UNIVERSIDADE DE SÃO PAULO
CENTRO DE ENERGIA NUCLEAR IN AGRICULTURE

MIRIAM GONÇALVES DE CHAVES

Multianalytical approach in the study of the composition and functionality of the microbiota associated with carbon and nitrogen cycles in soils under preserved and disturbed environments

Piracicaba
2020

MIRIAM GONÇALVES DE CHAVES

Multianalytical approach in the study of the composition and functionality of the microbiota associated with carbon and nitrogen cycles in soils under preserved and disturbed environments

Revised version according to the Resolution CoPGr 6018 from 2011

Thesis presented to Center for Nuclear Energy in Agriculture of the University of Sao Paulo as a requisite to the Doctoral Degree in Sciences

Concentration Area: Biology in Agriculture and Environment

Advisor: Prof. Dr. Acacio Aparecido Navarrete

Co-Advisor: Profa. Dra. Tsai Siu Mui

Piracicaba
2020

Dados Internacionais de Catalogação na Publicação (CIP)

Seção Técnica de Biblioteca - CENA/USP

Chaves, M. G.

Abordagem multianalítica no estudo da composição e funcionalidade da microbiota associada aos ciclos do carbono e nitrogênio em solos sob ambientes preservados e perturbados / Multianalytical approach in the study of the composition and functionality of the microbiota associated with carbon and nitrogen cycles in soils under preserved and disturbed environments / Miriam Gonçalves de Chaves; orientador Acacio Aparecido Navarrete; co-orientadora Tsai Siu Mui. - - Versão revisada de acordo com a Resolução CoPGr 6018 de 2011. - - Piracicaba, 2020.

125 p. : il.

Tese (Doutorado – Programa de Pós-Graduação em Ciências. Área de Concentração: Biologia na Agricultura e no Ambiente) – Centro de Energia Nuclear na Agricultura da Universidade de São Paulo, Piracicaba, 2020.

1. Biologia molecular 2. Carbono 3. Ecologia microbiana 4. Fertilizantes nitrogenados 5. Microbiologia do solo 6. Nitrogênio 7. Sequenciamento genético 8. Uso do solo 9. Vinhaça I. Título

CDU 631.461: 579.26

Elaborada por:
Marilia Ribeiro Garcia Henyei
CRB-8/3631
Resolução CFB Nº 184 de 29 de setembro de 2017

For my daughter Marina

Acknowledgments

I am grateful for my advisors Professor Dr. Acacio Navarrete and Professor Dra. Siu Mui Tsai by the support in my PhD and related studies. Professor Tsai motivated me perform my doctorate and is an example for me of her leading role in scientific research. Professor Acacio helps me in all steps and always encourages me in face of challenges…both are visionaries in Microbial Ecology and examples for me.

I am grateful to Center for Nuclear Energy in Agriculture – USP and your employees that support me in informatics, secretary, laboratory supplies, among others.

To my laboratories colleagues: Luis Fernando Merloti, Dayane Juliate, Leandro Lemos, Leandro Fonseca, Dasiel Obregon, Iara Barros Feitosa, Mariley Fonseca, Paula Henriques Suarez, Andressa Venturini, Alexandre Pedrinho, Julia Brandão Gontijo, Fernanda Nakamura, Caio Yushiura, and Aline Giovana da França for their suggestions in analisys and for good moments discussing science.

I would like to express my gratitude to researchers Dr. Lucas William Mendes and Dra. Fabiana Paula for their suggestions and example of professionalism.

To CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) - Finance Code 001, and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) - Project 142146/2017-0, for the doctoral scholarships.

Finally, I am very grateful for my family, Luiz Alexandre Simões de Castro (husband) and Marina Chaves Simões de Castro (daughter) that always support me, since we moved to Piracicaba in 2016.

ABSTRACT

De CHAVES, M. G. **Multianalytical approach in the study of the composition and functionality of the microbiota associated with carbon and nitrogen cycles in soils under preserved and disturbed environments.** 2020. 125 p. Tese (Doutorado) – Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2020.

