DNA-metabarcoding uncovers the diversity of soil-inhabiting fungi in the tropical island of Puerto Rico

Hector Urbina [1,*], Douglas G. Scofield [1,2], Matias Cafaro [3] and Anna Rosling [1]

[1] Department of Evolutionary Biology, Uppsala University, Norbyvägen 18D, Uppsala 752 36, Sweden

[2] Uppsala Multidisciplinary Center for Advanced Computational Science, Department of Information Technology, Uppsala University, Box 137, Uppsala 751 05, Sweden

[3] Department of Biology, University of Puerto Rico, Mayagüez 00681-9000, Puerto Rico

Key words

Communities, Dikarya, high-throughput sequencing, internal transcribed spacer, molecular operational taxonomic unit

Research article

CORRESPONDING AUTHOR

Hector Urbina, Department of Evolutionary Biology, Uppsala University, Norbyvägen 18D, Uppsala 752 36, Sweden. E-mail: hurbinay@gmail.com

Running title

Puerto Rican soil fungal diversity

Abbreviations

DNA: Deoxyribonucleic acid

EMF: Ectomycorrhizal fungi

HB: High abundance

ITS: Internal transcribed spacer 2 of the rRNA gene

LB: Low abundance

LSU: Large subunit of rRNA gene

MOTU: Molecular operational taxonomic unit

NCBI: National Center for Biotechnology Information

NF: National forest

NGS: Next generation sequencing

NMDS: Non-metric multidimensional scaling

PCR: Polymerase chain reaction

qPCR: Quantitative PCR

SF: State forest

TSC: Two step clustering

Abstract

Soil fungal communities in tropical regions remain poorly understood. In order to increase the knowledge of diversity of soil-inhabiting fungi, we extracted total DNA from top-organic soil collected in seven localities dominated by four major ecosystems in the tropical island of Puerto Rico. In order to comprehensively characterize the fungal community, we PCR-amplified the ITS2 fungal barcode using newly designed degenerated primers and varying annealing temperatures to minimize primer bias. Sequencing results, obtained using Ion Torrent technology, comprised a total of 566,613 sequences after quality filtering. These sequences were clustered into 4,140 molecular operational taxonomic units (MOTUs) after removing low frequency sequences and rarefaction to account for differences in read depth between samples. Our results demonstrate that soil fungal communities in Puerto Rico are structured by ecosystem. Ascomycota, followed by Basidiomycota, dominates the diversity of fungi in soil. Amongst Ascomycota, the recently described soil-inhabiting class Archaeorhizomycetes was present in all the localities and taxa in this class were among the most commonly observed MOTUs. The Basidiomycota community was dominated by soil decomposers and ectomycorrhizal fungi with a distribution strongly affected by local variation to a greater degree than Ascomycota.

Introduction

Recent studies using next generation sequencing (NGS) technologies have increased our understanding of the diversity of soil-inhabiting fungi enormously. Most surveys were carried out in the Northern Hemisphere, more specifically in Europe and North America (e.g., Clemmensen et al. 2013; Lentendu et al. 2014; Menkis et al. 2015), however a few investigations targeting soil-inhabiting fungi in tropical regions (e.g., Kemler et al. 2013; Tedersoo et al. 2014). Initial surveys indicate that tropical biomes contain the greatest species richness of soil fungi (McGuire et al. 2013), though across different biomes, soil fungal diversity appears to be similarly structured by abiotic and biotic factors (McGuire et al. 2013; Tedersoo et al. 2014). We do not yet have more extensive sampling from a single tropical region, and more efforts are needed to extend the characterization of tropical soil fungal communities.

The Caribbean tropical island of Puerto Rico exhibits a rich diversity of flora and fauna in just 8,900 km^2 (Gannon et al. 2005; Liogier and Martorell 2000). This landmass together with Cuba and Hispaniola form the Greater Antilles of the West Indies, which are all fragments of old continental crust (Ricklefs and Bermingham 2008). Puerto Rico has a complex geographical history due to several periods of separation and rejoining with the other Greater Antilles (Ricklefs and Bermingham 2008). Six major ecosystems have been identified: littoral zone forest, semi-deciduous subtropical dry forest, tropical- and subtropical-moist forest, and subtropical rain forest (Helmer et al. 2002).

Previous studies in Puerto Rico have aimed to characterize fungi adapted to hypersaline environments using classic molecular and culture-based approaches (Burgos-Caraballo et al. 2014; Cantrell and Duval-Perez 2012; Cantrell et al. 2007). More recently, the diversity of soil fungi has been characterized using NGS techniques at three localities in El Yunque National Forest (NF) (Tedersoo et al. 2014). Consequently, soil-inhabiting fungal communities remain largely uncharacterized throughout this island. The aim of this work is to describe the diversity of soil-inhabiting fungi across the major ecosystems in Puerto Rico.

Materials and Methods

Soil samples and total DNA extraction

Three plots were sampled at each of seven localities representing four different ecosystems in Puerto Rico (Fig. 1, Table 1). The top-organic layer of plant debris and rocks was removed in an approximately 1 m^2 grid and a table spoon of soil was scooped from each corner and the center of an internal 0.50 m^2 grid. Soils were pooled in a sterile plastic bag and manually homogenized for 1 min. Two sub-samples of approximately 0.5 g were added into separate 2.0 mL microtube containing 750 µL of lysis buffer (Xpedition^{TM} Soil/Fecal DNA miniprep, Zymo Research Corporation, Irvine, California, USA). Followed by cell disruption for 30 s using a TerraLyser^{TM} (Zymo Research Corporation). Samples were stored at room temperature and later at 4 °C until DNA extraction was carried out in the laboratory within one month of sampling, following the manufacturer's protocol.

DNA concentration and integrity was verified in 0.8 % agarose gel electrophoresis on 0.5 % Tris Acetate-EDTA buffer (Sigma-Aldrich, St. Louis, Missouri, USA) stained with 1 × GelRed™ (Biotium Inc., Hayward, California, USA). DNA from a pure culture of *Neurospora crassa* was included as a positive control.

Polymerase Chain Reaction (PCR) and Ion Torrent library preparation

A fragment of the 5.8S, the Internal transcribed spacer 2 (ITS2) and a fragment of the Large Subunit (LSU) of the rRNA genes was amplified using primers gITS7 *forward* (Ihrmark et al. 2012) and modified ITS4m *reverse* (5'-TCCTC[**C/G**][**G/C**]CTTATTGATATGC-3'), with both primers containing adequate barcode sequences for single-ended amplification (Table 1). The ITS locus is broadly accepted as a taxonomic barcode for *Fungi* (Schoch et al. 2012) and among ITS1 and ITS2 there are no significant differences in their power of discriminating species between fungal groups (Blaalid et al. 2013). Modifications on the reverse primer ITS4 (White et al. 1990) were included to reduce its known bias against the soil-inhabiting fungal class Archaeorhizomycetes (Schadt and Rosling 2015).

The PCR mixes were comprised of 10 - 20 ng of soil DNA, 1 × SSoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad Laboratories, Hercules, California, USA), and 0.8 nM of each primer in a final volume of 20 µL. PCR amplifications were carried out in a CFR96 Touch™ Real/Time PCR Detection system (Bio-Rad Laboratories) following the protocol, 10 min pre-

denaturation at 95 °C, 1 min DNA denaturation at 95 °C, 45 s at three independent annealing temperatures (50, 54 and 58 °C) to reduce primer bias Schmidt et al. (2013), 50 s of extension at 72 °C and 3 min final extension at 72 °C. We used a quantitative PCR (qPRC) that allowed us to adjust the number of cycles for each plate between 23 − 27, in order to ensure that the reactions were maintained within linear amplification. To reduce the chance of altering the relative abundance of fungi by biasing against long reads, we used tag primers directly thereby avoiding to an extra nested-PCR run. All reactions were carried out in duplicates, and all six runs were combined before purification using the ZR-96 DNA Clean & Concentrator™-5 (Zymo Research Corporation). DNA concentration was quantified on duplicates using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies Corporation, Carlsbad, California, USA) on a TECAN F500 microplate reader and the DNA integrity was checked by electrophoresis in 2 % agarose gel 0.5 × TAE buffer. A sequencing library was prepared by pooling 35 ng DNA from each sample, loaded onto a 318 chip for PGM Ion Torrent sequencing technology (Life Technologies Corporation) and sequenced in the facilities at Uppsala Genome Center (Uppsala University, Sweden).

Assembly and taxonomy of molecular operational taxonomic units (MOTUs)

We obtained a demultiplexed dataset comprising 1,440,213 sequences, with 1,359,657 sequences corresponding to samples and 80,556 to the positive control (Torrent Suite v 4.0.4) (Table 1, Supplementary Fig.S1A). The software

mothur v1.33.3 (Schloss et al. 2009) was used for sequence trimming of the FastQ files ($Q \geq 25$ in sequence average, minimum length 150 bp) (Table 1, Supplementary Fig. S1B). To improve the accuracy of clustering and assigned taxonomy of the MOTUs we used only the variable ITS2 locus, by trimming the fragments of 5.8S and LSU loci using the software ITSx v1.0.9 (Bengtsson-Palme et al. 2013). After this step, all sequences shorter than 80 bp and non-fungal sequences were removed; consequently our dataset was reduced to 648,535 sequences of fungal ITS2 locus exclusively (Supplementary Fig. 1C).

