Research paper

Contrasting the microbiomes from forest rhizosphere and deeper bulk soil from an Amazon rainforest reserve

Jose Pedro Fonseca\textsuperscript{a,b,⁎}, Luisa Hoffmann\textsuperscript{b}, Bianca Catarina Azeredo Cabral\textsuperscript{a}, Victor Hugo Giordano Dias\textsuperscript{a}, Marcio Rodrigues Miranda\textsuperscript{c}, Allan Cezar de Azevedo Martins\textsuperscript{a}, Clarissa Boschiero\textsuperscript{b}, Wanderley Rodrigues Bastos\textsuperscript{d}, Rosane Silva\textsuperscript{a}

\textsuperscript{a}Universidade Federal do Rio de Janeiro, Instituto de Biofísica Carlos Chagas Filho, Rio de Janeiro, RJ 21941-902, Brazil
\textsuperscript{b}The Noble Research Institute, Ardmore, OK 73401, USA
\textsuperscript{c}Universidade Federal de Rondônia, Núcleo de Ciência e Tecnologia, Porto Velho, RO 76815800, Brazil
\textsuperscript{d}Universidade Federal de Rondônia, Departamento de Biologia, Porto Velho, RO 76801059, Brazil

ARTICLE INFO

Keywords:
Amazon Tropical Forest Metagenomics Shotgun NGS

ABSTRACT

Pristine forest ecosystems provide a unique perspective for the study of plant-associated microbiota since they host a great microbial diversity. Although the Amazon forest is one of the hotspots of biodiversity around the world, few metagenomic studies described its microbial community diversity thus far. Understanding the environmental factors that can cause shifts in microbial profiles is key to improving soil health and biogeochemical cycles. Here we report a taxonomic and functional characterization of the microbiome from the rhizosphere of \textit{Brosimum guianense} (Snakewood), a native tree, and bulk soil samples from a pristine Brazilian Amazon forest reserve (Cuniã), for the first time by the shotgun approach. We identified several fungi and bacteria taxon significantly enriched in forest rhizosphere compared to bulk soil samples. For archaea, the trend was the opposite, with many archaeal phylum and families being considerably more enriched in bulk soil compared to forest rhizosphere. Several fungal and bacterial decomposers like \textit{Postia placenta} and \textit{Catenulispora acidiphila} which help maintain healthy forest ecosystems were found enriched in our samples. Other bacterial species involved in nitrogen (\textit{Nitrobacter hamburgensis} and \textit{Rhodopseudomonas palustris}) and carbon cycling (\textit{Oligotropha carboxidovorans}) were overrepresented in our samples indicating the importance of these metabolic pathways for the Amazon rainforest reserve soil health. Hierarchical clustering based on taxonomic similar microbial profiles grouped the forest rhizosphere samples in a distinct clade separated from bulk soil samples. Principal coordinate analysis of our samples with publicly available metagenomes from the Amazon region showed grouping into specific rhizosphere and bulk soil clusters, further indicating distinct microbial community profiles. In this work, we reported significant shifts in microbial community structure between forest rhizosphere and bulk soil samples from an Amazon forest reserve that are probably caused by more than one environmental factors such as rhizosphere and soil depth.

1. Introduction

Organisms that live in the soil include plants, microbes (archaea, bacteria, fungi, algae and lichens), invertebrates and vertebrates. They are important to soil health because of their role in decomposition, storage and flux of carbon, nutrients and soil structure. More importantly, these organisms are partners in symbiotic relationships with plants, supplying the nutrients necessary for plant productivity. Microbial community activity of the rhizosphere influences several plant physiological processes such as growth and energy metabolism affecting overall plant health. Thus, determining the taxonomical diversity and functional properties of the rhizosphere microbiome is essential to characterize the composition of its associated microbial communities. The Amazon forest is one of the most diverse biomes on Earth, with thousands of plant and animal species described throughout the region (Asner et al., 2014; Coley and Kursar, 2014). However, even

https://doi.org/10.1016/j.gene.2017.11.039
Received 30 December 2016; Received in revised form 31 October 2017; Accepted 13 November 2017
Available online 15 November 2017
0378-1119/ © 2017 Elsevier B.V. All rights reserved.
though microorganisms play a key role on the dynamics of all communities in the environment, little is known about their microbiome population structure and function. With the shotgun sequencing method (Venter et al., 1998) combined with today’s next generation sequencing (NGS) capabilities it is possible to survey, using metagenomic analysis, the microbiome of virtually any biome on the planet from the most inhospitable to the more densely populated areas.

