterns of fungal richness and community turnover across Amazonia and compare the results obtained from different genes and platforms.

Methods

Study area and sampling design

We sampled four localities across Brazilian Amazonia (Fig. 1) following the sampling design described by Tedersoo et al.¹². Detailed locality descriptions are available in Ritter et al.¹⁰. Briefly, we sampled all depths of the litter layer above the mineral soil (all organic matter, including leaves, roots, and animal debris) and the top 5 cm of the mineral soil in a total of 39 circular plots, each with a radius of 28 m. We chose 20 random trees inside each plot and collected litter and soil on both sides of each tree. We then pooled the samples by substrate to produce one litter sample and one soil sample per plot. The soil physicochemical properties were determined by a Brazilian company (EMBRAPA); additional details of the soil analysis can be found in Ritter et al.¹¹.

Data generation

For the nuclear ribosomal small subunit (SSU) 18S rRNA (18S) and the mitochondrial cytochrome c oxidase subunit I (COI) genes, we used the OTU table produced in Ritter et al.²⁵. We selected the OTUs assigned to the fungal kingdom based on SILVA³⁰ for 18S and GenBank³¹ for COI datasets, respectively, for all our analyses. We present here the results of both markers in light of the fact that the previous publication did not analyse fungi separately, which imposed limits on the fungal richness and community structure analyses employed at the time.

For the ITS, we used the approach described in Tedersoo et al.¹⁸. We used the forward primers ITS9MUNgs (5'-TACACACCGCCCGTCG-3'³²) and ITS4ngsUni (5'-CCTSCSCTTANTDATATGC-3'³²) to target the full ITS region (ITS1 - 5.8S - ITS2)". For amplification, we used a PCR mixture comprised 5 μ l of FirePol DNA polymerase master mix (Solis Biodyne, Tartu, Estonia), 0.5 μ l of each forward and reverse primer (20 mM), 1 μ l of DNA extract and 18 μ l ddH₂O. Thermal cycling included an initial denaturation at 95 °C for 15 min; cycles of denaturation for 30 s at 95 °C, annealing for 30 s, elongation at 72 °C for 1min; final elongation at 72 °C for 10 min and storage at 4 °C. The duplicate PCR samples were pooled; their relative quantity was estimated by running 5 μ l DNA on 1% agarose gel stained with ethidium bromide (Sigma-Aldrich, St Louis, MO, USA). We used negative (for DNA extraction and PCR) controls throughout the experiment. The amplicons were purified with FavorPrep PCR Clean Kit (FavorGen Biotech Corporation, Vienna, Austria). The libraries were prepared using PacBio amplicon library preparation protocol (Pacific Biosciences, Inc, Menlo Park, Ca, USA) and loaded to seven SMRT cells using the MagBead method. The libraries were sequenced using the PacBio RS II instrument using P6-C4 chemistry following the manufacturer's protocol.

Bioinformatics analyses were performed using the PipeCraft platform³³. PacBio circular consensus reads (CCS, reads_of_insert) were quality filtered with vsearch³⁴ (maxee = 2, maxns = 0, minlen = 150). Filtered reads were demultiplexed based on the unique sequence identifiers using mothur³⁵ (bdiffs = 1). Putative chimeric reads were filtered using denovo and reference-database-based methods in vsearch. Additionally, sequences where the full PCR primer was found anywhere in the read were filtered out using the PipeCraft built-in module, as these reads represent additional chimeras not detected by the vsearch method. The full ITS region was extracted using ITSx³⁶ and clustered using the UPARSE algorithm³⁷ with a 98% similarity threshold. Additionally, the post-clustering curation method LULU³⁸ was applied (minimum_ratio_type = "min", minimum_match = 98) to merge consistently co-occurring 'daughter' OTUs. Taxonomy annotation was performed using BLASTn³⁹ against the UNITE^{40, 41} and INSDC⁴² databases.

Statistical analysis

We performed all statistical analyses in R v.3.6.0⁴³. Two samples (SCUICAMP3 and LCUITFP3) had a very low number of reads in the ITS results, and were excluded from subsequent analyses of all markers. We use as a diversity estimative the effective number of OTUs, calculated with the unrarefied read counts as OTU abundance, using the exponential of the Shannon entropy diversity of order $q = 1$ ⁴⁴. This measure is more robust against biases arising from uneven sampling depth than the simple number of OTUs⁴⁵. For

the abundance-based community matrices, we transformed read counts using the “varianceStabilizingTransformation” function in DESeq2⁴⁶ as suggested by McMurdie and Holmes⁴⁵. This transformation normalizes the count data with respect to sample size (number of reads in each sample) and variances, based on fitted dispersion-mean relationships⁴⁶.

We tested the correlation between diversity of each marker through a Pearson correlation between each pair of markers. To test between the community composition correlation, we performed a Mantel test with the Jaccard dissimilarity matrices, using the Pearson correlation and 999 permutations for significance. Both analyses were performed using the vegan v.2.5.5 R package⁴⁷.

For soil physicochemical analysis, we first normalized all variables to mean = 0 and variance = 1. We then performed two principal component analyses (PCA), one for soil grain size and the other for chemical compounds, using the vegan package. We used the first axis of each PCA (explaining 56% and 69% of the total variation, respectively) in the subsequent linear models and multiple regressions analysis. Given the expected importance of soil organic carbon content^{11,48} and pH^{11, 49}, we used these as independent variables.