The implementation of different agricultural management practices and land use systems promote physicochemical changes in soil associated with loss of nitrogen (N) and carbon (C) from soil and greenhouse gases (GHG) emission. In Brazil, the expansion of sugarcane production and land use changes can affect microbial groups that play essential roles in biogeochemical cycles of N and C. This thesis focused on the microbial community structure, composition, and functions associated with the N and C cycling on sugarcane-cultivated soil and soils under different land use in the Brazilian Amazon. Using a culture-independent approach based on high-throughput DNA sequencing and microarray technology, this thesis is composed by three chapters: the first and the second chapters encompass studies applying GeoChip v.5.0M microarray technology to investigate N and C functional genes as potential bioindicators of vinasse organic residue in combination with N mineral fertilizer in archaeal and acidobacterial communities inhabiting sugarcane-cultivated soil. In the third chapter, the co-occurrence of archaeal taxonomic groups was evaluated in primary and secondary forests, agricultural and cattle pasture soils in the Brazilian Amazon based on high-throughput amplicon sequencing and qPCR of the 16S rRNA gene. Our results in sugarcane-cultivated soil revealed that 87% of genes families associated with N metabolism from soil microbiota were responsive to vinasse with N fertilizer in 7 first days after application. The main gene families responsives were related to processes of nitrification (*amo*A and *hao*), ammonification (*gdh* and *ure*C), and denitrification (*p450nor*). The potential O_2 decrease and the increase of K and P due vinasse addition can promote the growth of halophile Archaea (*Natronomonas*) and N_2O reducing bacteria (*Anaeromyxobacter*), but also decrease ammonia oxidizer bacteria (AOB). Regarding the Acidobacteria, subgroups Gp13 and Gp18 revealed positive correlations with the C gene families associated with degradation, especially hemicellulose, but low abundance in vinasse presence. On the other hand, Gp4 was the most abundant acidobacterial subgroup in the vinasse treatment but was not associated with C gene families. This soil management practice can reduce the total Acidobacteria abundance, including that potentially involved with C degradation in sugarcane crops. The co-occurrence of archaeal classes analysis revealed that forest to pasture or agriculture conversion in Amazonian may reduce the "syntropy" between groups, an important strategy from Archaea to get energy and can promote the reduction of "key" groups related to N metabolism as ammonia oxidizer Archaea (AOA). These findings of the microbial functionality associated with N and C cycle in tropical soils can support public policy to mitigate N losses in sustainable agroecosystems.

Keywords: Archaea. Acidobacteria. Amazon rain forest. GeoChip v.5.0M. High-throughput DNA sequencing. Vinasse. Nitrogen fertilizer. Carbon degradation. Co-occurence. Soil.

RESUMO

De CHAVES, M. G. **Abordagem multianalítica no estudo da composição e funcionalidade da microbiota associada aos ciclos do carbono e nitrogênio em solos sob ambientes preservados e perturbados**. 2020. 125 p. Tese (Doutorado) – Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2020.

A implantação de diferentes sistemas de uso da terra e de práticas de manejo na agricultura promovem mudanças físico-químicas no solo associadas com perdas de nitrogênio (N) e carbono (C) provenientes da emissão de gases de efeito estufa (GEE). No Brasil, a expansão da produção da cana-de-açúcar e dos diferentes usos da terra podem afetar grupos microbianos que atuam nos ciclos biogeoquímicos do N e do C. Essa tese aborda a estrutura, composição e funções de comunidades microbiana associadas com os ciclos do N e do C em solos sob cultivo de cana-de-açúcar e sob diferentes usos da terra na Amazônia brasileira. Usando metodologia independente de cultivo, com base em sequenciamento de DNA de alto rendimento e tecnologia de microarranjo de DNA, essa tese está organizada em três capítulos: no primeiro e no segundo capítulos foi utilizada a versão de microarranjo GeoChip 5.0.M para investigar genes funcionais associados ao metabolismo de N e C como potenciais bioindicadores do uso combinado de vinhaça, um subproduto da produção de etanol, em combinação com fertilizante N mineral em comunidades de Archaea e de Acidobacteria que habitam o solo cultivado com cana-de-açúcar. No terceiro capítulo, foi avaliada a co-ocorrência entre grupos de Archaea em florestas primária e secundária, em solos agrícolas e em pastagens na Amazônia brasileira utilizando a técnica de PCR quantitativo em tempo real (qPCR) e sequenciamento de amplicon do gene 16S rRNA. Nossos resultados com cultivo de cana-de-açúcar sob adição de vinhaça e do fertilizante nitrogenado revelaram que 87% das famílias gênicas associadas com o metabolismo de N provenientes da microbiota do solo foram responsivas aos 7 dias após a aplicação. As principais famílias gênicas responsivas estão relacionadas aos processos de nitrificação (*amo*A e *hao*), amonificação (*gdh* e *ure*C) e denitrificação (*p450nor*). A possível redução na disponibilidade de O_2 e o aumento de K e P no solo devido a adição da vinhaça podem promover o crescimento de Archaea halophilas (*Natronomonas*) e bactérias redutoras de N_2O (*Anaeromyxobacter*), e reduzir AOB (bactéria oxidadora de amônia). Os subgrupos Gp13 e Gp18 de Acidobacteria foram correlacionados com famílias gênicas relacionadas com a degradação de C, principalmente hemicelulose. Entretanto, apresentaram baixa abundância no tratamento com vinhaça. O subgrupo Gp4 foi o mais abundante no tratamento com vinhaça, mas não foi correlacionado às famílias gênicas de C. Essa prática de manejo do solo pode reduzir a abundância de Acidobacteria total, incluindo aquelas potencialmente envolvidas com a degradação de C em culturas de cana-de-açúcar. Os estudos realizados nessa tese contribuem com informações sobre a dinâmica funcional da microbiota ativa nos ciclos biogeoquímicos do N e C, orientando políticas públicas que visam a redução de perdas de N e C em agroecossistemas.