The rhizosphere can be broadly defined as a region in the soil that is under direct contact with plant roots, immediately surrounding roots or soil adjacent to root tips, and is influenced by root exudates, secretions and mucilages. In this environment, several biotic and abiotic interactions take place between rhizosphere-dwelling microbes and plant roots that can affect a plant’s physiological processes such as development, defense and nutrition (Philippot et al., 2013; Wieland et al., 2001; Marschner et al., 2001; Meier et al., 2015; Bais et al., 2006; Lundberg et al., 2012). Indeed, the rhizosphere microbiome can influence plant physiological processes such as growth and energy metabolism while plants in turn secrete metabolites and mucilages that can affect the growth of specific microorganisms nurturing a great microbial diversity.

Soils are heterogeneous environments where several physicochemical factors can affect the growth and diversity of its microbial population (Gupta, 2011). Indeed, previous studies had shown that soils are extremely rich in microbial diversity (Daniel, 2005), as reported before by Timmusk et al. (2011), that showed up to 4 × 10^6 different taxa.

Other species of bacteria can benefit the plant by the formation of protective biofilms in the roots and antibiotic secretion that can protect the plant against harmful pathogens (Kuiper et al., 2004). Plants in turn, can influence the microbiome around its roots by secretion of secondary metabolites, amino acids, sugars, enzymes, mucilage and dead cell debris which in turn can attract or inhibit the growth of specific microorganisms (Kuiper et al., 2004; Mendes et al., 2014) including disease agents.

Many metagenomic rhizosphere studies have previously focused on the description of model plants (Arabidopsis thaliana) or crop system (rice and soy) microbiomes, in cultivated conditions and mostly restricted to agricultural areas (Mendes et al., 2014; Lundberg et al., 2013). Few studies to date have characterized the microbiome from forest rhizosphere and bulk soil from a pristine site directly without further greenhouse cultivation methods (e.g. mesocosm experiments) from the Amazon rainforest in Brazil (Meyer et al., 2017; Paula et al., 2014; Navarrete et al., 2013). Brosimum guianense (snakewood) is a native tree widespread in the Amazon rainforest which is widely known in the region for its use in the timber industry (Scholz et al., 2007) and its macerated bark is used as anti-inflammatory and in the treatment of several skin infections by indigenous and local people. Other Brosimum sp. genera are used for latex production (Homobono et al., 2014).

In this work, we conducted a metagenomic analysis of the microbiome present in the Snakewood forest rhizosphere and bulk soil from the Amazon forest biome to characterize its microbiome taxonomic diversity and potential functional properties. Few studies to date have described the microbiota using shotgun sequencing directly from a pristine site in the Amazon forest biome. Changes that cause disequilibrium on the dynamics of the microbiome from the Amazon soil and rhizosphere can be detrimental to this important biodiversity hotspot.

2. Materials and methods

2.1. Ethics statement

For sampling in the protected federal reserve of Cuniã, the authors obtained all necessary permits from federal and local authorities. Permission number 63/2013 was issued by the Chico Mendes Institute in the Cuniã natural reserve and sampling was authorized for research purposes in the area. This area plays a significant role in protecting against deforestation and important headwaters. The access is restricted to scientific research, mainly inserted in the Biodiversity Research Program (PPBio) that promotes interdisciplinary research in the Brazilian Amazon region (http://ppbio.inpa.gov.br/en/home).

2.2. Sample collection and processing

The Cuniã reserve is a federal protected area (Fig. 1) located 120 km from Porto Velho (Brazil) in the Purus-Madeira interfluve that is part of the Brazilian Amazon rainforest (Global Positioning System coordinates – 8.376770 S, – 63.631439 W). Rhizosphere and bulk soil samples were collected in three different plots in the PPBio sample grid (25 km²) from the Cuniã reserve. Physicochemical parameters were measured for rhizosphere and bulk soil sites (S1 Table). Three rhizosphere and bulk soil samples were collected from three different plots (1 km²) belonging to the PPBio sample grid in June 2013. Each rhizosphere sample (soil up to 1 mm attached to roots) was collected from a thick root system from Brosimum guianense (Snakewood) weighting about 1 Kg covering the Amazon forest soil (1–30 cm in depth). Around 2 kg of bulk soil samples were also collected from the same plots belonging to the Cuniã grid, 10 cm below the rhizosphere, with the help of an auger in October 2013. All samples were transported in coolers with ice.