The ITS2 fungal sequence dataset was *de novo* clustered using the software TSC (Jiang et al. 2012) under the following parameters: data type 454 (Ion Torrent and 454 LifeSciences sequencing technology possesses similar sequencing errors (Yang et al. 2013)), single linkage clustering algorithm (recommended for ITS locus in fungi (Lindahl et al. 2013)), sequence identity 0.97 (suitable cutoff for species delimitation in fungi (Blaalid et al. 2013)). We tested three cutoff values for high (HB) and low abundance (LB) sequences (100, 500 and 1000) and obtained 18,961, 18,293 and 18,292 MOTUs respectively for each cutoff. Because cutoffs 500 and 1000 converged to similar total number of MOTUs, we choose 500 for further analysis. Singletons and doubletons were then discarded. The final ITS2 dataset contained 608,300 sequences and 4,378 MOTUs, representing 97.3 % of sequences and 23.9% of MOTUs from the ITS2 dataset, a reduction that allowed for robust analyses (Supplementary Fig. 1D). Predictably, the positive control showed higher retention than the total soil DNA samples, and we found no major differences among samples (Table 1). Ion

Torrent sequencing technology can prematurely truncate sequences (Salipante et al. (2014), thus the longest sequences representing each MOTUs were selected using the script *pick_rep_set.py* in QIIME v1.8.0 (Caporaso et al. 2010). The initial taxonomic assignment of MOTUs was performed using BLASTn 2.2.29+ (Madden 2002) with parameters E-value = $1e^{-10}$ and word size = 7, against the reference database 2015-03-02 UNITE + INDS (Koljalg et al. 2013) with only trimmed ITS2 region as described above. Unassigned MOTUs were blasted against the nucleotide database nt at NCBI and sequences of the ITS2 hits were extracted along with their respectively taxonomy using custom scripts deposited at https://github.com/douglasgscofield/add-to-Qiime-DB. For any sequences with taxonomic levels not matching those in the 2015-03-02 UNITE + INDS database, we corrected their taxonomy using Index Fungorum (http://www.indexfungorum.org). We also added Archaeorhizomycetes representative sequences (Menkis et al. 2014).

The final taxonomic identification of each representative sequence was corrected based on 97, 90, 85, 80, 75 and 70 % of sequence identity for assigning MOTUs with names of a species, genus, family, order, class, or phylum, respectively, as proposed by Tedersoo et al. (2014).

The positive control using DNA of *N. crassa* was used to verify the entire protocol from sampling and tag primer arrangement to clustering and taxonomic assignment. The majority of sequences (95%) in this sample were clustered into a unique MOTU that was identified as *N. crassa*. The error of 5 % was distributed across 53 MOTUs. Eleven of the MOTUs were unique of the positive control (288

sequences) the rest represents MOTUs also identified in the main dataset and showed variable sequence abundance from 1 – 243 sequences. We concluded that this error originates predominantly from the demultiplexing step and it did not affect the abundance of any particular MOTU significantly (less than 0.1%). The positive control was eliminated from further analysis. The script *compute_core_microbiome.py* in Qiime was used to account for the difference in sequence depth among samples by the rarefaction method using the lowest number of sequences (11,455) found in one of the samples as recommended by (McMurdie and Holmes 2014). This resulted in a reduction to 4,140 MOTUs. Sequence representatives for each MOTU that satisfied GenBank requirements are available under the accession numbers (KT241043 - KT245134). Sequence data and abundance information per soil sample and locality are available in the Supplementary online File S1.

Statistical analyses

Sequences abundance plots were generated using the library *phyloseq* (McMurdie and Holmes 2013) in R v3.0.2 (http://www.r-project.org). In order to address differences in MOTUs composition and diversity, rarefaction curves, Chao-1 and Shannon diversity indexes and the non-metric multidimensional scaling (NMDS) based on Bray Curtis distance were computed in the R package *ampvis* (http://madsalbertsen.github.io/ampvis/). To address the statistically significance of the variables studied here the multivariate analysis of variance *Adonis* implemented in Qiime was used. Double clustering analysis based on the

100 most abundant MOTUs was carried out using the R script *heatmap.2* (http://www.inside-r.org/packages/cran/gplots/docs/heatmap.2). The g-Test of independence was used to calculate the significance of association of MOTUs with locality and ecosystem and corrected by Bonferroni and False Discovery Ration test with 1000 permutations, all as implemented in the Qiime script *group_significance.py*.

Results and Discussion

Fungal identification based on MOTUs

We identified a total of 4,140 MOTUs at different taxonomic levels after rarefaction: 559 to species (13.5 %), 965 to genus (23.3 %), 919 to family (22.2 %), 785 to order (19.0 %), 354 to class (8.6 %), 548 to phylum (13.2 %) and only 10 MOTUs (0.2 %) remained unclassified but assigned to kingdom fungi. The lack of ITS2 reference sequences of fungi from tropical regions as well as fungal endemism in Puerto Rico are likely reasons for the identification of a high number of MOTUs resolved only to high taxonomic levels such as phylum or class. This phenomenon is common in many recent studies, e.g., in Menkis et al. (2015) 13 of the 30 most common MOTUs were identified only to phylum, and in Toju et al. (2014) 10 of the 25 most common MOTUs were identified only to phylum. There is clearly a great need for more fungal inventories from the tropics, including Puerto Rico, that contain both specimen vouchers – i.e. fruiting bodies, axenic cultures – and their molecular characterization.

Among localities the number of common MOTUs and their percentages are shown in the Table 2. On average, the number of MOTUs shared between pairs of localities was 20.8 %. Specifically, we reported 1,499 MOTUs from El Yunque (Cubuy and El Cacique) while independently Tedersoo et al. (2014) found 1,652 fungal MOTUs in the same forest locality but at different sampling sites (Supplementary Fig. S2) with 264 MOTUs (17.8 %) in common between both studies based on a blast sequence similarity search using 97% similarity against our MOTU dataset. These results support our observation that the large majority of MOTUs are unique to each sampling locality. Interestingly Tedersoo et al. (2014) and our study are independent and employing differing sampling and DNA extraction methodologies yet both captured similar MOTU richness at the same locality, despite high local variation and relatively low MOTU overlap. Field surveys should take into account the degree of spatial separation between sites (McGuire et al. 2013) to increase the amount of expected overlapping species between individual samples in order to capture a more complete set of fungal species.

In terms of taxonomic abundance, Dikarya dominates the diversity of soil-inhabiting fungi in Puerto Rico (Fig. 2A, Supplementary online file S1). MOTUs classified in Ascomycota were the most abundant (2,967 MOTUs; 79.6 % sequences) followed by Basidiomycota (1,022 MOTUs; 17.5 % sequences). Much less abundant fungi included Glomeromycota (206 MOTUs; 1.6 % sequences;), Chytridiomycota (35 MOTUs; 0.2 % sequences), Zygomycota (68 MOTUs; 0.2 % sequences), and Cryptomycota (67 MOTUs; 0.3 % sequences).

In accordance with other studies, we observe Dikarya to be the most abundant group as they dominate the diversity of soil fungi in most ecosystems worldwide (Tedersoo et al. 2014; Toju et al. 2014). Species in this group include the principal decomposers of organic matter and many taxa establish symbiotic relationships with plant roots (Smith and Read 2008).

Soil-inhabiting Ascomycota in Puerto Rico

Among Ascomycota the classes with major sequence abundance were Archaeorhizomycetes (20.9 %), Sordariomycetes (13.9 %), Eurotiomycetes (11.4 %), Dothideomycetes (10.6 %) and Leotiomycetes (10.6 %) (Figs. 2A & 3). Among ascomycetes, 26 MOTUs were found across all localities: *Archaeorhizomyces* (4 MOTUs), *Pestalotiopsis* (2 MOTUs), *Aspergillus, Bionectria, Chaetomium, Cladosporium, Cylindrocladium, Haematonectria, Lasiodiplodia, Neurospora, Penicillium, Phialocephala* and *Talaromyces* (1 MOTU each) (Supplementary online File S1).

Focusing on the most abundant class Archaeorhizomycetes, a total number of 50,379 sequences were assigned to the class, these grouped into 190 MOTUs (4.6% of the total MOTUs) (Supplementary online File S1). Only eight of these MOTUs were identified to the species level using environmental sequence previously identified as belonging to the class (Menkis et al. 2014). Probably because available sequence were obtained predominantly from studies performed in non-tropical regions. Our modifications to the traditional ITS4 primer (White et al. 1990) to reduce its bias against Archaeorhizomycetes (Schadt and

Rosling 2015) were clearly successful; see Material and Methods for further details. Our results are in agreement with previous observations that the diversity of Archaeorhizomycetes is highest in tropical regions (Tedersoo et al. 2014), however our observed abundances are much higher than in that study resulting from our use of less biased primers.