/

To test the effect of soil properties on fungal OTU richness, we performed a Bayesian general linear model (GLM) analysis, as implemented in the R-INLA v.17.6.20 R package⁵⁰. The response variables were the OTU diversity by soil layer (litter and soils) and marker (18S, ITS and COI), giving a total of six models. In each case, the soil properties (PC1 for the physical, PC1 for the chemical, organic carbon content, and pH both standardized to mean = 0 and variance = 1) were used as explanatory variables. We tested the effect of spatial autocorrelation by comparing analyses of standard GLMs with GLM analysis using stochastic partial differential equations (SPDE) that explicitly consider spatial correlation.

To test the effect of soil properties on fungal community turnover, we used multiple regressions on dissimilarity matrices (MRM) with the R package ecodist v.2.0.1⁵¹. The response variables were dissimilarity matrices calculated using the Jaccard dissimilarity. In each case, the explanatory variables were the distance matrices based on soil properties (physical PC1, chemical PC1, organic carbon, and pH) and one geographical distance matrix (all calculated using Euclidean distances). Statistical significance of the regression coefficients was determined using 10,000 permutations.

For the analysis of differences of community composition by locality and habitat, we performed a non-metric multi-dimensional scaling (NMDS) analysis using the Jaccard dissimilarity matrix and tested the significance of groups using the envfit test, which fits vectors of continuous variables - in this case the NMDS axes - and centroids of levels of class variables (locality, habitat, and soil layer) using the vegan package. Additionally, we performed a permutational analysis of variance (PERMANOVA) to test the significance of each factor (locality, habitat, soil layer, first PC of both PCAs, pH, and carbon) in the community composition of each dataset (18S, COI, and ITS) using the vegan package. We furthermore performed an analysis of indicator OTUs of each locality, habitat, and soil layer using the R package indicspecies v.1.7.6⁵² using the matrix of relative abundance. This analysis identifies the species, in our case the OTUs, that are associated with a determined group. We performed the analysis three times with each dataset (18S, COI, and ITS): the first grouping the OTUs by locality, the second by habitat, and the third by soil layer. We tested the significance with 9,999 permutations, from which we quantified the number of indicator OTUs for each group with an alpha < 0.05. Following this, based on literature and experience, V.X.L. classified all possible indicator OTUs in their functional guild group, such as mycorrhizae, phytopathogen, and saprobe based in their assigned taxonomy (Table S3). We calculated the mean number of OTUs by each factor (locality, habitat, and soil layer) in each dataset (18S, COI, and ITS) using the vegan R package. We produced a Venn diagram for visualization of the number and proportion of exclusive and shared OTUs for each factor (locality, habitat, and soil layer) in each dataset (18S, COI, and ITS) using the online tool Venny 2.0⁵³. Additional R packages used for data curation were tidyverse v.1.2.1⁵⁴ and ggplot2 v.3.1.1⁵⁵. All scripts and data used in the analyses are available as supplementary material.

Results

OTU classification and marker correlation

After sequencing, processing, and filtering of short-read (Illumina) reads, we found a total of 10,745 OTUs, of which 2,212 (20%) were identified as fungi for the 18S dataset. For COI we found a total of 6,227 OTUs, of which 2,161 (35%) were fungal. For the long-read (PacBio) reads of ITS, we obtained a total of 3,711 OTUs, of which 3,039 (82%) were fungal. The majority of the fungal OTUs were found to belong to the phylum Ascomycota, followed by Basidiomycota (Fig. 2). The 18S dataset was found to contain a higher proportion of non-Dikarya (Ascomycota plus Basidiomycota) than did the other datasets (Fig. 2). All the following results are based only on OTUs classified as Fungi.

The Effective number of OTU showed a weak correlation across datasets, with COI being more correlated with 18S ($r = 0.36$). The ITS was non-correlated with either 18S ($r = -0.08$) or COI ($r = -0.02$). The Mantel tests showed a significant ($p = 0.001$) correlation in all matrices of similarity, with the strongest correlation between 18S and COI ($r = 0.52$) and a weaker correlation with the ITS datasets (ITS and COI $r = 0.30$, ITS and 18S $r = 0.17$).

Soil characteristics and their effect on fungal diversity and composition

The principal component analysis (PCA) recovered more than 56% of data variability in the first principal component axis (PC1) in both physical and chemical properties. The PC1 of each PCA was used in further analyses (Fig. 3). In our PCA for physical characteristics, the negative values represent fine texture soils (silt and clay), which are predominantly present in seasonally flooded forests – igapós and várzeas (Fig. 3A). The campinas had plots at both extremes of PC1, having the plots in Jaú and Cuieras localities with fine texture and the others plots localized in Caxiuana with coarse soil textures (Fig. 3A). Terra-firme was more spread across different gradients of the soil texture (Fig. 3A). In the PCA for chemical compounds, positive values in PC1 represent low-fertility soils. Campina and terra-firme were more associated with low-fertility soils, while várzea forests showed different fertility levels (Fig. 3B). Plots in igapó forests also showed low soil fertility except for the plots in Benjamin Constant (Fig. 3B). For details of soil characteristics see Ritter et al¹¹.

Only the mineral soil had some soil properties with significant effect on the OTU Shannon diversity, an effect that varied by marker (Table 1). For 18S, only the organic carbon (C) content was significant, with a negative effect. Organic carbon was also significant and negative for soil ITS diversity. Chemical PC1 was significant for COI and ITS soil diversity, with a higher effective number of OTUs increase following decreasing soil fertility. The pH and soil texture had no significant effect on OTU diversity.

Geographical distance was significant for all datasets. However, since juxtaposed localities are usually similar in many respects, we cannot differentiate the level of spatial correlation from the effect of soil properties in our analysis of community turnover (Table 2). For community turnover, organic carbon and pH were significant for all soil communities (18S, COI and ITS), as was pH for all litter communities. Organic carbon was also significant for the COI litter dataset. Soil texture was significant in all communities except for the ITS soil dataset (Table 2). The PC1 for chemical properties was significant for the 18S and COI litter communities. In the PERMANOVA analysis, the soil properties were all significant with a weak effect on all datasets (Table S3).