2.3. Extraction of total community DNA and sequencing

Total community DNA was obtained from 0.3 g of rhizosphere or bulk soil using the Power Soil DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA) for each sample collected. Overall DNA quality was estimated by electrophoresis in 1% agarose gels in 1× TAE buffer (40 mM Tris, pH 8.3; 20 mM acetic acid; 1 mM EDTA), stained with 0.5 μg/mL ethidium-bromide solution and quantified using a spectrophotometer instrument (Thermo Scientific NanoDrop 2000). Total community DNA was sequenced at Laboratório de Metabolismo Macromolecular Firmino Torres de Castro (Biophysics Institute, Rio de Janeiro Federal University) using an Ion Xpress PGM sequencer (Life Sciences). The shotgun library kit used for sequencing was Ion Xpress Plus Fragment Library kit (Life Biosciences) with barcodes (Ion Xpress Barcode Adapters kit, Life Sciences) followed by PCR emulsion using the Ion PGM Template OT2 400 kit and the Ion PGM Sequencing 400 kit (Life Sciences) for shotgun sequencing. The three barcoded rhizosphere samples were run in parallel in a 318 chip yielding 1.5 Gigabase of total sequence information. Bulk soil samples were also run in triplicates in a 318 chip generating 1.6 Gigabase of total sequence information. The metagenomic data were stored on the MG-RAST server (http://metagenomics.anl.gov) under id numbers (4535167,3, 4535136.3, 4535204.3 for forest rhizosphere, and 4579275, 4579276, 4579277 for bulk soil) and it is available upon request.

2.4. Annotation of metagenomic sequence reads and data analysis

A read-based shotgun sequencing method was used in combination with MG-RAST server semi-automated pipeline to perform read annotation and obtain taxon relative abundance values (DeLong, 2005; Kusand et al., 2012; Thomas et al., 2012). All FASTQ files were uploaded to the MG-RAST server and processed using the metagenomics RAST pipeline version 4.0.2 (Meyer et al., 2008; Aziz et al., 2008) with the default parameters for the initial quality control (QC) filtering of raw reads, such as removal of low quality reads and dereplication (removal of artificial sequences produced by sequencing artifacts). Organism abundance and functional analysis were performed using a “Best Hit Classification” approach with a maximum e-value of 1e − 05, minimum identity of 60% and 50 bp for minimum alignment length using IMG (Markowitz et al., 2008) and Subsystems (Overbeek, 2005; Overbeek et al., 2014) databases to annotate source data. In order to reduce the impact of experimental noise, the normalized data option of MG-RAST was used. Data were additionally normalized and all values were
transformed to log base 2. Box plot distribution analysis of functional categories was computed using ggplot from R (Wickham, 2009). Rarefaction curves of annotated species richness were generated from the total number of distinct species annotated as a function of the number of sequences sampled. Principal coordinate analysis (PCoA) was generated using Euclidean distances from taxonomy relative abundance data for contrasting rhizosphere and bulk soil samples. We selected bulk soil samples from the Amazon region (MGRAST ids 4493544, 4493546, 4493545, 4493652, 4493650) and soya mesocosm rhizosphere (MG RAST ids 4477749, 4477751, 4477755, 4478290, 4478291, 4478292) from the Amazon forest soil which was the most similar Amazon rhizosphere sample found in the database at the time of this work.

2.5. Statistical analysis

Reads were classified and normalized into MG-RAST taxonomy and functional annotation pipeline (IMG and subsystem respectively). MG-RAST normalized relative abundance values were log transformed and used to compute statistical significance between two groups (rhizosphere and bulk soil) for each taxon as differences amongst means (n = 3) for each group by multiple t-test (Frese et al., 2015; White et al., 2008) using alpha = 0.05 with GraphPad Prism 6 software (GraphPad software Inc.). For the heatmap figures, hierarchical clustering of distances using taxonomy relative abundance data (phylum and species level) was generated using Euclidean distance calculation to obtain a matrix and dendrogram construction using R packages.