Palavras-chave: Archaea. Acidobacteria. Amazônia. GeoChip v.5.0M. Seqüenciamento de DNA de alto rendimento. Vinhaça. Fertilizante nitrogenado. Degradação de carbono. Co-ocorrência. Solo.

SUMMARY

1. INTRODUCTION

The soils play an essential role in the control of storage, transformation, and flow of nutrients through the lithosphere and biosphere because they are the main terrestrial reservoir of organic carbon (C) and nitrogen (N) (QUINTON et al., 2010; BATJES, 1996; BERNOUX et al., 2002). The impact on global climate changes involves these biogeochemical cycles because soil management is still not well known (DAVIDSON; JANSSENS, 2006; HEIMANN; REICHSTEIN, 2008; SULEIMAN et al, 2018, NAVARRETE et al., 2015a).

In the C cycle, this element goes through the layers of Earth (biosphere, pedosphere, and lithosphere) and is fixed from the atmosphere by photosynthetic plants, protists, cyanobacteria and bacteria with bacterioclorofila. The C stored in vegetable biomass is part of gross primary production (GPP), which is the amount of C that can be retained from the atmosphere. The GPP rules most of the C flux between the biosphere and atmosphere, being the base of food and wood production (XIAO et al., 2019). Although part of the C absorbed through photosynthetic life returns naturally to the atmosphere by respiration, a large part stays gathered as soil organic C (SOC). The litter (remains from plants that fall on the ground), root exudation, and C from death body decomposition, determine the SOC size (CAO et al., 2019). Greater SOC in soil means more moisture and nutrient retention to plants and microbiota success. Processes such as deforestation, harvest, and fire can lead to C losses because they transfer C directly from plants to the atmosphere, skipping steps that should occur on the forest ground through organisms. As a result, critical damage to the ecosystem follow, which interfere in the cycle dynamic reducing or increasing the C stocks (XIAO et al., 2019; JUNGKUNST; FIEDLER, 2007), like as the use of residue from sugarcane production in Brazil (GALDOS; CERRI; CERRI, 2009; ZANI et al., 2018). Unfortunately, authors have demonstrated that carbon dioxide (CO_2) and methane (CH_4) global fluxes increase because of intensive agriculture, logging, and cattle pasture implantation (BAHN et al., 2010; SCHULZE et al., 2009; SCAVINO et al., 2013).