MOTUs classified in Archaeorhizomycetes were found in all the localities (Fig. 2B). Archaeorhizomycetes were most abundant in the wet forests El Yunque (Cubuy) and Maricao, followed by moist forest Guajataca, then dry forests Guanica and Boqueron (Figs. 2B & 3). Only a few reads were classified as Archaeorhizomycetes in samples from the littoral locality Isabela dominated by introduced *Casuarina* trees. In contrast to other wet localities few reads were classified as Archaeorhizomycetes from the wet forest site El Yunque (El Cacique). We do not have an explanation for the low numbers of Archaeorhizomycetes from the El Cacique site, though this could be due to local soil conditions, or site history as the degree of forest preservation is higher in Cubuy than El Cacique. Current sampling could not resolve whether locality or ecosystem had a greater effect on the distribution of Archaeorhizomycetes in the island; we found that 84 MOTUs differed in abundances among localities (Bonferroni $P < 0.01$; Supplementary Table S1) while 58 MOTUs differed among ecosystems (Bonferroni $P < 0.01$; data not shown). This high variation in Archaeorhizomycetes abundance among localities is in accordance with earlier observations of local variation of Archaeorhizomycete (Porter et al. 2008; Rosling et al. 2013). Species abundances have been suggested to be affected by biotic

and abiotic factors including type of vegetation, soil horizon, season and pH (Rosling et al. 2013).

Another abundant Ascomycota MOTU was MOTUHB 5, identified as *Phialocephala* (Vibrisseaceae). This MOTU occurred across all localities (Figs. 2B & 3; Supplementary online File S1), but was particularly abundant in littoral Isabela, which is dominated by introduced *Casuarina* trees. Species belonging to this genus are soil- and root- inhabiting fungi commonly found in alpine and boreal ecosystems in association with pine roots (Jacobs et al. 2003). High abundance of *Phialocephala* MOTUHB 5 in this locality could be explained as a result of association with mycorrhizal activity in *Casuarina* roots (Wang and Qiu 2006). Other studies have also reported *Phialocephala* as a common root-associated fungi (Bougoure and Cairney 2005) and important decomposers of organic matter in soil in tropical regions (DeAngelis et al. 2013).

Soil-inhabiting Basidiomycota in Puerto Rico

Among Basidiomycota, the class Agaricomycetes was the most diverse with (669 MOTUs; 15.3 % sequences) followed by unclassified Basidiomycota (207 MOTUs, 1.3 %, Fig. 2C). The most diverse families were Agaricaceae, Clavariaceae, Mycenaceae, Thelephoraceae and Psathyrellaceae, most members of which are saprophytic (Supplementary online File S1). Ectomycorrhizal fungi (EMF) classified in the Boletaceae, Inocybaceae, Pluteaceae and Sebacinaceae (108 MOTUs) (Supplementary online File S1) were also detected in Puerto Rico (Fig. 2B). Guajataca was the locality in which

basidiomycetes were most abundant (134 MOTUs; 8,103 sequences) followed by Maricao (256 MOTUs; 7,700 sequences). The most abundant MOTU was classified as unidentified Pluteaceae (MOTUHB 31; 3,326 sequences). It was identified at two localities, most abundantly in Guanica (3,044 sequences) with markedly fewer sequences in Guajataca (255 sequences). This pattern of strongly site-specific abundances was observed among the most abundant basidiomycete MOTUs: 15, 1060, 120, 21, 13104, 33 and others (Fig. 3, Supplementary online File S1). The majority of basidiomycete MOTUs had fewer than 20 sequences in total (746 MOTUs, 73.1 %) and no Basidiomycota MOTU was common across all localities nor were any found at all localities, in contrast to Ascomycota (see above).

We identified basidiomycete taxa known to contain EMF fungi, e.g., Inocybaceae, Entolomataceae, Sebacinaceae (Supplementary online File S1). EMF basidiomycetes have been observed in the rhizosphere of native plants on Tenerife in the Canary Islands (Zachow et al. 2009), suggesting that the EMF symbiotic relationship is common in tropical islands as well. The strong local variation we observed in the basidiomycete communities of Puerto Rico has also been observed in other tropical as well as boreal ecosystems (McGuire et al. 2013; Tedersoo et al. 2012). Such strong local variation could be an effect of the presence of clusters of EMF trees that drastically change the local fungal community (McGuire et al. 2013), and also could reflect strong effects of precipitation and temperature on basidiomycete communities (Tedersoo et al. 2012).

In general our results are consistent with suggestions of lower diversity of EMF in tropical ecosystems than in northern temperate ecosystems (Tedersoo et al 2012, 2014). In our study, soil fungal communities were dominated by fungi identified belonging to Ascomycota, in contrast with observations by Tedersoo and co workers (2014) who found greater abundance of Basidiomycota relative to Ascomycota in tropical soils. Clearly more sampling is required throughout the tropics in order to resolve these general patterns of soil fungal communities in tropical soils.

Soil-inhabiting Glomeromycota in Puerto Rico

We found Glomeromycota to be most abundant in Isabela (46 MOTUs; 1,281 sequences), the littoral *Casuarina* site, and in contrast to nearly all other localities, its abundance was consistent among samples (Fig. 2B). Similar to the ascomycete *Phialocephala*, this may reflect an association with *Casuarina* roots, with the consistency of abundance due to the local ubiquity of *Casuarina*. Whether our results reflect the actual abundance of Glomeromycota at this and other sites is unclear, given the known difficulties in amplifying ITS regions from this group of fungi in general (Hart et al. 2015).

Soil-inhabiting fungal community diversity across localities and ecosystems

The MOTU accumulation curves indicate high variation in number of MOTUs per individual sample (200 – 800 MOTUs in total, Supplementary Fig. S2) and localities (900 – 1200 MOTUs in total, Fig. 5). The observed MOTU

accumulation curves continued to accumulate total alpha diversity, while both the Chao-1 richness and Shannon diversity measures reached or approached a plateau in the majority of localities (Fig. 4) as well as samples (Supplementary Fig. S2). This indicates that we have sampled the very large majority of soil fungal diversity at our localities, while some rare taxa likely remains undetected.

Based on the curves for observed MOTU and Chao-1 richness, the moist Guajataca has the highest richness of soil fungal MOTUs and El Yunque (El Cacique) the lowest (Fig. 4). The low fungal richness of El Yunque El Cacique is reflected in our results and discussion of specific taxonomic groups above, and remains a puzzling observation.

The Shannon curve, which takes into account abundance and evenness, indicates that all protected localities (national and state forests) have higher diversity of soil fungi than Isabela, which has low plant diversity and is dominated by introduced *Casuarina* trees (Fig. 4). The low abundance and diversity of soil fungi at Isabela is also reflected in the analysis of most common MOTUs (Fig. 3 & Supplementary Fig. S3). Above we have outlined some reasons for particular taxonomic representation of soil fungi at Isabela, because of the dominance of exotic plant species *Casuarina*. Whether lower diversity at this site also reflects more general effects of anthropogenic activities on soil fungal communities in Puerto Rico or other tropical areas will require specific sampling efforts.

We used non-metric multidimensional scaling (NMDS) of the 21 samples to identify how fungal community composition differed with locality and ecosystem (Fig. 5). This analysis showed that the soil fungal communities are

distinct depending on both the locality (r^2 = 0.844, P < 0.01) as well as the ecosystem (r^2 = 0.781, P < 0.01) (Fig. 5), reflecting the MOTU-specific analysis above and confirmed using clustering analysis (Fig. 3). This same pattern was also recovered analyzing the Ascomycota and Basidiomycota communities separately; (Ascomycota, locality: r^2 = 0.831, P < 0.01; ecosystem: r^2 = 0.7794, P < 0.01), Basidiomycota, locality: r^2 = 0.796, P < 0.01; ecosystem: r^2 = 0.532, P < 0.02; Supplementary Fig. S2). More detailed locality information than we gathered would be required to separate the effects of specific climatic and edaphic conditions, including any effect of covariance of soil fungal communities with soil calcium content (Tedersoo et al. 2014).

Conclusion

Soil fungal communities in Puerto Rico are organized similarly to other mature tropical and temperate soil fungal communities, with the Ascomycota dominating, followed by the Basidiomycota. In particular, we have shown that Archaeorhizomycetes is one of the most abundant classes in many localities; this was possible only after addressing known primer and amplification biases. Soil fungal community structure also varies significantly among localities and ecosystems, with just a handful (26) MOTUs present at all seven localities, all common soil decomposer ascomycetes. Basidiomycetes were mostly saprophytic and they had a more local distribution compared to Ascomycetes. MOTU accumulation analysis showed that we have identified the majority of fungal taxa present in each sample and locality. This reflects our comprehensive

methodology, which includes *in situ* DNA extraction, the use of deep sequencing output obtained from the Ion Torrent platform, and increasing the breadth of taxonomic identification in our bioinformatic pipeline.

Acknowledgements

This work was supported by the Swedish Research Council VR and the Carl Trygger Foundation for Scientific Research. We thank Stefan Bertilsson, Department of Limnology, Uppsala University, for providing access to the CFR96 Touch™ Real/Time PCR Detection system. We acknowledge the Departamento de Recursos Naturales y Ambientales of Puerto Rico for granting a collecting permit (2014-IC-082) to MJC.