3. Results

3.1. Sequencing and initial filtration

A total of 6,619,179 raw reads (1,514,299,093 bp) were generated from three rhizosphere shotgun metagenomic replicates representative of the Amazon forest pristine site with a mean length of 228 bp. For bulk soil samples a total of 6,536,882 raw reads with a mean length of 228 bp were generated (Table 1). The raw read data were subjected to the MG-RAST QC filtering pipeline to remove low quality and artificial duplicate reads. A total of 157,456 (2.3%) duplicated reads and 479,839 (7.24%) low-quality reads were removed from the total reads.

Table 1

<table>
<thead>
<tr>
<th>Description</th>
<th>Rhizosphere 1 (FR-1)</th>
<th>Rhizosphere 2 (FR-2)</th>
<th>Rhizosphere 3 (FR-3)</th>
<th>Bulk soil 1 (BS-1)</th>
<th>Bulk soil 2 (BS-2)</th>
<th>Bulk soil 3 (BS-3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG-RAST ID</td>
<td>4535167.3</td>
<td>4535136.3</td>
<td>4535204.3</td>
<td>4579277.3</td>
<td>4579275.3</td>
<td>4579276.3</td>
</tr>
<tr>
<td>Total read number</td>
<td>2,715,594</td>
<td>1,293,070</td>
<td>2,610,515</td>
<td>2,694,917</td>
<td>1,725,895</td>
<td>2,116,070</td>
</tr>
<tr>
<td>Total bp</td>
<td>630,081,836</td>
<td>296,443,716</td>
<td>587,773,541</td>
<td>648,451,865</td>
<td>413,336,354</td>
<td>435,889,590</td>
</tr>
<tr>
<td>Average length</td>
<td>232 ± 83</td>
<td>229 ± 83</td>
<td>225 ± 87</td>
<td>240 ± 90</td>
<td>239 ± 90</td>
<td>205 ± 84</td>
</tr>
<tr>
<td>Total read post-QC</td>
<td>2,534,137</td>
<td>1,209,060</td>
<td>2,396,143</td>
<td>1,762,019</td>
<td>1,084,955</td>
<td>1,412,839</td>
</tr>
<tr>
<td>Total bp post-QC</td>
<td>500,766,202</td>
<td>235,050,858</td>
<td>451,546,352</td>
<td>224,041,336</td>
<td>137,393,334</td>
<td>178,780,356</td>
</tr>
<tr>
<td>Average length post-QC</td>
<td>197 ± 75</td>
<td>194 ± 74</td>
<td>188 ± 75</td>
<td>127 ± 61</td>
<td>126 ± 60</td>
<td>126 ± 55</td>
</tr>
<tr>
<td>Processed protein features</td>
<td>2,245,636</td>
<td>1,076,907</td>
<td>2,104,854</td>
<td>1,321,026</td>
<td>821,804</td>
<td>1,085,823</td>
</tr>
<tr>
<td>Processed rRNA features</td>
<td>367,121</td>
<td>179,713</td>
<td>365,664</td>
<td>30,380</td>
<td>18,554</td>
<td>24,289</td>
</tr>
</tbody>
</table>

FR-forest rhizosphere, BS-bulk soil.
of the forest rhizosphere generated during QC trimming before the metagenomic annotation procedure. A total of 2,277,069 (34.8%) low quality reads and 245,666 (3.75%) duplicated reads were removed from the bulk soil samples.

### 3.2. Taxonomic analysis of metagenomes

Relative abundance at the domain level revealed that bacteria is the major domain present in the forest rhizosphere (> 93%), with Eukaryota and Archaea being less prevalent (< 5% and < 1%, respectively) by the IMG protein database (Figs. S1A and S1B). Bacteria also dominated the bulk soil samples (96%) with only a small fraction contributed by Eukaryotes (1%) and Archaea (< 2%). Annotation quality indicated consistent e-value scores for all databases used with the majority of sequence matches above the e-value cutoff of 1e-05 (S1B Fig.). To estimate the fraction of species sequenced in the surveyed environments rarefaction curves were generated for forest rhizosphere and bulk soil metagenomes (Fig. 2A). Bulk soil metagenomes generated flatter curves with lower species count than rhizosphere metagenomes indicating that nearly all species for this metagenome were sampled. Given that all sequenced metagenomes displayed a similar read coverage and depth, the observed rarefaction curves indicate a lower taxonomical richness for bulk soil samples in comparison to forest rhizosphere, which is a hotspot for microbial diversity (Table 1 and Fig. 2A). The fact that each environmental metagenome type might have a distinct GC-signature (Reichenberger et al., 2015) prompted a measure of GC content for Amazon forest bulk soil and rhizosphere samples. A similar GC distribution was observed between forest rhizosphere and bulk soil metagenomes with a higher read percentage showing between 60 and 70% GC content (Fig. 2B). The fact that both metagenomes showed similar taxonomical diversity at the domain level (> 90% of both metagenomes is composed of bacteria) could account for this observed similarity, although some difference can be observed in regions with a lower overall GC content (30–40%).