The N is an essential chemical element to build mainly nucleic acid and proteins in cells, and the lack of it is a limiting factor to the growth of organisms. Although it may be found widely as N_2 gas in the atmosphere, few bacteria and Archaea can fix it and make it available in reactive forms as ammonia (NH_3^-) or nitrate (NO_3) (KUYPERS; MARCHANT; KARTAL, 2018). The availability of these substrates depends on the metabolic diversity present in the prokaryotic cells reactions. These reactions transform less reactive NH_3^- into more reactive forms as NO_3^- and (nitrite) NO_2^-. They integrate the large network of processes and biochemical

pathways that constitute the N cycling on Earth and are important to technological applications and the environment (KUYPERS; MARCHANT; KARTAL, 2018). In the N cycle, different processes happen in an orderly manner. First, an N_2 molecule is reduced to ammonia (NH_3^-) in a process named N fixation. Then, this compound is oxidized to nitrate (NO_3^-) through the nitrification process ($NH_3^- \rightarrow NO_2^- \rightarrow NO_3^-$) in two steps. In the first process, when ammonia oxidation occurs, NH_3^- is oxidized to nitrite (NO_2^-) ($NH_3^- \rightarrow NO_2^-$). This process is performed by ubiquitous soil microorganisms as Archaea ammonia oxidizer (AOA), and Bacteria ammonia oxidizer (AOB) (LEHTORVITA-MORLEY, 2018; PJEVAC et al., 2017; KONNEKE et al., 2005). There are two intermediate compounds before NO_2^- formation: hydroxylamine (NH_2OH) and nitric oxide (NO). The NH_2OH is the result of ammonia oxidized by ammonia monooxygenase (AMO) enzyme. NO is the result of NH_2OH oxidized by hydroxylamine dehydrogenase or oxidoreductase (HAO) enzyme. AMO belongs to an ammonia superfamily with methane and alkane monooxygenases (LEHTORVINA-MORLEY, 2018). In the end of the nitrification process, the NO_3 produced may return to N_2 molecule by a "denitrification" process ($NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2$) or anaerobic ammonium oxidation ("anammox" process: $NO_2^- + NH_4^- \rightarrow N_2$). In both processes, "nitrification" and "denitrification", the production of a powerful greenhouse gas - nitrous oxide (N_2O) - may occur.

The N fixation process can be affected by soil deficiency of N nutrient (TANG et al., 2019). As an alternative, N fertilizer from fossil fuels is widely used to improve global food production (PEOPLES et al., 2019). However, the plants roots do not take all N applied to the soil (PEOPLES et al., 2019). Consequently, the large amount of fertilizer used by the human population in agriculture production generates twice as much N available in terrestrial ecosystems (GALLOWAY et al., 2008). Because much of N is leached or lost as environmentally harmful gas emissions (PEOPLES et al., 2019), the main topic of recent studies in microbial ecology has been the N_2O gas (HINK, 2017; SULEIMAN et al., 2018; NAVARRETE et al., 2015a; LOURENÇO et al., 2018).

Variations in intrinsic soil properties can result from removal of crops (QUINTON et al., 2010; NIETZSCHE et al., 2016; MENDES et al., 2015a) or addition of organic residue (de CHAVES et al., 2019; NAVARRETE et al., 2015a; SULEIMAN et al., 2018) as well as the application of chemical fertilizer (LIN et al., 2019; LI et al., 2012; SHARMA et al., 2014; SILVA et al., 2016). Through slash-and-burn technique, the forest to pasture conversion in Amazon rainforest promotes changes in pH and organic matter of the soil leading to CO_2, N_2O and CH_4 emissions by soil microbial communities (SULEIMAN et al., 2018; MENDES et al.,

2015a; BERNOUX et al., 1998; STEUDLER et al., 1996). Studies using culture independent-technique have shown not only microbial diversity losses in tropical forest converted to pastures (JESUS et al., 2009; RODRIGUES et al., 2013; PAULA et al., 2014) but also functional changes associated with biogeochemical cycles (PEDRINHO et al., 2019). Some "key" groups such as Acidobacteria and Archaea ammonia oxidizers (AOA) were suggested as a management-indicator to discriminate between sustainable *versus* non-sustainable agricultural practices or forest-to-pasture conversion (NAVARRETE et al., 2011; 2015a; FIGUEROLA et al., 2012). Because of their wide distribution, easy detection, and essential functions in soil, these indicators can be used for quality monitoring, which may be the secret to prevent the loss of functionality in the agroecosystem (SCHLOTER et al., 2018). Beyond these groups, specific genes are associated with the transformation of C, N, sulfur (S) and phosphorous (P) (SCHLOTER et al., 2018; DINI-ANDREOTE et al., 2018) can also be used. Genes *nif*H (N fixation), *amo*A (N nitrification), *nor*B/C, and *nir*K/S (N denitrification) have been proposed for evaluating N cycling functions in soil (ANAND, 2012; PROSSER, 2012; PHILLIPOT, 2002). These genes can be monitored by molecular techniques such as microarray, real time quantitative polymerase chain reaction (RT-qPCR), high-throughput sequencing, and emerging technologies (SCHLOTER et al., 2018; CLOUTIER et al., 2019) to better understand physiological mechanisms in soil microbiota (LEHTOVIRTA-MORLEY, 2018).