Cited references

Bengtsson-Palme J, Ryberg M, Hartmann M, Branco S, Wang Z, Godhe A, de Wit PJ, Sánchez-García M, Ebersberger I, de Sousa F, Amend A, Jumpponen A, Unterseher M, Kristiansson E, Abarenkov K, Bertrand Y, Sanli K, Eriksson K, Vik U, Veldre V, Nilsson RH, 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution* **4**, 914-919.

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.

Bougoure DS, Cairney JW, 2005. Fungi associated with hair roots of *Rhododendron lochiae* (Ericaceae) in an Australian tropical cloud forest revealed by culturing and culture-independent molecular methods. *Environ Microbiol* **7**, 1743-1754.

Burgos-Caraballo S, Cantrell SA, Ramirez A, 2014. Diversity of benthic biofilms along a land use gradient in tropical headwater streams, Puerto Rico. *Microb Ecol* **68**, 47-59.

Cantrell SA, Duval-Perez L, 2012. Microbial mats: an ecological niche for fungi. *Front Microbiol* **3**, 424.

Cantrell SA, Hanlin RT, Emiliano A, 2007. *Periconia variicolor* sp. nov., a new species from Puerto Rico. *Mycologia* **99**, 482-487.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R, 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335-336.

Clemmensen KE, Bahr A, Ovaskainen O, Dahlberg A, Ekblad A, Wallander H, Stenlid J, Finlay RD, Wardle DA, Lindahl BD, 2013. Roots and associated fungi drive long-term carbon sequestration in boreal forest. *Science* **339**, 1615-1618.

DeAngelis KM, Chivian D, Fortney JL, Arkin AP, Simmons B, Hazen TC, Silver WL, 2013. Changes in microbial dynamics during long-term decomposition in tropical forests. *Soil Biology & Biochemistry* **66**, 60-68.

Gannon MR, Kurta A, Rodriguez-Duran A, Willig MR, 2005. *Bats of Puerto Rico*. Texas Tech University, Lubbock, Texas.

Hart MM, Aleklett K, Chagnon PL, Egan C, Ghignone S, Helgason T, Lekberg Y, Opik M, Pickles BJ, Waller L, 2015. Navigating the labyrinth: a guide to sequence-based, community ecology of arbuscular mycorrhizal fungi. *New Phytologist* **207**, 235-247.

Helmer EH, Ramos O, Lopez TD, Quinones M, Diaz W, 2002. Mapping the forest type and land cover of Puerto Rico, a component of the Caribbean biodiversity hotspot. *Caribbean Journal of Science* **38**, 165-183.

Ihrmark K, Bodeker IT, Cruz-Martinez K, Friberg H, Kubartova A, Schenck J, Strid Y, Stenlid J, Brandstrom-Durling M, Clemmensen KE, Lindahl BD, 2012. New primers to amplify the fungal ITS2 region--evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiol Ecol* **82**, 666-677.

Jacobs A, Coetzee MP, Wingfield BD, Jacobs K, Wingfield MJ, 2003. Phylogenetic relationships among *Phialocephala* species and other ascomycetes. *Mycologia* **95**, 637-645.

Jiang XT, Zhang H, Sheng HF, Wang Y, He Y, Zou F, Zhou HW, 2012. Two-stage clustering (TSC): a pipeline for selecting operational taxonomic units for the high-throughput sequencing of PCR amplicons. *PLoS One* **7**, e30230.

Kemler M, Garnas J, Wingfield MJ, Gryzenhout M, Pillay KA, Slippers B, 2013. Ion Torrent PGM as tool for fungal community analysis: a case study of endophytes in Eucalyptus grandis reveals high taxonomic diversity. *PLoS One* **8**, e81718.

Koljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AF, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Duenas M, Grebenc T, Griffith GW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lucking R, Martin MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Poldmaa K, Saag L, Saar I, Schussler A, Scott JA, Senes C, Smith ME, Suija A, Taylor DL, Telleria MT,

Weiss M, Larsson KH, 2013. Towards a unified paradigm for sequence-based identification of fungi. *Mol Ecol* **22**, 5271-5277.

Lentendu G, Wubet T, Chatzinotas A, Wilhelm C, Buscot F, Schlegel M, 2014. Effects of long-term differential fertilization on eukaryotic microbial communities in an arable soil: a multiple barcoding approach. *Mol Ecol* **23**, 3341-3355.

Lindahl BD, Nilsson RH, Tedersoo L, Abarenkov K, Carlsen T, Kjoller R, Koljalg U, Pennanen T, Rosendahl S, Stenlid J, Kauserud H, 2013. Fungal community analysis by high-throughput sequencing of amplified markers--a user's guide. *New Phytol* **199**, 288-299.

Liogier AH, Martorell LF, 2000. *Flora of Puerto Rico and adjacent islands : a systematic synopsis*, 2nd ed. Editorial de la Universidad de Puerto Rico, San Juan, P.R.

Madden T, 2002. The BLAST Sequence Analysis Tool, *The NCBI Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2002-*.

McGuire KL, Allison SD, Fierer N, Treseder KK, 2013. Ectomycorrhizal-dominated boreal and tropical forests have distinct fungal communities, but analogous spatial patterns across soil horizons. *PLoS One* **8**, e68278.

McMurdie PJ, Holmes S, 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217.

McMurdie PJ, Holmes S, 2014. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* **10**, e1003531.

Menkis A, Marciulynas A, Gedminas A, Lynikiene J, Povilaitiene A, 2015. High-Throughput Sequencing Reveals Drastic Changes in Fungal Communities in the Phyllosphere of Norway Spruce (*Picea abies*) Following Invasion of the Spruce Bud Scale (*Physokermes piceae*). *Microb Ecol*. Epub DOI 10.1007/s00248-015-0638-z.

Menkis A, Urbina H, James T, Rosling A, 2014. *Archaeorhizomyces borealis* sp. nov. and a sequence-based classification of related soil fungal species. *Fungal Biology*. *Fungal Biol.* **118**, 934-955.

Porter TM, Schadt CW, Rizvi L, Martin AP, Schmidt SK, Scott-Denton L, Vilgalys R, Moncalvo JM, 2008. Widespread occurrence and phylogenetic placement of a soil clone group adds a prominent new branch to the fungal tree of life. *Mol Phylogenet Evol* **46**, 635-644.

Ricklefs R, Bermingham E, 2008. The West Indies as a laboratory of biogeography and evolution. *Philos Trans R Soc Lond B Biol Sci* **363**, 2393-2413.

Rosling A, Timling I, Taylor DL, 2013. Archaeorhizomycetes: Patterns of distribution and abundance in soil, in: Horwitz BA, al. e (eds), *Genomics of soil- and plant-associated fungi*. Springer-Verlag, Berlin.

Salipante SJ, Kawashima T, Rosenthal C, Hoogestraat DR, Cummings LA, Sengupta DJ, Harkins TT, Cookson BT, Hoffman NG, 2014. Performance Comparison of Illumina and Ion Torrent next-generation dequencing platforms for 16S rRNA-based bacterial community profiling. *Appl Environ Microbiol* **80**, 7583-7591.

Schadt CW, Rosling A, 2015. Comment on “Global diversity and geography of soil fungi”. *Science* **348**, 1438.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF, 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**, 7537-7541.

Schmidt PA, Balint M, Greshake B, Bandow C, Rombke J, Schmitt I, 2013. Illumina metabarcoding of a soil fungal community. *Soil Biology & Biochemistry* **65**, 128-132.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding C, Fungal Barcoding Consortium Author L, 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci U S A* **109**, 6241-6246.

Smith SE, Read DJ, 2008. *Mycorrhizal symbiosis*, 3nd ed. Academic Press, San Diego, Calif.

Tedersoo L, Bahram M, Polme S, Koljalg U, Yorou NS, Wijesundera R, Villarreal Ruiz L, Vasco-Palacios AM, Thu PQ, *et al*., 2014. Fungal biogeography. Global diversity and geography of soil fungi. *Science* **346**, 1256688.

Tedersoo L, Bahram M, Toots M, Diedhiou AG, Henkel TW, Kjoller R, Morris MH, Nara K, Nouhra E, Peay KG, Polme S, Ryberg M, Smith ME, Koljalg U, 2012. Towards global patterns in the diversity and community structure of ectomycorrhizal fungi. *Mol Ecol* **21**, 4160-4170.

Toju H, Sato H, Tanabe AS, 2014. Diversity and spatial structure of belowground plant-fungal symbiosis in a mixed subtropical forest of ectomycorrhizal and arbuscular mycorrhizal plants. *PLoS One* **9**, e86566.

Wang B, Qiu YL, 2006. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* **16**, 299-363.

White TJ, Bruns SL, Taylor W, 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, in: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds), *PCR protocols: A guide to methods and applications*. Academic Press, Inc., New York, pp. 315-322.

Yang X, Chockalingam SP, Aluru S, 2013. A survey of error-correction methods for next-generation sequencing. *Brief Bioinform* **14**, 56-66.

Zachow C, Berg C, Muller H, Meincke R, Komon-Zelazowska M, Druzhinina IS, Kubicek CP, Berg G, 2009. Fungal diversity in the rhizosphere of endemic plant species of Tenerife (Canary Islands): relationship to vegetation zones and environmental factors. *ISME J* **3**, 79-92.