### 3.3. Bacterial taxonomy analysis between forest rhizosphere and bulk soil metagenomes

A detailed comparison of microbial community structure at the phylum level between forest rhizosphere and bulk soil demonstrated that Proteobacteria, Actinobacteria and Acidobacteria were the predominant phyla for both types of sample (Fig. 3). Several other bacterial phyla also showed a significantly (p < 0.05) higher sequence enrichment such as Bacteroidetes, Proteobacteria and Planctomycetes for the forest rhizosphere compared to bulk soil. A heatmap consisting of only statistically significant bacterial species that were highly enriched in the forest rhizosphere compared to the bulk soil samples (p < 0.05) was generated. The increased relative abundance observed for these bacterial species in forest rhizosphere samples suggests that the rhizosphere has the potential to select in favor of beneficial bacteria that actively participate in basic metabolic functions such as nitrogen, carbon cycling and decomposition (S2 Fig.). Several bacterial species commonly found in forest soils, that are related to nitrogen and/or carbon cycling as well as the genes involved in the respective pathways, were identified in our data such as Catenulispora acidiphila, Xanthobacter autotrophicus and Rhodopseudomonas palustris (S2 Table).

### 3.4. Fungal taxonomy analysis between forest rhizosphere and bulk soil metagenomes

For fungi, the Ascomycota and Basidiomycota phyla were highly enriched (p < 0.05) in the forest rhizosphere samples when compared to bulk soil (Fig. 4A). Interestingly, Ascomycota was the most abundant phyla present for forest rhizosphere. The ascomycetes Eurotiomycetes, Sordariomycetes and Dothideomycetes were the predominant classes in forest rhizosphere samples (Fig. 4B). Amongst the Basidiomycota the most over-represented classes were Agaricomycetes, Tremellomycetes and Ustilaginomycetes for all samples (Fig. 4B). Interestingly, amongst the species of Basidiomycota found significantly enriched in the forest rhizosphere samples is the brown-rot fungal decomposer Postia placenta commonly found in forest ecosystems (S3 Fig.). Indeed, forest rhizosphere samples displayed a higher organic matter content than bulk soil samples providing a richer substrate for fungal and bacterial decomposing agents (S1 Table). Another species of basidiomycete highly enriched in forest rhizosphere was the ectomycorrhizal (EMC) symbiont Laccaria bicolor that promotes plant growth and the biotrophic plant pathogen Ustilago maydis which also has a role in degrading lignocellulose (S3 Fig.).

### 3.5. Relative abundances of archaea between forest rhizosphere and bulk soil metagenomes

For archaeal phylum relative abundance, the trend was surprisingly the opposite of fungi and bacteria with two phyla (Thaumarchaeota and Crenarchaeota) being significantly more enriched in the bulk soil samples than forest rhizosphere (Fig. 5). The most abundant archaea phyla observed in both bulk soil and forest rhizosphere samples were Crenarchaeota and Euryarchaeota with Korarchaeota being the least abundant phyla in the samples. Relative abundance profiles for archaea at the family level showed the same trend with a higher sequence enrichment observed for bulk soil compared with forest rhizosphere samples such as the acidophilic Ferroplasmaceae, Cenarchaeaceae, Thermofilaceae, Nitrospumilaceae, Picrophilaceae, Desulphurococaceae and Thermoplasmataceae (S4 Fig.).

### 3.6. Hierarchical clustering of forest rhizosphere and bulk soil samples

Hierarchical clustering based on relative abundance grouped the forest rhizosphere and bulk soil samples in two distinct clades at the microbial phylum level (Fig. 6), and also at the bacterial species level (S2 Fig.). For the microbial phylum heatmap (Fig. 6) all bacterial and fungal phylum appear overrepresented in forest rhizosphere samples compared to bulk soil but archaea samples showed the opposite trend being highly enriched in bulk soil compared to the rhizosphere. At the bacterial species level, relative abundances for > 100 bacterial species only significantly enriched in forest rhizosphere compared to bulk soil revealed several bacterial species with a role in carbon and nitrogen cycling as well as the breakdown of organic matter in the soil (S2 Fig. and S2 Table).