The access and development to new techniques based on genomic information (ROSSELLÓ-MÓRA; WHITMAN, 2019; ROESCH et al., 2007; WOODCROFT et al., 2018) as well as software and bioinformatic tools for data analysis (EDGAR, 2013; WU et al., 2016; LEMOS et al., 2017), have allowed to overcome the challenges of studies involving microbial taxonomic and functionality in complex environments such as the soil. This thesis sought to evaluate the taxonomic composition and the functionality of soil microbial communities associated with N and C cycling. Therefore, these studies are driven by the following questions: "Who is in this soil?", "How many?" and "What are they doing?". Simpson (1961) defines taxonomy as "the theory study of organisms classification or ordination in groups based on your morphological similarities". The microorganisms characterization nowadays is not limited anymore by the difficulties associated to isolation and cultivation techniques. By using techniques of environmental DNA sequencing, especially the ones that use the ribosomal gene 16S rRNA and PCR (TEELING; GLOCKNER, 2012), has become possible to understand the exceptional degree of prokaryote diversity. Currently with complete genome sequencing of a microbial community from a soil sample is possible to infer the origin of genes and their metabolic functions (MAHATO et al., 2017). To discover which taxa a microorganism belongs

to allows us to predict its adaptations to the environment and biotechnological potential (ROMALDE; BALBOA; VENTOSA, 2019). The 16S rRNA amplicons sequencing study may quantify community composition while in the metagenomic analysis is also possible to associate microbial taxa with the soil processes (FIERER et al., 2012). The metagenomic analysis from the first sequencing generation used cloning coupled to Sanger sequencing method (SANGER; NICKLEN; COULSON, 1977), a difficulty that was overcome through new generation sequencing (NGS), more specifically in the third generation. In the second generation, it was still necessary to prepare a DNA/RNA clone library, which is a long and expensive process (THOMPSON; MILOS, 2011). Besides, the NGS provides extremely high-throughput (in the range of gigabase - Gb) from multiple samples at the same time (MARDIS, 2011). This NGS quality provided a faster acquisition of results, decreasing the time necessary to generate Gb sequences. An example is the genome sequence of J. C. Venter, in Human Genome Project, that was finished in 15 years by Sanger (LEVY et al., 2007) and by NGS (454 Genome Sequencer FLX) the same Gb sequencing would be reduced to only 2 months (WHEELER et al., 2008). With the Illumina platform (https://www.illumina.com) better yields of reaction and decrease of prices happen, making NGS technique even more attractive (TEELING; GLOCKNER, 2012).

New bioinformatics tools have been developed to overcome challenges of storage, analysis, and interpretation of NGS data, also revolutionizing this field of knowledge (WANG; GERSTEIN; SNYDER, 2009; RABBINI; TEKIN; MAHDIEH, 2014; LEO et al., 2015; LAND et al., 2015; EL-METWALLY et al., 2013). There are at least four steps in sequence analysis of nucleotides from NGS platform (LAND et al., 2015; EL-METWALLY et al, 2013). In the first step, the ~~gross~~ raw data from the software integrated to the sequence analyzer are turned into short reads of nucleotide sequences (FASTq format in general). Quality score are necessary for this step to remove reads with low phred score levels, sequence errors, primers, vectors, tags, and tails that were introduced experimentally during the preparation of the sequencing libraries (HONG et al., 2013). The second step is the alignment and assembly of contigs and/or scaffolds. Annotation for functional detection and identification of genes, data integration and assemble sequence visualization is the third step. Finally, in fourth, all data is submitted to databases that contain tools for taxonomy and functionality identification of biological interest. The gross data from NGS technology may be put in open databases such as the NCBI Sequence Read Archive or MG-RAST (http://metagenomics.anl.gov) for metagenomic datasets (WILKE et al., 2016).