List of Figures

List of Supplementary Figures

Figure S1: Read length frequency for all the samples: A) after demultiplexing; B) after quality filtering using mothur; C) after ITSx trimming and removal of nonfungal sequences; D) final read set after removal of singletons and doubletons

Figure S2: Map of Puerto Rico showing our sampling sites together with the El Yunque sampling sites in Tedersoo et al. (2014).

Figure S3: Non-metric multidimensional scaling (NMDS) of all Ascomycota MOTUs (left) and Basidiomycota MOTUs (right) in each sample, characterized by locality and ecosystem.

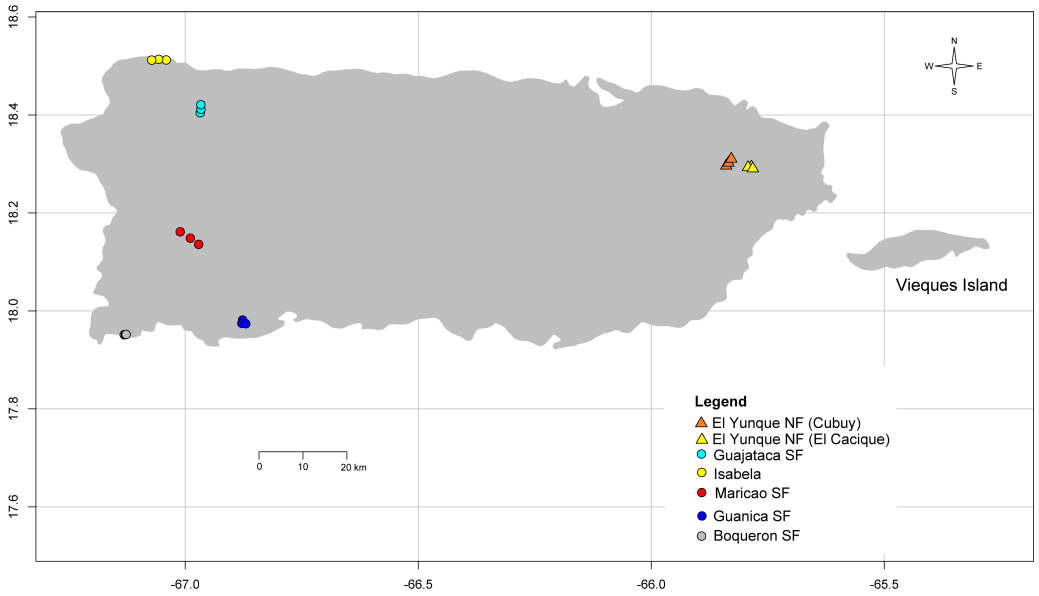
Figure S4. MOTU accumulation curve of alpha diversity by samples among observed species and the Chao-1 and Shannon diversity indices.

List of Tables

Supplementary Tables

Supplementary Table S1: g-Test results for Archaeorhizomycetes across localities. The probabilities of the *post hoc* statistical test of False Discovery Rate (FDR) and Bonferri correction are also included. For all MOTUs the $p < 0.001$ and FRD $p < 0.001$ and Bonferri $p < 0.001$.

Figure 1



1 **Figure 2**

2

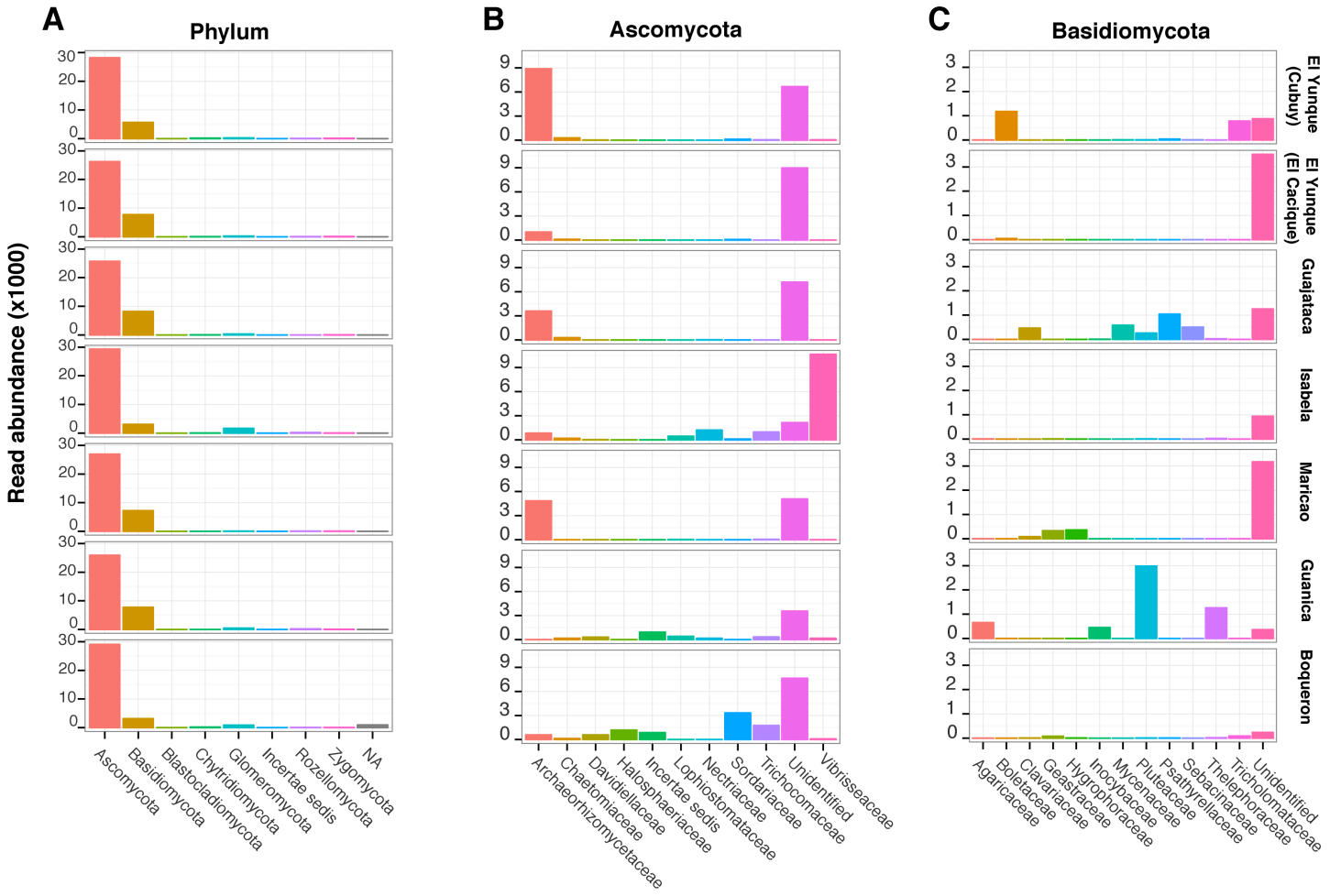

A
Phylum
Read abundance (x1000)
Ascomycota
Basidiomycota
Blastocladiomycota
Chytridiomycota
Glomeromycota
Incertae sedis
Rozellomycota
Zygomycota
NA
B
Ascomycota
Archaeorhizomycetaceae
Chaetomiaceae
Davidiellaceae
Halosphaeriaceae
Incertae sedis
Lophiostomataceae
Nectriaceae
Sordariaceae
Trichocomaceae
Unidentified
Vibrisseaceae
C
Basidiomycota
Agaricaceae
Boletaceae
Clavariaceae
Geastraceae
Hygrophoraceae
Inocybaceae
Mycenaceae
Pluteaceae
Psathyrellaceae
Sebacinaceae
Thelephoraceae
Tricholomataceae
Unidentified
El Yunque (Cubuy)
El Yunque (El Cacique)
Guajataca
Isabela
Maricao
Guanica
Boqueron