Differences in microbial community structure between forest rhizosphere and bulk soil compared to publicly available metagenomes from MG-RAST database can be visualized in Fig. 7 by Principal Coordinate Analysis (PCoA). Microbiome samples from the same biome (e.g. rhizosphere or bulk soil samples from the Amazon forest reserve) appeared consistently grouped rather than with a different microbiome type, indicating distinct microbial community profiles.

### 3.7. Functional potential of rhizosphere and bulk soil microorganisms from the Cunti forest reserve

Metagenomic read sequences from rhizosphere and bulk soil microorganisms displaying matches to known proteins from MG-RAST Subsystem 1 database were annotated and relative abundancies of functional categories plotted in a boxplot (Fig. 8). In a similar way to the taxonomical analysis, the forest rhizosphere samples were highly enriched (p < 0.05) for all basic metabolic pathways such as energy metabolism (Nitrogen, Phosphorus and Sulphur metabolism) than bulk soil samples. Basic housekeeping functions such as energy metabolism, cellular processes and genetic information processing were the most predominant, while more specific functions such as photosynthesis, dormancy and sporulation and secondary metabolism were less predominant. Functional categories associated with plant processes like...
photosynthesis were significantly more abundant in the rhizosphere metagenomes than bulk soil samples as expected since bulk soil has no direct connection to roots or any plant system.

4. Discussion

It has been estimated that soil organisms excluding roots and vertebrate animals comprise only about 1–2% of total ecosystem biomass, but their importance in ecosystem processes and biodiversity is essential (Vogt et al., 2015). The soil is teeming with microbes that break down senesced plant tissue. This decomposition activity releases nutrients to the plants and helps maintain a healthy forest ecosystem. The three most over-represented bacterial phyla for both forest rhizosphere and bulk soil samples were Proteobacteria, Actinobacteria and Acidobacteria. Other studies comparing rhizosphere and bulk soil microbiome also observed similar changes in microbial community structure (Mendes et al., 2014). The fact that Acidobacteria ranked as the third most abundant group in all our samples is consistent with an acidic pH observed at Cuniã reserve Amazon soil’s biome (Sait et al., 2006). Similarly, soil bacterial community structure analysis in broadleaf forest biomes (Wang et al., 2016) also showed a similar enrichment of Acidobacteria in forest biomes.

A similar distribution was observed between forest rhizosphere and bulk soil metagenomes in GC rich regions but not in low GC content regions (Fig. 2B). It has been previously suggested that both environment and phylogeny can shape nucleotide composition in metagenomes (Reichenberger et al., 2015) therefore the observed differences seen here cannot be explained by phylogeny alone, although it is likely the main driver of nucleotide bias for metagenomes.

Microorganisms can affect biogeochemical cycling and have an important role in the carbon-nitrogen cycle. Bacteria can accumulate energy through nitrogen cycling processes and this phenomenon has been reported previously in soils and other environments (Gao et al., 2016; Uroz et al., 2016). In our samples, the presence of ammonia oxidizing bacteria such as Nitrosospira, Nitrosococcus and Nitrosonomas suggests they can perform autotrophic nitrification which is an indication of CO2-fixation-coupled ammonia oxidation process. Beneficial interaction of rhizosphere microbial communities and plant roots suggest that plants can select taxonomical and functional gene groups. A number of genes involved in the nitrogen cycle were described here from bacterial species overrepresented on rhizosphere and bulk soil samples (S2 Table), suggesting their importance for the biogeochemical cycles in their respective environments.