Besides NGS technology, the RT-qPCR or microarray can be used to characterize microbial communities (ARYA et al., 2005; De ANGELIS et al., 2015; HE et al., 2010a; TU et al., 2014; 2017; ZHOU et al., 2015). These high-throughput technologies have provided much information about new microbial processes and functions (VALDES; GLASS; SPECTOR, 2013). When comparing NGS technology with microarrays, there are differences in random sampling susceptibility and errors of specificity, quantification, and data analysis (ZHOU et al., 2015). Both have advantages and disadvantages in relation to specificity, sensibility, and quantification (ZHOU et al., 2015). However, if these techniques are used together (in a complementary way), they can answer important questions in microbial ecology (ZHOU et al., 2015). Microarray can be used to detect and quantify phylogenetic markers and functional genes (ZHOU et al., 2015). The GeoChip is a generic functional gene arrays (FGAs) that aim to detect hundreds of genic categories, being currently the most used method to analyze biogeochemical cycles genes (SHI et al., 2019). It is a quantitative technique capable of accessing very small amounts of nucleic acid to analyze the functional structure of microbial communities in environments such as soil (HE et al., 2010b; ZHOU et al, 2012; XUE et al., 2016a; 2016b), aquatic ecosystems (KIMES et al., 2010), extreme environments (MASSON et al., 2010), contaminated habitats (LIANG et al., 2011; XU et al., 2010; HE et al., 2018), and bioreactors (WU et al, 2017; ZANG et al., 2017). Technical problems related to this technology have been solved by implementing improvements in the new versions (SHI et al., 2019).

The GeoChip v.5.0 is the most representative generation of FGA, which has 1000 new functional gene families compared to the previous version, due to the rapid expansion of information in databases. This version was systematically evaluated for specificity, sensitivity, and quantitative capacity, and proved to be highly specific, sensitive, and quantitative to analyze the functional profile of microbial communities (SHI et al., 2019). It was synthesized *in situ* by Agilent's Sure Print Technology (SHI et al., 2019), which has a matrix with greater sensitivity and quantitative capacity when compared to other technologies (FULMER-SMENTEK et al., 2016). Computational and experimental evaluations indicated that designed probes are highly specific and can detect as little as 0.05 ng of pure culture DNA/1 g of community DNA (equivalent to 0.005 % of the population). To ensure reliability in the results of the analysis, chips have standard positive and negative probes for hybridization (SHI et al., 2019). Additionally, the hybridization temperature and the addition of formamide to the buffer were adjusted to 67 °C and 10 %, according to previous tests, adding specificity and efficiency, especially in environmental DNA samples (HE et al., 2010a; TU et al., 2014).

Two formats of GeoChip 5.0 were designed: a smaller, named GeoChip 5.0S, that cover genes of C, N, sulfur (S) cycle, and a larger, GeoChip 5.0M, for gene categories of C, N, S, and P cycles, and also several metals, response to stress, microbial defense, electron transport, growth promote of plant, virulence, *gyr*B, and specific fungi, protozoa, and virus genes. This last one has 161,961 oligonucleotide probes covering 365,000 genes belonging to 1,447 gene families. These gene families come from large and functionally divergent taxonomic groups, including Bacteria (2,721 genus), Archaea (101), Fungi (297), Protists (219), and Viruses (167). GeoChip v.5.0M is the most complete FGA at moment, linking genes or microbial populations to ecosystem functions. Changes in DNA concentration were also made in the 5.0M format, allowing the use of a bigger concentration of DNA (1000 ng) to improve coverage of detection (SHI et al., 2019). However, it is possible to use a lower concentration of DNA, as long as the same amounts are used for all samples. GeoChip technology can offer more advantages than other molecular techniques when-analyzing functional genes with low abundance in the community, such as *pmo*A (to methane oxidation) (PAULA et al., 2014). The 16S rRNA amplicon sequencing has shown to be inefficient for quantification of this gene in complex communities (PINTO; RASKIN, 2012; ZHOU et al., 2011), due to the difficulty in obtaining conserved sequences to design specific primers for RT-qPCR analysis. A comparative analysis between GeoChip and RT-qPCR with the same genes resulted in a high correlation between the number of copies and the signal intensity (HE et al., 2010b).