Figure 3

Figure 4

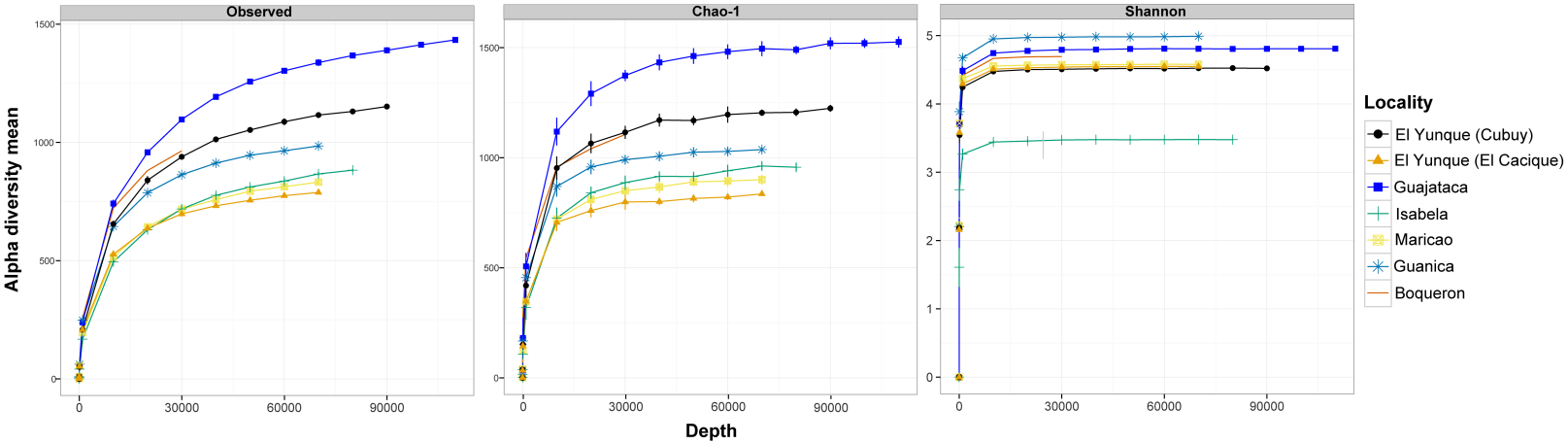
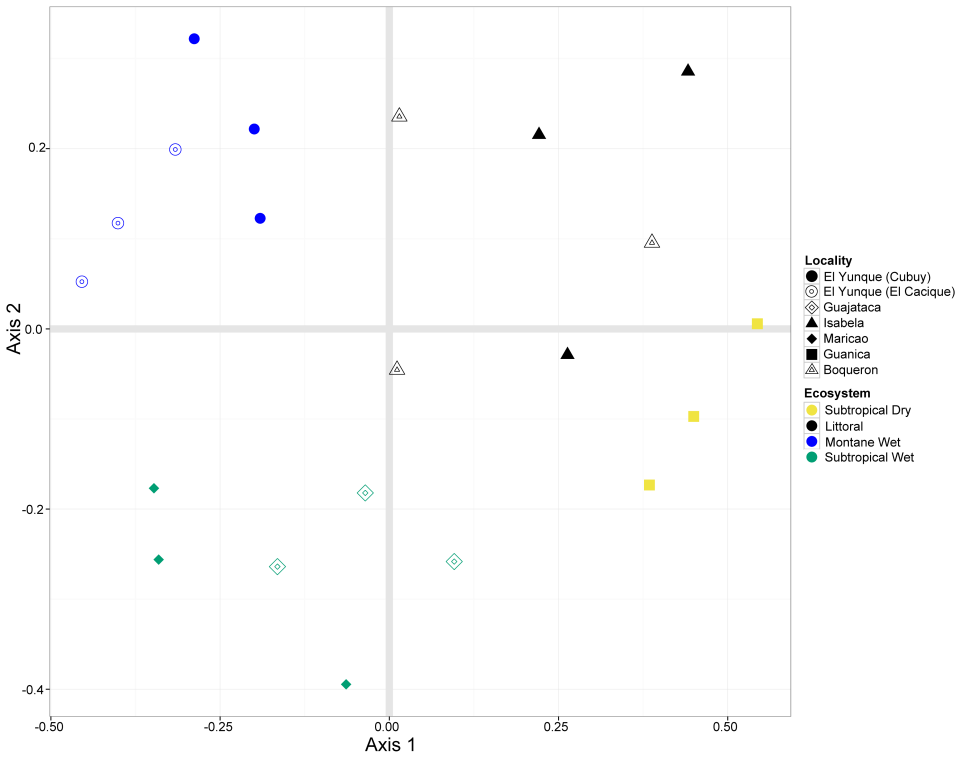
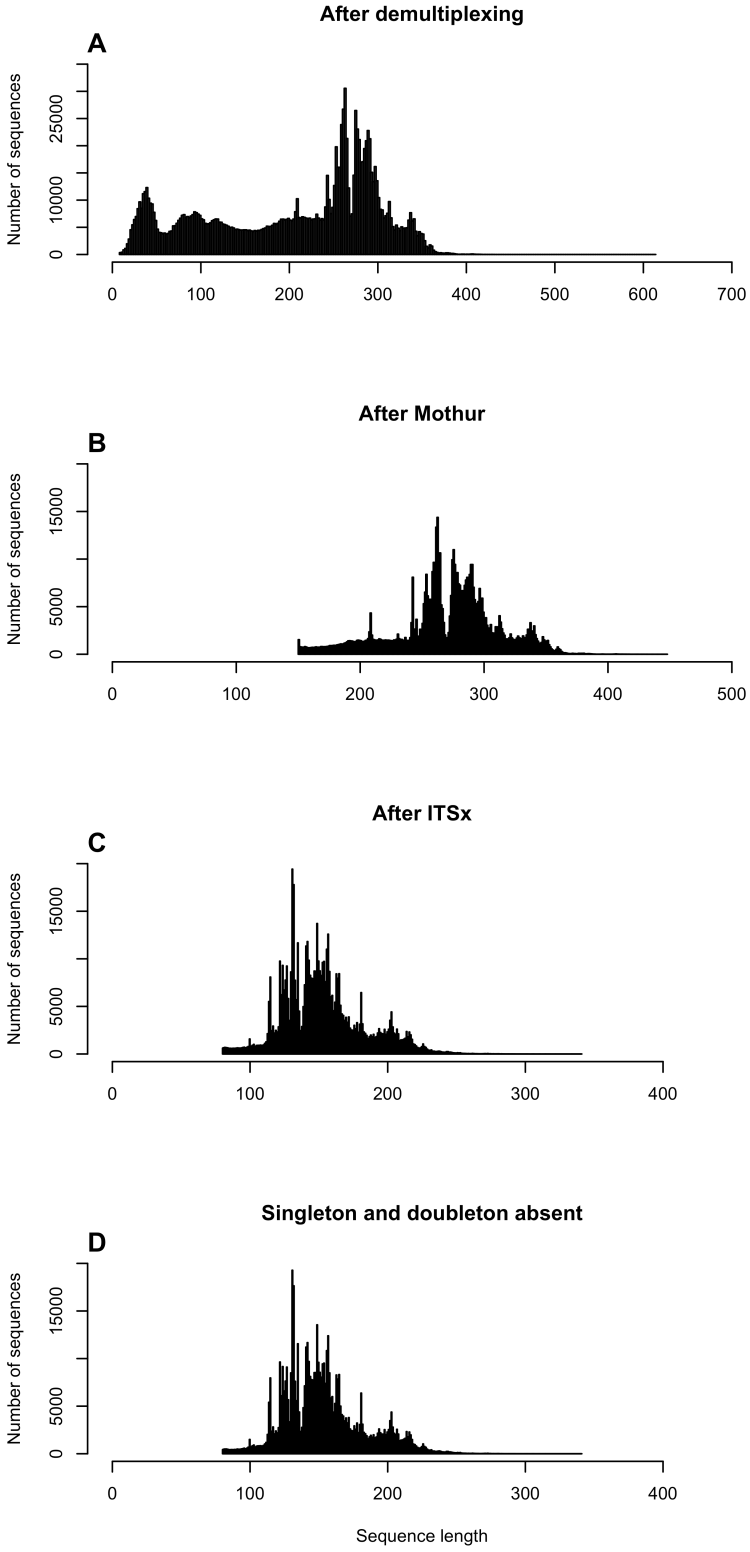

Observed
Chao-1
Shannon
Alpha diversity mean
Depth
0
30000
60000
90000
500
1000
1500
1
2
3
4
5
Locality
El Yunque (Cubuy)
El Yunque (El Cacique)
Guajataca
Isabela
Maricao
Guanica
Boqueron