The soil layer, organic litter, and mineral soil had a weak but significant effect on the number of OTUs (PERMANOVA results: $p < 0.001$ for all datasets, 18S – $R^2 = 0.05$, COI – $R^2 = 0.04$, and ITS – $R^2 = 0.03$). There were small differences between the soil and litter communities in the two axes of non-metric multi-dimensional scaling (NMDS) in all datasets (Fig. 4). The litter COI and ITS datasets had a higher mean number of OTUs, where a higher number of OTUs is considered litter indicators (OTUs with a significantly higher probability to be found in litter than soil; Table 3), and a high number of exclusive OTUs than 18S (Fig. 5). For 18S, the results contrast with those of the other markers, showing soil as the most diverse substrate, with the highest number of exclusive and indicator OTUs (Table 3, Fig. 5C). The majority of

indicator OTUs for both layers are saprotrophs (Table S2).

The effect of localities

Regarding locality, Benjamin Constant had the most differentiated community in all datasets (Fig. 4). The effect of localities was significant ($p < 0.001$) and had a higher effect than the soil layer factor in explaining the community composition in all datasets (18S – $R^2 = 0.10$, COI – $R^2 = 0.12$, and ITS – $R^2 = 0.11$). In general, the pattern of highest mean, number of exclusive, and number of indicator OTUs by locality varied between markers (Table 3, Fig. 5). For 18S and ITS, Benjamin Constant had the highest mean number of OTUs (Table 3) and the highest number of exclusive OTUs (Fig. 5). Benjamin Constant also had the highest number of OTUs considered indicators of this locality for the 18S dataset (Table 3). Cuieras had the lowest number of exclusive OTUs for COI and 18S, but had the highest number of indicator OTUs for the COI datasets (Fig. 5). The majority of indicator OTUs at all localities were saprotrophs, followed by a high proportion of OTUs that could not be classified by their functional group (Table S2).

The effect of habitats

Habitat type was the strongest factor ($p < 0.002$) in explaining community composition in the PERMANOVA analysis (18S – $R^2 = 0.12$, COI – $R^2 = 0.18$, and ITS – $R^2 = 0.08$), with the exception of the ITS dataset. In NMDS, the seasonally flooded forests, igapós, and várzeas were more similar to each other than to campinas and terra-firmes, which were the most similar to each other (Fig. 5). Campinas had the highest mean number of OTUs and the highest number of OTUs considered indicators of this habitat for all datasets (Table 3). Regarding the number of exclusive OTUs, campinas had the highest number of OTUs in the COI datasets (Fig. 5E). Terra-firme was the habitat with the highest number of exclusive OTUs for the 18S and ITS datasets (Figs. 5B and 5H). In all habitats, the majority of indicator OTUs were saprotrophs, followed by a high proportion of OTUs that could not be classified by their functional group (Table S2). Campinas have a moderate proportion of phytopathogen indicator OTUs and terra-firme a moderate proportion of parasite indicator OTUs (Table S2).

Discussion

Our results highlight the importance of habitat type for fungal community composition in Amazonia and suggest that Amazonian fungi have different diversity patterns for habitat and locality variables, with the importance of each predictor varying between markers. By contrast, community turnover shows a consistent pattern, with habitat being a strong factor explaining community similarity between plots. This is likely to be because different areas can have similar species richness but different species composition due to historical, geographic, and environmental factors. For instance, in a study of leaf litter fungi in Central Amazonia, the abundance and richness of fungal morphospecies did not change between low and high rainfall periods, but there was a low proportion of shared morphospecies between periods⁵⁶. Our results also showed a low proportion of shared OTUs when compared with a HTS study of micro-organisms in general in the same area¹⁰ (Fig. 5).

Soil texture did not explain fungal diversity, while chemical soil characteristics was of importance for COI and ITS soil communities, indicating a high diversity in less fertile soil (Table 1). Although it appears counter-intuitive, the habitat with lowest soil fertility was the one with highest fungal and other microbial diversity: the campinas^{10, 25}. These results suggest that factors other than soil properties explain a habitat's fungal diversity and community composition.

The soil diversity of the 18S dataset was negatively related with carbon, while the specifics of the other datasets were not related to carbon. This could be explained by taxonomic coverage of the 18S dataset, which included the Chytridiomycota and Mucoromycota, which are mostly saprobe groups^{57, 58}. Saprobes decompose matter into various constituent components, making the nutrients available to other organisms. Saprobes are, in other words, important agents in carbon cycling⁵⁹. Hence, a high fungal richness may lead to a faster carbon decomposition in soil as well as a faster carbon assimilation in the above-ground biomass.

This is in agreement with Liu et al.⁶⁰, who found that phylotype richness and phylogenetic diversity of black soil fungi responded negatively to total carbon content in China. Experiments controlling the variables and quantifying the above-ground biomass are necessary to further verify these observations.

Contrary to our expectations, pH had no effect on fungal richness. This finding was surprising, since soils with more neutral pH generally have a higher richness of micro-organisms^{11,61,62,63}. Our soil samples were all acidic, with the pH varying between 3.5 and 5.14. Soil fungi studied by Liu et al.⁶⁰ showed a similar pattern as reported in this study – a higher relative influence of soil carbon content than of soil pH. They also noted that fungi often have a wider tolerance to pH variation than other micro-organisms, suggesting that in soils with low pH variation such as presented here, the acidity impact should be less striking⁶⁰. On the other hand, pH was important to explain community turnover for all datasets (Table 2). Furthermore, in tropical areas the relationship between fungal communities and soil pH is affected by the fungal trophic guilds⁶⁴. It may indicate that in a highly diverse area, such as Amazonia, fungal diversity will not be impacted by pH variation but there will be a turnover of fungal species related with the pH range.