In the last years, different methods have been useful to study ecological aspects of microbial communities. However, there is a lack of integrative approaches based on molecular techniques in soil microbiology, and soil processes and properties as well as statistical analyses to understand the impacts of bioenergy cropping systems and land use changes in tropical soils. Here, we used an integrative approach based on metagenomics, amplicon sequencing, microarray, quantitative real-time PCR (RT-qPCR), soil physical and chemical analysis, and statistical methods for data analysis and integration. This integrative approach was applied to evaluate soil Archaea and Bacteria responses associated with C and N cycles in soil from preserved and disturbed environments in Brazil, with sugarcane crops land use and other different systems. Our hypotheses were defined based on “gaps” left by previous studies in tropical soil that showed that applying sugarcane organic residue and N fertilizer alter soil physicochemical properties, soil bacterial community abundance and composition associated with CO_2 and N_2O emissions from soil (NAVARRETE et al., 2015a). According to Navarrete et al. (2015a), Acidobacteria is one of the most responsive bacterial groups to this agricultural management in soil from sugarcane production fields. However, these studies in sugarcane-

cultivated soils did not investigated the microbial functionality associated with N metabolism neither revealed information about Acidobacteria capacities of C degradation under N fertilization combined with organic residue. In the Western Brazilian Amazon soils, previous studies demonstrated that physicochemical characteristics in soils from different land use systems shape the bacterial community structure (JESUS et al., 2009) and the AOA diversity (NAVARRETE et al., 2011). However, these studies in Amazonian soil did not investigate the archaeal community in a deep resolution using an integrative approach. In this sense, this thesis has a goal to fill the "gaps" of these previous studies that were performed in the same area and soil conditions, using a non-integrative analytical approach.

1.1. Hypothesis

1.1.1. General hypothesis

Agricultural management of the soil with vinasse combined to mineral N fertilizer in sugarcane-cultivated soil and, changes in the way this land has been used, lead to changes in in the Amazonian tropical soil properties which affect the abundance of specific soil microbial groups related to C and N cycles.

1.1.2. Specific hypothesis

Chapter 1. Responses of Archaea and Bacteria N-cycling genes and functional processes in sugarcane soils under vinasse and nitrogen fertilization as indicated by metagenome and GeoChip v.5.0M.

The organic fertilization combined with mineral N addition to sugarcane-cultivated soil can promote changes in soil properties that may affect the abundance of specific functions from archaeal and bacterial groups related to N metabolism.

Chapter 2. Acidobacteria subgroups and their metabolic potential for carbon degradation in sugarcane soil amended with vinasse and nitrogen fertilizers.

This agriculture management can promote changes in functions associated with carbon-degrading metabolism from specific classes of Acidobacteria.

Chapter 3. Ecological co-occurrence and soil physicochemical factors drive the archaeal community in Amazonian soils

Due to physicochemical factors changes in soils from primary forest to establishment of pasture and agriculture in the Amazon rain forest, the structure of archaeal community associated to biogeochemical cycles also changes.

1.2. Objectives

1.2.1. General objective

To study the composition and functionality of Archaea and Bacteria associated with biogeochemical cycles of C and N in tropical soils under preserved and disturbed environments using an integrative analytical approach to fill the "gaps" in previous studies that were performed in the same area and soil conditions.

1.2.2. Specific objectives

Chapter 1.

I. To determine which gene families and processes associated with N metabolism could be changing in abundance in soil archaeal and bacterial groups under vinasse and mineral N fertilizer.

II. To determine what chemical soil factors can promote changes in archaeal and bacterial groups associate with N cycle under vinasse combined with urea agricultural management.

Chapter 2.

I. To determine which classes of Acidobacteria and which carbon degradation gene families from this phylum can be changed under vinasse and mineral N fertilizer.

II. To determine which chemical soil factors can promote changes in acidobacterial classes associated with C degradation metabolism in sugarcane soil fertilizer.

Chapter 3.

I. To determine what is the main factor which drives archaeal community to assemble in Amazon soils under different types of land use: co-occurrence of Archaea classes or their interactions with soils factors.

II. To determine what soil factor mainly drives archaeal community in Amazon lands.

1.3. Structure of the thesis

This thesis comprises an introduction and three chapters. The three chapters are presented as a scientific manuscript written in the English language. The supplementary materials by chapters 2 and 3 are available in the Appendix section.