The presence of microorganisms that colonize the rhizosphere helping plants to acquire nutrients such as phosphorus and potassium and enhancing nitrogen uptake from the soil are consistent with the results obtained here. Supplemental Fig. S2 displays the 121 bacterial species, from the top 300 most overrepresented bacterial species in our samples that are significantly enriched in the forest rhizosphere compared to bulk soil. Many bacterial species commonly found in forest biomes were found overrepresented in our samples. Several have a role in carbon and nitrogen cycle, such as Candidatus solibacter usitatus, Acidobacterium capsulatum and Rhodopseudomonas palustris, amongst others (S2 Table and S2 Fig.). A. capsulatum and C. usitatus belong to the Acidobacteria phylum, which is one of the most abundant in soils (Challacombe et al., 2011). Several genes involved in carbon and nitrogen cycle such as crA (carbon storage regulator A), CODH (carbon monoxide dehydrogenase), cbb family (Calvin-Benson-Bassham cycle) and rbCl (ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit) present in the genome of bacterial species overrepresented in our samples were listed (S2 Table). The bacteria A. capsulatum can metabolize diverse carbon sources and it can grow up in soils with pH up to 6.0, correlating with its vast geographical distribution (Rawat et al., 2012). Bacterial genes important for nitrogen cycling from bacteria present in our Amazon forest samples such as the nir family (nitrite reductase), nif family (nitrogen fixation) and nar family (nitrate
reductase) were also identified (S2 Table).

Hierarchical clustering based on abundance profiles for fungi, archaea and bacteria at the phylum level grouped rhizosphere and bulk soil samples in two distinct clades with fungi and bacteria taxa highly enriched in rhizosphere samples while archaea displayed the opposite trend being more enriched in bulk soil samples (Fig. 6). This separates clustering of rhizosphere and bulk soil samples from the heatmaps (Fig. 6 and S2 Fig.) was also observed for the PCoA analysis results (Fig. 7), further validating the distinct microbial profiles between rhizosphere and bulk soil microbial communities.

The phyla Basidiomycota and Ascomycota were found to be highly enriched for the forest rhizosphere when compared to bulk soil, whereas Ascomycota was the predominant phyla for forest rhizosphere in agreement with previous shotgun-sequenced soil microbiome data (Souza et al., 2013). A recent study also captured a reduction in the relative abundance of microbes, specifically methanotrophs, when comparing pristine forest and pasture in the Amazon region (Meyer et al., 2017). Fungi play a pivotal role as decomposers of plant matter in forest ecosystems and several Ascomycetes and Basidiomycetes are known decomposers of organic matter (Stursova et al., 2012). The wood-decaying basidiomycete Postia placenta was enriched in forest rhizosphere samples compared to bulk soil. P. placenta is a well-known decomposer of organic matter such as wooden structures and an essential role to carbon cycling in forest soil ecosystems (Martinez et al., 2009). Secreted mucilages and dead plant debris present around the rhizosphere at the plant root system provide the substrate whereby fungi and bacteria can act as decomposing agents, playing a key role in degrading such compounds. The basidiomycete plant pathogen Ustilago maydis, also found enriched in forest rhizosphere samples, has been recently implicated in lignocellulose-degradation and can act as an organic matter decomposer in forest ecosystems (Couturier et al., 2012). Additionally, the EMC fungi Laccaria bicolor found enriched in our Amazon forest samples can support a healthy forest ecosystem by providing nutrients to their plant host through a mutualistic symbiotic relationship with host roots (Plett et al., 2014).

Archaea relative abundance data at the phylum and family levels showed the opposite trend to fungal and bacterial abundances, being more enriched in the bulk soil samples than in forest rhizosphere (Fig. 5 and S4 Fig.). Significantly enriched archaea taxa such as Nitrosopumilus which belongs to the Thaumarchaeota phyla, is an ammonia oxidizing archaea, suggesting a potential role on nitrite reductase activity of the soil. Thaumarchaeota, that was found significantly overrepresented in
bulk soil samples compared to forest rhizosphere, is recognized to have a significant role in geochemical cycles (Brochier-Armanet et al., 2008), it is also known to have copper-dependent nitrite reductases (NirK) (De Mandal et al., 2017). Other archaea families participate in many diverse functional processes like energy metabolism such as the sulphur metabolizing Desulphurococcaceae. This shift observed in archaea abundance could be explained in terms of soil depth. It has been shown previously that the archaeal domain was more abundant in the deepest soil layers than at the surface (Eilers et al., 2012), in agreement with the results shown here. Alternatively, it is also possible that archaea are less influenced than other microorganisms, such as bacteria and fungi, by root exudates or it is simply outcompeted by its fungal or bacterial counterparts near the root. Indeed, archaea are likely contributors to soil biogeochemical cycles but their relatively low abundances compared to other domains in the microbiota such as fungi and bacteria hint at a potential compartmentalized role in the microbiome (S1A Fig.). Several significantly enriched archaea families found in bulk soil compared to rhizosphere are related to acidic soils such as the acidophilic Ferroplasmaceae and Picophilaceae.