Spatial differences

Different Amazonian habitats varied considerably in their biotic composition^{10, 65}. Habitat was the most significant factor explaining community turnover in the 18S and COI datasets. Although habitat was also significant for ITS, community turnover was better explained by locality for this dataset. In the ITS data, we found that campinas and igapós are dissimilar in their communities (Fig. 4A). This can be explained by the physicochemical soil properties (Fig. 3). When it comes to chemical properties, campinas and igapós were placed at opposite extremes of PC1 and PC2 (Fig. 3B). With respect to the physical properties, campinas have plots in both extremes of PC1, but igapós were better explained by clay content (Fig. 3A). Clay content was an important factor in explaining leaf litter in central Amazon fungi⁵⁶. For the 18S and COI, the similarity between habitats is better explained by comparing seasonally flooded and non-flooded habitats (Fig. 4). In both communities, igapós and várzeas are similar to each other and distinct from terra-firme and campinas. This is in agreement with results from studies of micro-organisms in general in the same areas¹⁰. These results were expected as the flooded period is a powerful factor that selects for a very specific vegetation type^{66, 67, 68,26}. Igapós and várzeas are more restricted to a fine soil texture, while in terra-firme and campinas the soil texture varies more (Fig. 3A). However, regarding the chemical properties, terra-firme and campinas have almost exclusively poor soils, while igapós and várzeas present different gradients of soil fertility (Fig. 3B). These distinct patterns among markers might be explained by the differences in taxonomic coverage of each marker, since different species of fungi have distinct habitat preferences¹².

We were surprised to find that Campinas were, on average, the richest habitat for fungi. This stands in contrast to patterns observed for animals and plants^{69, 70}, and fungi in Colombian Amazonia¹⁴. One explanation for the campinas being the richest environment may be the need for plants to associate with micro-organisms that fix nutrients in the poor soil habitats. For instance, some studies of campinas in Amazonia address the diversity of ectomycorrhizal fungi^{71, 72, 73, 74, 75}. The general pattern is that the diversity of ectomycorrhizal fungal diversity is the highest in temperate zones^{12,76,77}, but due to the poor soil in Campinas, the ectomycorrhizal fungi will be more diverse.

The origin of the campinas in Amazonia is a debated topic⁷⁰, but the nature of their soil, caused by high drainage and high acidity, is considered one of the poorest in the world⁷⁸. In this context, Singer et al.⁷³ hypothesized that the ectomycorrhizal fungi increase the ability of their host plant to acquire nutrients and water in these very stressful habitats. We found a high richness and number of indicator OTUs in campinas (Fig. 5, Table 3), adding more evidence that the campinas may be hotspots for the diversity of fungi and other micro-organisms. However, we registered very few ectomycorrhizal indicator OTUs, although these results could be biased by the lack of representative DNA sequences from tropical areas⁷⁹. It is interesting that várzea areas have on average 43 to 53% fewer OTUs of known mycorrhizal species for the three markers. Of the four habitats analysed, várzea soils are of higher fertility as they are flooded by nutrient-rich waters, decreasing the necessity for plants to associate with mycorrhizal fungi, in accordance with the hypothesis proposed by Singer et al.⁷³.

Comparison between short and long reads

Our results showed a similar pattern for the habitat diversity of long and short-reads, corroborating the patterns previously reported^{10, 11, 25}. These similarities support the view that our findings are real and independent of any possible methodological biases introduced by the different markers and platforms.

The importance of soil properties on the diversity and community turnover varied among markers. We acknowledge the different taxonomic coverages of each marker and the limitations of the available databases. For instance, the diversity of the early-diverging fungal lineages Chytridiomycota, Cryptomycota, and Zoopagomycota using 18S is higher and it is in stark contrast with the ITS and COI data. This difference may be the result of either PCR biases or of shortages of the reference databases used. The COI is usually used as barcode for metazoans⁸⁰, with lower sequence available for fungi. Our COI data showed around 40% of unidentified OTUs²⁵, which could represent at least in part some fungal lineages without public reference sequences. Uneven availability of reference sequences may have had impact on our diversity and community composition results for the various markers used, with the highest effect for the COI results.

The use of short-read fragments (for both 18S and COI) resulted in a higher number of OTUs, for all organisms, than did the long-read technique. Long-read ITS, on the other hand, registered more fungal OTUs even though the total number of OTUs was smaller than for short reads. It is important to stress here that, unlike for the ITS region, for short-reads we used general primers targeting all eukaryotes and not just fungi, such that only a portion of reads belonged to fungi in the 18S and COI datasets. Although the differences in primer design preclude us to reliably identify the “best” marker or sequencing platform choice for fungal assessments in general, we highlight the main advantages and disadvantages of those used here.

On the one hand, we showed that the use of 18S under the Illumina platform provides the overall highest taxonomic coverage. So for studies aiming to compare diversity and community turnover the use of short-reads can be recommended. In economic terms, this is also the more cost-efficient option at the moment. On the other hand, due to the short fragment size of Illumina reads, some OTUs could be potentially misidentified or categorised only at, for example, the family or genus level. For instance, in an earlier study comparing the taxonomic identification of short-read HTS, the choice of the ITS sub-region, ITS1 or ITS2, affected 51% of fungal identifications¹⁶. Long-read HTS methods have the potential to identify fungi with higher accuracy, despite recording fewer sequences per sample¹⁸. In our data, PacBio registered the highest number of OTUs classified as fungi but the lowest number of total OTUs. This is expected, since PacBio platforms have a small number of reads in total⁸¹ and also will not sequence partially degraded DNA. Additionally, long reads have the potential of combining population analysis with environmental data. This is limited with short-reads, which provide a more limited genetic variation for environmental diversity analysis or require the sequencing of several markers for a limited number of target individuals.