Figure 5

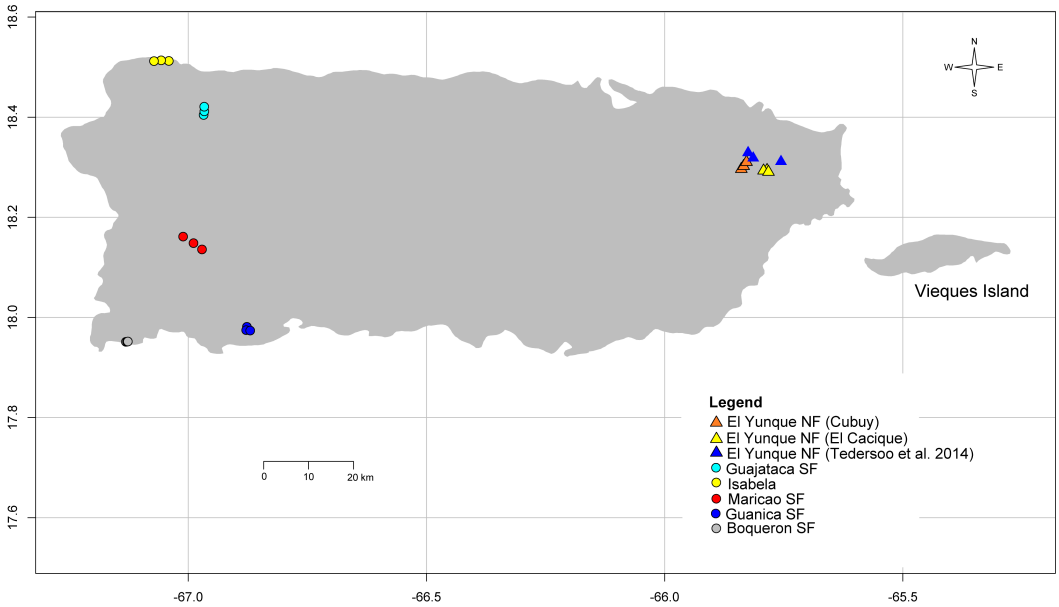


Supplementary Figures

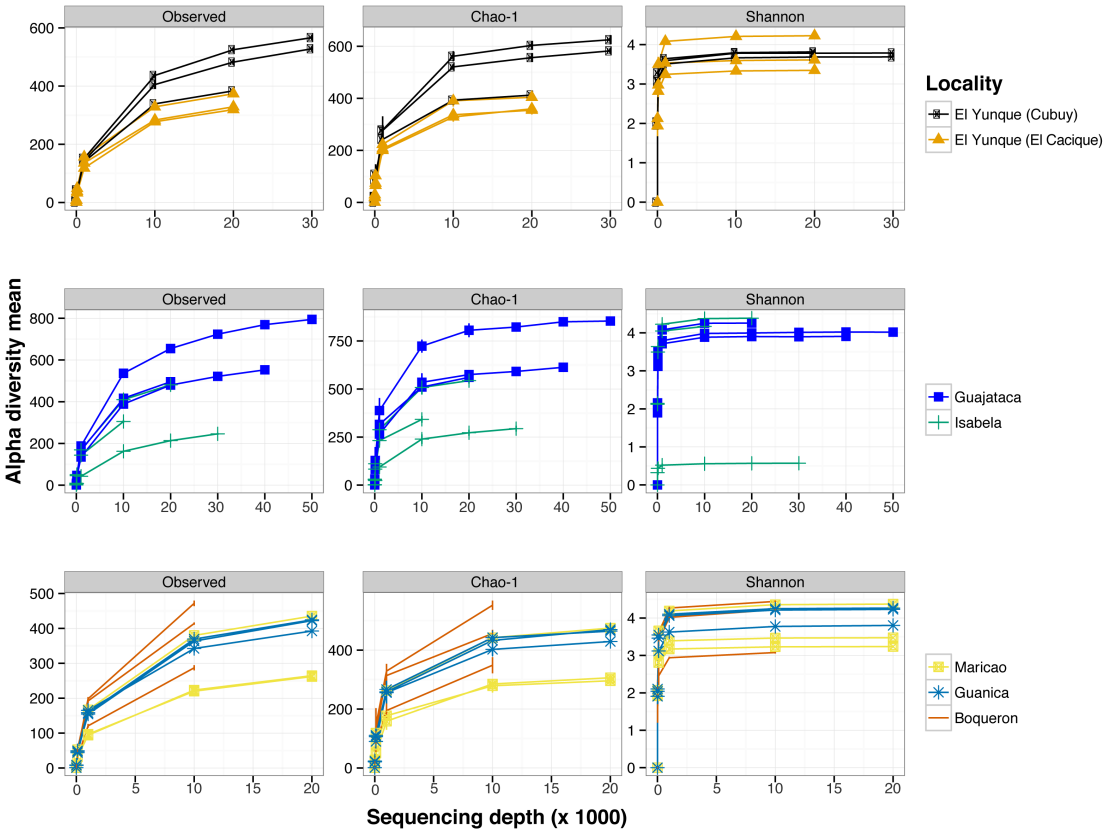
Supplementary Fig. S1



Supplementary Fig. S2



Supplementary Figure S3



Supplementary Figure S4

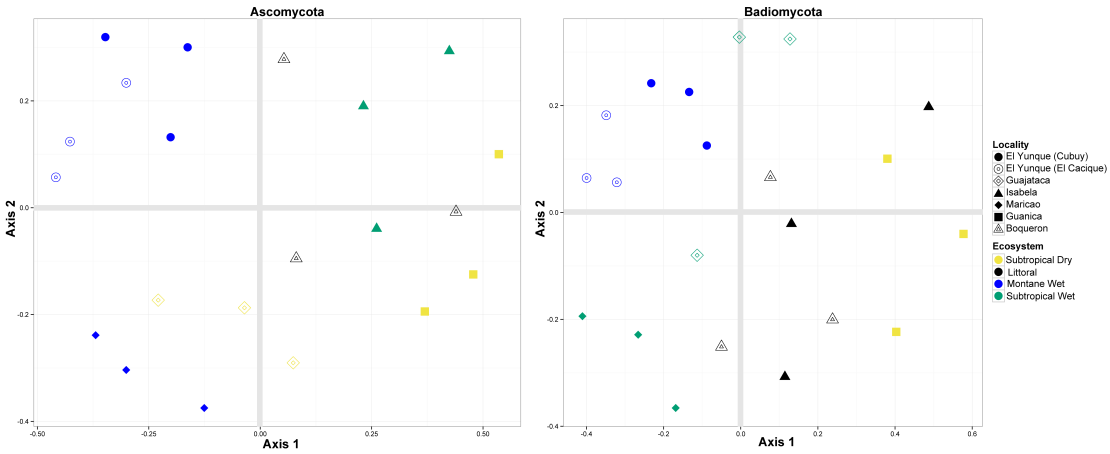


Table 1

Sample Number	Locality	Coordinates	Ecosystem*	Barcode primer 5' – 3'	Total number of sequences		
					After demultiplexing	Final dataset*	Sequence retention (%)
1	El Yunque NF (Cubuy Dam)	18.294128, -65.785082	Lower Montane Wet Forest	TTAAGCGGTCGAT	80,722	40,259	49.87
2		18.293129, -65.792442		TCGAAGGCAGGCGAT	59,341	27,009	45.51
3		18.290552, -65.782132		CGAGGTTATCGAT	75,344	33,879	44.97
4	El Yunque NF (El Cacique sector)	18.296329, -65.839069	Lower Montane Wet Forest	TCCTGGCACATCGAT	58,243	25,858	44.4
5		18.302196, -65.834263		CTGCAAGTTCGAT	49,970	23,121	46.27
6		18.310345, -65.828255		TCTGGATGACGAT	60,935	26,506	43.5
7	Guajataca SF	18.404516, -66.967743	Subtropical Moist Forest	TAAGGAGAACGAT	92,470	45,570	49.28
8		18.411682, -66.966284		TTGGCATCTCGAT	112,266	52,651	46.9
9		18.420966, -66.966498		TCAGGAATACGAT	62,050	22,291	35.92
10	Isabela	18.512390, -67.040742	Littoral; Introduced *Casuarina* trees	CTGACATAATCGAT	85,735	38,702	45.14
11		18.513611, -67.056878		ATCCGGAATCGAT	48,296	22,771	47.15
12		18.512065, -67.071898		CTAGGACATTCGAT	73,434	30,131	41.03
13	Maricao SF	18.161489, -67.010787	Subtropical Wet Forest	TGAGGCTCCGACGAT	51,733	24,344	47.06
14		18.148440, -66.988986		AACCTCATTCGAT	63,342	30,978	48.91
15		18.135879, -66.971133		TTCTCATTGAACGAT	64,673	28,129	43.49
16	Guanica SF	17.981283, -66.876891	Subtropical Dry Forest	CTTGTCCAATCGAT	53,310	24,832	46.58
17		17.974670, -66.878436		CAGCATTAATTCGAT	50,953	24,479	48.04
18		17.973813, -66.869982		TCACTCGGATCGAT	67,312	28,341	42.1

19		17.951319, -67.130843		CGATCGGTTCGAT	48,091	16,207	33.7
20	Boqueron SF	17.951602, -67.128362	Littoral; Mangrove	CCAGCCTCAACGAT	48,509	17,904	36.91
21		17.951948, -67.126791		TGGAGGACGGACGAT	52,928	19,774	37.36
22	Positive Control	-	-	CCTACTGGTCGAT	80,556	43,956	54.6
-	Total	-	-	**-**	**1,440,213**	**648,535**	**45.21**

Ion Torrent adaptors	**Sequence 5' – 3'**
Reverse	CCTCTCTATGGGCAGTCGGTGAT
Forward	CCATCTCATCCCTGCGTGTCTCCGACTCAG

* After removal of singleton and doubletons

Table 2

MOTU number / %	El Yunque (Cubuy)	El Yunque (El Cacique)	Guajataca	Maricao	Isabela	Boqueron	Guanica	Avegare	SDV
El Yunque (Cubuy)	**1017**	272	224	155	146	258	84	**189.8**	73.3
El Yunque (El Cacique)	26.8	**754**	148	177	94	181	62	**155.7**	73.8
Guajataca	22	19.6	**1207**	259	171	229	238	**211.5**	42.6
Maricao	15.2	23.5	21.5	**797**	148	364	93	**199.3**	97.0
Isabela	14.4	12.5	14.2	18.6	**795**	311	283	**192.2**	85.5
Boqueron	25.4	24	19	45.7	39.1	**1120**	243	**264.3**	64.5
Guanica	8.3	8.2	19.7	11.7	35.6	26	**936**	162.2	97.6
Average	**18.7**	**19.1**	**18.1**	**24.3**	**23.5**	**24**	**17.7**		
STD	7.2	7.3	4	12.7	11.3	5.8	10.6		