Overall, the shifts in microbial community profile reported in this study likely reflects the influence of, not only B. guianense rhizosphere, but also other factors such as soil depth which can affect multiple environmental factors like nutrient availability, carbon amounts, pH and organic matter content. It has been shown previously a significant association between the distribution of particular taxa amongst the archaeal, bacterial and fungal communities and certain soil parameters (Uroz et al., 2016). Their study correlation analyses revealed higher scores for soil parameters such as pH, carbon and nitrogen as well as exchangeable iron for fungi compared to bacteria and archaea. In another study, it was reported that Firmicutes responded to additions of organic material to the soil (Cleveland et al., 2007). Also, a significant correlation of actinobacterial community response to different soil factors, such as aluminum, magnesium and carbon was found using 16S pyrosequencing analysis (Barbosa Lima et al., 2015).

In our work the distribution of subsystems 1 functional categories showed basic functions such as metabolism, cell and genetic processes were predominant and highly enriched for rhizosphere samples when compared to bulk soil (Fig. 8). Other more specific subsystems functions
were less predominant and also less enriched for bulk soil compared to rhizosphere such as photosynthesis and dormancy and sporulation, highlighting the importance of these functional groups for plants. The clear trend observed for enriched functional categories in the forest rhizosphere compared to bulk soil suggests that plant roots might provide a more effective substrate in which bacteria and fungi could thrive as opposed to bulk soil. Alternatively, the fact that bulk soil samples were collected at depths 10 cm below the rhizosphere could also play a role in the observed differences, but it is unlikely that soil depth constrains such as reduced oxygen can increase overall housekeeping genes function. It has also been estimated that a significant percentage of the carbon assimilated by the plant (around 20% of the root allocated carbon) is secreted through the root system and utilized by soil microbes as a carbon source for nutrition (Timmusk et al., 2011). This carbon cycling, from the root of the plant to the soil, has an important role in the regulation of beneficial symbiotic interactions between the plant and the soil microbial community (Timmusk et al., 2011).

Mapping the microbial communities present in plant roots such as the rhizosphere can have a broad impact on several economically important areas such as agriculture and natural ecosystem preservation. A recent study using forest biometric data has shown that the Amazon forest carbon sink has started to decline (Brienen et al., 2015). Therefore, novel methods are needed to monitor and characterize changes in microbial community structure and function in key natural ecosystems and commodity crop production that could affect plant productivity and biogeochemical cycles.

5. Conclusion

The results obtained here provide an initial insight into the microbiome diversity and function present in two contrasting microbiomes, forest rhizosphere and bulk soil, from an Amazon forest reserve. Bacterial, fungal and archaea taxonomic analysis indicated a significant shift in microbial population structure between forest rhizosphere and bulk soil samples. Several bacterial and fungal taxa were highly enriched in the forest rhizosphere samples compared to bulk soil, but the opposite trend was observed for archaea with higher sequence enrichment for bulk soil compared to forest rhizosphere samples. Although we identified several microbial taxa that exhibited pronounced changes in relative abundance between bulk soil and rhizosphere, it was difficult to determine the specific factors driving these microbial profile changes since several environmental factors change between rhizosphere and bulk soil and soil depth (nutrient availability, organic matter, carbon availability). A more comprehensive understanding of the drivers of the observed changes in microbial profile distribution reported here will be possible in the future with an increase in the sampling number of different soil depths. The Amazon biome is undergoing constant changes possible in the future with an increase in the sampling number of different soil depths. The Amazon biome is undergoing constant changes due to environmental effects such as bioremediation, metagenomic exploration and perhaps the impact of climate change.

The same agents responsible for decomposition in the soil biota are the ones that produce organic matter, hence if their population is reduced, soil health and plant nutrient acquisition could also decrease (Vogt et al., 2015). Therefore, changes that cause disequilibrium on the dynamics of the microbiome of Amazon soil and rhizosphere can be detrimental to the Amazon forest and the environment. Microbial replacement to a productive soil and rhizosphere biome interaction can be a possible mitigation for anthropogenic damage to the forests.

The data sets supporting the results of this article are available in
https://metagenomics.anl.gov/ under ids: MG-RAST: 4535167.3, 4535136.3, 4535204.3 for forest rhizosphere, and 4579275, 4579276, 4579277 for bulk soil. Supplementary data associated with this article can be found in the online version, at https://doi.org/10.1016/j.gene.396