Conclusions

Tropical fungal diversity is surprisingly high and poorly understood. In our study, we found that the equivalent to a teaspoon of Amazonian soil can contain as many as 1 800 OTUs, which up to 400 were classified as fungi. It might therefore not be an exaggeration to call fungal diversity the ‘dark matter’ of life on Earth, alongside many other poorly studied groups. Our results highlight the importance of habitat type for fungal community composition. We also show that the known general patterns found for macro-organisms in Amazonia may not apply to fungi. It is important to improve our understanding of the patterns and drivers of fungal diversity and community composition since this is one of the most diverse eukaryotic kingdoms, whose members play key roles in nutrient cycling and biotic interactions in terrestrial ecosystems. Deforestation of Amazonia is increasing rapidly⁸³, and to protect this forest it is fundamental to understand the processes underpinning ecosystem stability. For this, we have to focus on the distribution and diversity of organisms essential for the ecosystem functionality, including fungi.

Acknowledgements

We thank the Brazilian authorities for permits: ICMBio (registration number 48185-2) and IBAMA (registration number 127341). We thank Francisco Diniz for early comments on the manuscript and Rhian Smith for improving the English. We acknowledge funding from the Alexander von Humboldt Foundation and CN-Pq (Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brazil: 249064/2013-8) to C.D.R.; the Deutsche Forschungsgemeinschaft (#DU1319/5-1) to M.D.; the Swedish Research Council, Swedish Foundation for Strategic Research, the Knut and Alice Wallenberg Foundation and the Royal Botanic Gardens, Kew to A.A.

Author contributions

C.D.R., A.A., L.T and R.H.N. designed the study; C.D.R. conducted fieldwork; C.D.R. led all analyses with contributions from S.A.; C.D.R. and S.A. conducted lab work; C.D.R., M.D. and V.X.L. discussed the results. C.D.R. led the writing of the manuscript with contributions from all authors.

Competing interests

The authors declare no conflict of interest.

Data accessibility

All raw 18s and COI sequences are available in GenBank under Bioproject PRJNA464362. PacBio raw sequences are available in GenBank under Bioproject XXXX.

References

1. Moore, D. *Developmental Biology of Higher Fungi: Symposium of the British Mycological Society Held at the University of Manchester April 1984*. (Cambridge University Press, 1985).
2. Hawksworth, D. L. The magnitude of fungal diversity: the 1·5 million species estimate revisited. *Mycol. Res.* **105**, 1422–1432 (2001).
3. Lodge, D. J. *et al.* A survey of patterns of diversity in non-lichenized fungi. *Mitteilungen der Eidgenössischen Forschungsanstalt für Wald, Schnee und Landschaft* **70**, 157–173 (1995).
4. Hawksworth, D. L. & Rossman, A. Y. Where are all the undescribed fungi? *Phytopathology* **87**, 888–891 (1997).
5. Wardle, D. A. *et al.* Ecological linkages between aboveground and belowground biota. *Science (80-)*. **304**, 1629–1633 (2004).
6. Fearnside, P. M. Amazon forest maintenance as a source of environmental services. *An. Acad. Bras. Cienc.* **80**, 101–114 (2008).
7. Ojea, E., Martin-Ortega, J. & Chiabai, A. Defining and classifying ecosystem services for economic valuation: the case of forest water services. *Environ. Sci. Policy* **19**, 1–15 (2012).
8. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science (80-)*. **342**, 850–853 (2013).
9. Antonelli, A. *et al.* Amazonia is the primary source of Neotropical biodiversity. *Proc. Natl. Acad. Sci.* **115**, 6034–6039 (2018).
10. Ritter, C. D. *et al.* Locality or habitat? Exploring predictors of biodiversity in Amazonia. *Ecography (Cop.)*. **42**, (2019).