Supplementary Table S1

MOTU	*Test-Statistic*	Read abundance mean							Taxonomy
		El Yunque (Cubuy)	El Yunque (El Cacique)	Guajataca	Isabela	Maricao	Guanica	Boqueron	
MOTUHB 1	4918.177	0.000	0.000	0.000	10.333	1359.333	1.333	62.333	Archaeorhizomyces sp.
MOTUHB 25	962.737	0.000	0.000	0.000	0.667	252.333	0.333	0.333	Archaeorhizomyces sp.
MOTUHB 36	777.997	0.000	0.000	70.333	0.000	214.000	0.000	0.667	Archaeorhizomyces sp.
MOTUHB 42	541.375	0.000	0.000	1.000	0.000	143.333	0.000	0.333	Archaeorhizomyces sp.
MOTULB 128	435.721	0.000	56.667	0.667	1.000	129.333	0.333	9.000	Archaeorhizomyces sp.
MOTULB 2252	108.103	31.000	0.000	0.000	0.000	0.000	0.000	2.667	Archaeorhizomyces sp.
MOTULB 272	157.131	0.000	41.667	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 475	97.322	28.000	0.000	0.000	0.000	0.000	0.000	2.333	Archaeorhizomyces sp.
MOTULB 91	240.207	31.000	1.000	0.000	64.333	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 97	253.189	70.667	0.000	0.000	0.000	0.000	0.000	5.000	Archaeorhizomyces sp.
MOTUHB 12	1810.686	554.333	12.667	52.333	0.000	1.667	0.000	19.667	Archaeorhizomyces sp.
MOTUHB 18	1709.525	483.333	4.667	0.000	0.000	0.000	0.000	44.333	Archaeorhizomyces sp.
MOTUHB 19	1166.578	8.333	314.333	0.000	0.000	0.000	0.000	0.667	Archaeorhizomyces sp.
MOTUHB 29	707.052	0.000	0.000	0.000	0.000	183.000	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 3	4038.065	1100.667	0.000	0.000	0.333	0.000	0.000	58.667	Archaeorhizomyces sp.
MOTUHB 6	2592.543	670.333	0.000	0.000	0.000	0.000	0.000	1.000	Archaeorhizomyces sp.
MOTUHB 7	1645.185	0.000	0.000	427.000	0.000	0.667	0.000	0.333	Archaeorhizomyces sp.
MOTUHB 8	2610.898	0.000	0.000	710.667	0.000	250.000	0.333	1.667	Archaeorhizomyces sp.
MOTULB 100	565.658	0.000	146.667	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 104	154.418	44.000	0.000	0.000	0.000	0.000	0.000	4.667	Archaeorhizomyces sp.
MOTULB 137	201.658	0.000	0.000	1.000	0.000	55.333	0.000	0.333	Archaeorhizomyces sp.
MOTULB 142	122.147	0.000	0.000	0.000	0.000	32.667	0.000	0.000	Archaeorhizomyces sp.
MOTULB 155	309.971	86.333	0.000	0.000	0.000	0.000	0.000	6.333	Archaeorhizomyces sp.
MOTULB 168	149.355	0.000	39.667	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 186	127.418	0.000	0.000	0.000	0.000	35.000	0.000	0.667	Archaeorhizomyces sp.
MOTULB 187	136.397	0.000	0.000	0.000	0.000	36.333	0.000	0.000	Archaeorhizomyces sp.
MOTULB 211	110.491	0.000	29.667	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 328	292.271	0.000	0.000	0.000	0.000	82.667	0.000	12.333	Archaeorhizomyces sp.
MOTULB 4180	502.097	0.000	130.333	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 10	2297.682	194.667	628.000	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 11	1145.860	359.333	10.667	4.667	2.333	202.667	1.000	69.000	Archaeorhizomyces sp.
MOTUHB 20	3023.307	799.000	11.667	0.000	0.000	0.333	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 22	608.916	0.000	0.000	158.667	0.000	0.333	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 37	485.235	0.000	0.000	126.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 4	3168.571	0.000	0.000	831.333	0.000	3.333	0.667	2.333	Archaeorhizomyces sp.
MOTUHB 47	4112.814	0.000	1059.333	0.000	0.333	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 9	2165.707	562.000	0.000	0.333	0.667	0.333	0.000	0.000	Archaeorhizomyces sp.
MOTULB 107	302.351	79.000	0.000	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.

Supplementary Table S1 (continuation)

MOTU	*Test-Statistic*	Read abundance mean							Taxonomy
		El Yunque (Cubuy)	El Yunque (El Cacique)	Guajataca	Isabela	Maricao	Guanica	Boqueron	
MOTULB 10881	417.785	0.000	0.000	108.667	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 123	268.632	0.000	0.000	0.000	0.000	70.333	0.000	0.000	Archaeorhizomyces sp.
MOTULB 179	92.369	25.000	0.000	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 2389	219.765	1.333	59.667	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 3310	279.007	0.000	0.000	73.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 384	119.557	0.000	32.000	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 55	778.843	217.333	0.000	0.000	0.000	0.000	0.000	23.667	Archaeorhizomyces sp.
MOTULB 56	639.597	0.000	0.000	0.000	0.000	165.667	0.000	0.000	Archaeorhizomyces sp.
MOTULB 6606	275.116	0.000	0.000	72.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 79	511.098	0.000	0.000	0.000	3.000	149.000	0.333	16.000	Archaeorhizomyces sp.
MOTULB 80	707.399	184.000	0.000	0.000	0.333	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 26	876.740	244.000	0.000	0.000	0.000	0.000	0.000	24.333	Archaeorhizomyces sp.
MOTUHB 45	434.504	0.000	0.000	115.333	0.000	1.333	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 48	1829.246	0.000	0.000	475.667	0.333	0.333	0.333	0.333	Archaeorhizomyces sp.
MOTULB 83	221.793	0.000	0.000	61.333	0.000	19.667	0.000	0.000	Archaeorhizomyces sp.
MOTUHB 46	905.297	148.000	6.667	6.667	267.000	3.667	1.333	6.000	Archaeorhizomyces DQ309123
MOTULB 11873	228.195	35.333	6.667	7.667	84.333	0.000	6.333	1.000	Archaeorhizomyces HM069393
MOTULB 101	163.361	7.667	56.000	0.333	6.000	0.000	0.000	3.333	Archaeorhizomyces HQ212004
MOTULB 3764	83.314	0.000	22.667	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 148	82.021	0.000	0.000	22.333	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 196	76.849	21.000	0.000	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 591	72.972	0.000	0.000	20.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 347	69.155	0.000	0.000	0.000	0.333	0.000	20.000	0.333	Archaeorhizomyces sp.
MOTULB 492	69.096	0.000	19.000	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 13130	64.275	18.667	0.000	0.000	0.000	0.000	0.000	1.000	Archaeorhizomyces sp.
MOTULB 240	62.215	18.333	0.000	0.000	0.000	0.000	0.000	1.667	Archaeorhizomyces sp.
MOTULB 356	58.769	0.000	16.333	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 314	58.307	0.000	0.000	0.000	0.000	16.667	0.000	0.333	Archaeorhizomyces sp.
MOTULB 545	57.479	0.000	16.000	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 795	56.190	0.000	15.667	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 716	51.875	0.000	0.000	0.667	0.000	15.667	0.000	0.333	Archaeorhizomyces sp.
MOTULB 430	50.622	0.000	0.000	12.000	0.000	9.333	0.000	0.000	Archaeorhizomyces sp.
MOTULB 262	49.748	14.000	0.000	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 10989	42.438	0.667	0.000	0.667	0.000	13.333	0.000	0.000	Archaeorhizomyces sp.
MOTULB 231	40.128	3.333	0.333	0.000	0.667	3.333	0.667	16.000	Archaeorhizomyces DQ233781
MOTULB 581	39.464	0.000	0.000	11.333	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 251	37.052	0.000	0.000	11.333	0.000	1.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 1541	33.475	0.000	4.333	0.000	0.000	10.333	0.000	0.667	Archaeorhizomyces sp.

Supplementary Table S1 (continuation)

MOTU	*Test-Statistic*	Read abundance mean							Taxonomy
		El Yunque (Cubuy)	El Yunque (El Cacique)	Guajataca	Isabela	Maricao	Guanica	Boqueron	
MOTULB 598	30.987	0.000	2.333	10.000	0.000	0.000	0.333	0.000	Archaeorhizomyces GQ160171
MOTULB 613	29.232	8.667	0.000	0.000	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 601	25.417	0.000	0.000	7.667	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 1301	22.884	0.000	0.000	0.000	0.000	7.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 804	22.753	0.000	0.000	0.000	0.000	7.333	0.000	0.667	Archaeorhizomyces sp.
MOTULB 1062	17.853	0.000	0.000	5.667	0.000	0.000	0.000	0.000	Archaeorhizomyces sp.
MOTULB 673	16.999	5.667	0.000	0.000	0.000	0.000	0.000	0.333	Archaeorhizomyces sp.
MOTULB 709	16.999	5.667	0.000	0.000	0.000	0.000	0.000	0.333	Archaeorhizomyces sp.