11. Ritter, C. D. *et al.* High-throughput metabarcoding reveals the effect of physicochemical soil properties on soil and litter biodiversity and community turnover across Amazonia. *PeerJ* **2018** , (2018).
12. Tedersoo, L. *et al.* Global diversity and geography of soil fungi. *Science* **346** , 1052–3 (2014).
13. Dunthorn, M., Kauserud, H., Bass, D., Mayor, J. & Mahé, F. Yeasts dominate soil fungal communities in three lowland Neotropical rainforests. *Environ. Microbiol. Rep.* **9** , 668–675 (2017).
14. Vasco-Palacios, A. M., Bahram, M., Boekhout, T. & Tedersoo, L. Carbon content and pH as important drivers of fungal community structure in three Amazon forests. *Plant Soil* 1–21 (2019).
15. Schoch, C. L. *et al.* Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc. Natl. Acad. Sci.* **109** , 6241–6246 (2012).
16. Nilsson, R. H., Ryberg, M., Abarenkov, K., Sjökvist, E. & Kristiansson, E. The ITS region as a target for characterization of fungal communities using emerging sequencing technologies. *FEMS Microbiol. Lett.* **296** , 97–101 (2009).
17. Nilsson, R. H. *et al.* Mycobiome diversity: high-throughput sequencing and identification of fungi. *Nat. Rev. Microbiol.* **17** , 95–109 (2019).
18. Tedersoo, L., Tooming-Klunderud, A. & Anslan, S. PacBio metabarcoding of Fungi and other eukaryotes: errors, biases and perspectives. *New Phytol.* **217** , 1370–1385 (2018).
19. Purahong, W., Mapook, A., Wu, Y. & Chen, C. Characterization of the *Castanopsis carlesii* Deadwood Mycobiome by Pacbio Sequencing of the Full-Length Fungal Nuclear Ribosomal Internal Transcribed Spacer (ITS). *Front. Microbiol.* **10** , 983 (2019).
20. Rhoads, A. & Au, K. F. PacBio sequencing and its applications. *Genomics. Proteomics Bioinformatics* **13** , 278–289 (2015).
21. Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17** , 333 (2016).
22. ter Steege, H. T. *et al.* A spatial model of tree alpha-diversity and tree density for the Amazon. *Biodivers. Conserv.* **12** , 2255–2277 (2003).
23. Hoorn, C. *et al.* Amazonia Through Time : Andean. *Science (80-.)*. **330** , 927–931 (2010).
24. Zizka, A., Steege, H. ter, Pessoa, M. do C. R. & Antonelli, A. Finding needles in the haystack: where to look for rare species in the American tropics. *Ecography (Cop.)*. **41** , 321–330 (2018).
25. Ritter, C. D. *et al.* The pitfalls of biodiversity proxies: Differences in richness patterns of birds, trees and understudied diversity across Amazonia. *Sci. Rep.* **9** , 1–13 (2019).
26. Myster, R. W. The physical structure of forests in the Amazon Basin: a review. *Bot. Rev.* **82** , 407–427 (2016).
27. Vogel, T. M. *et al.* TerraGenome: a consortium for the sequencing of a soil metagenome. (2009).
28. Laurance, S. G. W. *et al.* Influence of soils and topography on Amazonian tree diversity: a landscape-scale study. *J. Veg. Sci.* **21** , 96–106 (2010).
29. Higgins, M. A. *et al.* Geological control of floristic composition in Amazonian forests. *J. Biogeogr.* **38** , 2136–2149 (2011).
30. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41** , D590–D596 (2012).
31. Benson, D. A. *et al.* GenBank. *Nucleic Acids Res.* **46** , D41–D47 (2018).

32. Tedersoo, L. & Lindahl, B. Fungal identification biases in microbiome projects. *Environ. Microbiol. Rep.* **8** , 774–779 (2016).
33. Anslan, S., Bahram, M., Hiiesalu, I. & Tedersoo, L. PipeCraft: Flexible open-source toolkit for bioinformatics analysis of custom high-throughput amplicon sequencing data. *Mol. Ecol. Resour.* **17** , e234–e240 (2017).
34. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahe, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4** , e2584 (2016).
35. Schloss, P. D. *et al.* Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ. Microbiol.* **75** , 7537 LP – 7541 (2009).
36. Bengtsson-Palme, J. *et al.* Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods Ecol. Evol.* **4** , 914–919 (2013).
37. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10** , 996 (2013).
38. Froslev, T. G. *et al.* Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nat. Commun.* **8** , 1188 (2017).
39. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10** , 421 (2009).
40. Abarenkov, K. *et al.* The UNITE database for molecular identification of fungi – recent updates and future perspectives. *New Phytol.* **186** , 281–285 (2010).
41. Nilsson, R. H. *et al.* The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* **47** , D259–D264 (2018).
42. Cochrane, G., Karsch-Mizrachi, I., Takagi, T. & Sequence Database Collaboration, I. N. The international nucleotide sequence database collaboration. *Nucleic Acids Res.* **44** , D48–D50 (2016).
43. R Core Team. The R development core team. *R: A Language and Environment for Statistical Computing* **1** , (2003).
44. Jost, L. Entropy and diversity. *Oikos* **113** , 363–375 (2006).
45. McMurdie, P. J. & Holmes, S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput. Biol.* **10** , e1003531 (2014).
46. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15** , 550 (2014).
47. Oksanen, J. *et al.* Vegan: community ecology package. R package version 1.17-4. <http://cran.r-project.org>. *Acesso em* **23** , 2010 (2010).
48. Nielsen, U. N., Ayres, E., Wall, D. H. & Bardgett, R. D. Soil biodiversity and carbon cycling: a review and synthesis of studies examining diversity–function relationships. *Eur. J. Soil Sci.* **62** , 105–116 (2011).
49. Lauber, C. L., Hamady, M., Knight, R. & Fierer, N. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.* **75** , 5111–5120 (2009).
50. Rue, H. *et al.* INLA: functions which allow to perform a full Bayesian analysis of structured additive models using Integrated Nested Laplace Approximation. *R Packag. version 0.0* (2009).
51. Goslee, S. C. & Urban, D. L. The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.* **22** , 1–19 (2007).

52. De Caceres, M., Jansen, F. & De Caceres, M. M. Package ‘indicpecies’. *Relatsh. between species groups sites. R Packag. version 1* , (2016).
53. Oliveros, J. C. VENNY. An interactive tool for comparing lists with Venn Diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html> (2007).
54. Wickham, H. tidyverse: Easily Install and Load “Tidyverse” Packages (Version R package version 1.1.1). (2017).
55. Wickham, H. *ggplot2: elegant graphics for data analysis* . (Springer, 2016).
56. Braga-Neto, R., Luizao, R. C. C., Magnusson, W. E., Zuquim, G. & de Castilho, C. V. Leaf litter fungi in a Central Amazonian forest: the influence of rainfall, soil and topography on the distribution of fruiting bodies. *Biodivers. Conserv.* **17** , 2701–2712 (2008).
57. Barr, D. J. S. Chytridiomycota. in *Systematics and Evolution*93–112 (Springer, 2001).
58. Kosa, G. *et al.* High-throughput screening of Mucoromycota fungi for production of low-and high-value lipids. *Biotechnol. Biofuels* **11** , 66 (2018).
59. Swift, M. J. Basidiomycetes as components of forest ecosystems. *Decomposer basidiomycetes their Biol. Ecol.* (1982).
60. Liu, J. *et al.* Soil carbon content drives the biogeographical distribution of fungal communities in the black soil zone of northeast China. *Soil Biol. Biochem.* **83** , 29–39 (2015).
61. Wang, J. T. *et al.* Soil pH determines the alpha diversity but not beta diversity of soil fungal community along altitude in a typical Tibetan forest ecosystem. *J. Soils Sediments* **15** , 1224–1232 (2015).
62. Rousk, J. *et al.* Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J.* **4** , 1340 (2010).
63. Glassman, S. I., Wang, I. J. & Bruns, T. D. Environmental filtering by pH and soil nutrients drives community assembly in fungi at fine spatial scales. *Mol. Ecol.* **26** , 6960–6973 (2017).
64. Partel, M., Bennett, J. A. & Zobel, M. Macroecology of biodiversity: disentangling local and regional effects. *New Phytol.* **211** , 404–410 (2016).
65. Borges, S. H. *et al.* Bird communities in Amazonian white-sand vegetation patches: effects of landscape configuration and biogeographic context. *Biotropica* **48** , 121–131 (2016).
66. ter Steege, H. & Hammond, D. S. Character convergence, diversity, and disturbance in tropical rain forest in Guyana. *Ecology***82** , 3197–3212 (2001).
67. Haugaasen, T. & Peres, C. A. Floristic, edaphic and structural characteristics of flooded and unflooded forests in the lower Rio Purus region of central Amazonia, Brazil. *Acta Amaz.* **36** , 25–35 (2006).
68. Assis, R. L. *et al.* Patterns of tree diversity and composition in Amazonian floodplain paleo-varzea forest. *J. Veg. Sci.* **26** , 312–322 (2015).
69. Damasco, G., Vicentini, A., Castilho, C. V, Pimentel, T. P. & Nascimento, H. E. M. Disentangling the role of edaphic variability, flooding regime and topography of A mazonian white-sand vegetation. *J. Veg. Sci.* **24** , 384–394 (2013).
70. Adeney, J. M., Christensen, N. L., Vicentini, A. & Cohn-Haft, M. White-sand ecosystems in Amazonia. *Biotropica* **48** , 7–23 (2016).
71. Singer, R. & Araujo, I. de J. da S. Litter decomposition and ectomycorrhiza in Amazonian forests. 1. A comparison of litter decomposing and ectomycorrhizal basidiomycetes in latosol-terra-firme rain forest and white podzol campinarana. *Acta Amaz* 25–42 (1979).

72. Singer, R. & Aguiar, I. A. Litter decomposing and ectomycorrhizal Basidiomycetes in an igapo forest. *Plant Syst. Evol.* **153** , 107–117 (1986).
73. Singer, R., Araujo, I. & Ivory, M. H. The ectotrophically mycorrhizal fungi of the neotropical lowlands, especially central Amazonia. (Litter decomposition and ectomycorrhiza in Amazonian forests 2.). *Beihefte zur Nov. hedwigia* (1983).
74. Roy, M. *et al.* Diversity and distribution of ectomycorrhizal fungi from Amazonian lowland white-sand forests in Brazil and French Guiana. *Biotropica* **48** , 90–100 (2016).
75. Vasco-Palacios, A. M., Hernandez, J., Penuela-Mora, M. C., Franco-Molano, A. E. & Boekhout, T. Ectomycorrhizal fungi diversity in a white sand forest in western Amazonia. *Fungal Ecol.* **31** , 9–18 (2018).
76. Tedersoo, L. & Nara, K. General latitudinal gradient of biodiversity is reversed in ectomycorrhizal fungi. *New Phytol.* **185** , 351–354 (2010).
77. Tedersoo, L. *et al.* Towards global patterns in the diversity and community structure of ectomycorrhizal fungi. *Mol. Ecol.* **21** , 4160–4170 (2012).
78. Janzen, D. H. Tropical blackwater rivers, animals, and mast fruiting by the Diptercarpaeae. *Biotropica* 69–103 (1974).
79. Looney, B. P., Ryberg, M., Hampe, F., Sanchez-Garcia, M. & Matheny, P. B. Into and out of the tropics: global diversification patterns in a hyperdiverse clade of ectomycorrhizal fungi. *Mol. Ecol.* **25** , 630–647 (2016).
80. Huang, D., Meier, R., Todd, P. A. & Chou, L. M. Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. *J. Mol. Evol.* **66** , 167–174 (2008).
81. Quail, M. A. *et al.* A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* **13** , 341 (2012).
82. Prodan, A. *et al.* Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLoS One* **15** , e0227434 (2020).
83. Pereira, E. J. de A. L., Ferreira, P. J. S., de Santana Ribeiro, L. C., Carvalho, T. S. & de Barros Pereira, H. B. Policy in Brazil (2016–2019) threaten conservation of the Amazon rainforest. *Environ. Sci. Policy* **100** , 8–12 (2019).
84. Dinerstein, E. *et al.* An ecoregion-based approach to protecting half the terrestrial realm. *Bioscience* **67** , 534–545 (2017).
85. Team, Q. D. QGIS geographic information system. *Open Source Geospatial Found. Proj. Versao* **2** , (2015).

TABLES:

Table 1. Soil effects on OTU Shannon diversity by marker. The table shows the coefficients of each predictor in four Bayesian general multivariate regression models using stochastic partial differential equations (SPDE) that explicitly consider spatial correlation, modelling OTU diversity dependent on soil properties for Amazonian fungi in litter and soil. Since the organic carbon content and pH are considered important variables for soil biota, we use them as independent variables. Bold indicates important predictor variables (credible intervals not crossing zero). The importance of soil properties differed between markers and were significant only for the soil diversity. Carbon content was important for 18S and ITS soil, and chemical PC1 was important for COI and ITS.

Marker	Soil layer	Variable	Mean	SD	0.025	0.5	0.975
18S	Litter	Intercept	3.871	11.000	-19.413	4.034	25.957
		pH	0.067	0.061	-0.054	0.067	0.187
		Carbon	-0.057	0.077	-0.209	-0.056	0.089
		Chemical	0.109	0.090	-0.069	0.109	0.278
		Physical	0.017	0.043	-0.070	0.018	0.100
	Soil	Intercept	3.393	0.703	1.804	3.456	4.653
		pH	-0.019	0.045	-0.108	-0.019	0.070
		Carbon	-0.287	0.048	-0.384	-0.287	-0.194
		Chemical	0.049	0.034	-0.011	0.048	0.122
		Physical	0.029	0.031	-0.033	0.030	0.089
COI	Litter	Intercept	3.871	11.000	-19.413	4.034	25.957
		pH	0.067	0.061	-0.054	0.067	0.187
		Carbon	-0.057	0.077	-0.209	-0.056	0.089
		Chemical	0.109	0.090	-0.069	0.109	0.278
		Physical	0.017	0.043	-0.070	0.018	0.100
	Soil	Intercept	-1.670	12.588	-26.767	-1.834	24.665
		pH	0.085	0.055	-0.022	0.085	0.192
		Carbon	0.109	0.074	-0.037	0.109	0.253
		Chemical	0.620	0.082	0.460	0.619	0.782
		Physical	-0.019	0.035	-0.087	-0.019	0.050
ITS	Litter	Intercept	3.871	11.000	-19.413	4.034	25.957
		pH	0.067	0.061	-0.054	0.067	0.187
		Carbon	-0.057	0.077	-0.209	-0.056	0.089
		Chemical	0.109	0.090	-0.069	0.109	0.278
		Physical	0.017	0.043	-0.070	0.018	0.100
	Soil	Intercept	-1.400	10.548	-22.631	-1.470	20.545
		pH	-0.114	0.058	-0.229	-0.114	0.000
		Carbon	-0.389	0.085	-0.557	-0.388	-0.224
		Chemical	0.319	0.081	0.161	0.319	0.480
		Physical	-0.046	0.037	-0.119	-0.046	0.027

Table 2. Association between environmental distance and community turnover. The multiple regressions were based on the geographical distance, Euclidean distance matrices of soil properties, and community Jaccard dissimilarity index values. Bold indicates significant results. Community dissimilarity is significantly associated with geographical distance (Geo.Distance) for Amazonian fungal communities in soil and litter. All community turnovers were significant using 10,000 permutations ($p < 0.05$) with the following R^2 : 18S litter = 0.18 ($F = 31.6$) and soil = 0.18 ($F = 30.1$), COI litter = 0.26 ($F = 50.2$) and soil = 0.28 ($F = 54.5$), ITS litter = 0.12 ($F = 18.8$) and soil = 0.18 ($F = 30.1$).

Marker	Predictor	Litter		Soil	
		Coefficients	p value	Coefficients	p value
18S	Intercept	94.615	1.000	77.103	1.000
	Geo.Distance	0.144	0.003	0.084	0.050
	pH	0.193	0.002	0.143	0.026
	Carbon	0.110	0.096	0.286	0.001
	Chemical	0.168	0.015	0.109	0.162
	Physical	0.115	0.035	0.160	0.017
COI	Intercept	18.726	1.000	-1.402	1.000
	Geo.Distance	0.114	0.007	0.192	0.000

		Litter	Litter	Soil	Soil
ITS	pH	0.175	0.008	0.130	0.030
	Carbon	0.267	0.001	0.299	0.000
	Chemical	0.177	0.023	0.137	0.069
	Physical	0.215	0.002	0.246	0.000
	Intercept	157.504	1.000	110.212	1.000
	Geo.Dist	0.116	0.015	0.094	0.033
	pH	0.229	0.006	0.180	0.010
	Carbon	0.111	0.223	0.362	0.000
	Chemical	-0.115	0.227	0.046	0.589
	Physical	0.212	0.006	0.005	0.945

Table 3. Mean number of OTUs and number of indicator OTUs of Amazonian fungi by markers in each locality, habitat, and soil layer. Localities are ordered west to east: BC = Benjamin Constant, JAU = Jaú, CUI = Cuieras, and CXN = Caxiuanã. Habitats are ordered by plant and vertebrate diversity gradient: TF = Terra-firme, VZ = Várzea, IG = Igapó, and CAM = Campina. The highest number in each group is given in bold. Although the richest locality and soil layer varies depending on marker, for habitats campinas are consistently the richest for all markers.

		18S	18S	COI	COI	ITS	ITS
		Mean	Indicator	Mean	Indicator	Mean	Indicator
Locality	BC	436	90	107	75	165	174
	JAU	369	73	176	98	111	43
	CUI	338	58	181	173	142	189
	CXN	386	52	222	153	148	58
Habitat	TF	376	36	179	108	139	58
	VZ	399	101	145	184	127	118
	IG	370	61	133	79	144	73
	CAM	404	173	252	358	156	144
Soil layer	Litter	375	42	209	169	176	98
	Soil	393	106	142	29	109	12