References

ANAND, R. Biological nitrogen fixation. In: HUANG PM, SUMMER, L. Y. **Handbook of soil sciences, properties, and processes.** Boca Raton: CRC Press, 2012. p 27.1-27.7.

ARYA, M. et al. Basic principles of real-time quantitative PCR. **Expert Review of Molecular Diagnostics**, Salt Lake, v. 5, p. 209 –219. 2005. doi: 10.1586/14737159.5.2.209.

BAHN, M. et al. Soil respiration at mean annual temperature predicts annual total across vegetation types and biomes. **Biogeoscience,** Göttingen, v.7, p. 2147-2157. 2010.

BATJES, N.H. Total carbon and nitrogen in the soils of the world. **European Journal of Soil Science**, Amsterdam, v. 65, p. 4–21, 1996. doi: 10.1111/j.1365-2389.1996.tb01386.x.

BERNOUX, M. et al. Bulk densities of Brazilian Amazon soils related to other soil properties. **Soil Science Society of America Journal**, Madison, v. 62, n.3, p. 743-749, 1998.

BERNOUX, M. et al. Brazil's soil carbon stocks. **Soil Science Society of America Journal**, Madison, v. 66, n.3, p. 888-896, 2002.

CAO, J. et al. Leaf litter contributes more to soil organic carbon than fine roots in two 10-year-old subtropical plantations. **Science of the Total Environment,** Amsterdam, v. 704, p. 135341, 2019. doi: 10.1016/j.scitotenv.2019.135341.

CLOUTIER, M.L. et al. Innovative technologies can improve understanding of microbial nitrogen dynamics in agricultural soils. **Agricultural and Environmental Letters**, Toronto, v.4, p. 190032, 2019. doi.org/10.2134/ael2019.08.0032.

COTTIER, F. et al. Advantages of meta-total RNA sequencing (MeTRS) over shotgun metagenomics and amplicon-based sequencing in the profiling of complex microbial communities. **NPJ Biofilms and Microbiomes**, London, v. 4, p. 2. 2018. doi: 10.1038/s41522-017-0046-x.

DAVIDSON, E.A.; JANSSENS, I.A. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. **Nature,** London, v. 440, p. 165–173, 2006. doi.org/10.1038/nature04514.

DE ANGELIS, K. M. et al. Long-term forest soil warming alters microbial communities in temperate forest soils. **Frontiers in Microbiology**, Lausanne, v. 6, p. 104. 2015. doi: 10.3389/fmicb.2015.00104.

DE CHAVES, M.G. et al. Acidobacteria subgroups and their metabolic potential for carbon degradation in sugarcane soil amended with vinasse and nitrogen fertilizers. **Frontiers in Microbiology,** Lausanne, v. 10, p. 1680, 2019. doi: 10.3389/fmicb.2019.01680.

DINI-ANDREOTE, F. et al. Dispersal-competition tradeoff in microbiomes in the quest for land colonization. **Scientific Reports,** London, v. 8, p. 9451, 2018. doi: 10.1038/s41598-018-27783-6.

EDGAR, R. UPARSE: highly accurate OTU sequences from microbial amplicon reads. **Nature Methods**, London, v. 10, p. 996-998, 2013. doi.org/10.1038/nmeth.2604.

EL-METWALLY, S. et al. Next-generation sequence assembly: four stages of data processing and computational challenges. **PLoS Computational Biology,** San Francisco, v. 9, p. e1003345. 2013. doi: 10.1007/s10142-015-0433-4.

FIERER, N. et al. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. **Proceedings of the National Academy of Science of the USA**, Washington, DC, v. 109, n. 52, p. 21390–21395, 2012.

FIGUEROLA, E. L. M. et al. Bacterial indicator of agricultural management for soil under no-till crop production. **PLoS ONE,** San Francisco, v. 7, n. 11, p. e51075, 2012.

FULMER-SMENTEK, S. et al. Assessment of the sensitivity and dynamic range of new generation microarray technology using the MAQC samples. Poster. Agilent Technologies, Santa Clara, CA. 2016. https://www.agilent.com/cs/library/posters/public/Chips2HitsvFINAL.pdf.

GALDOS, M.V.; CERRI, C.C.; CERRI, C.E.P. Soil carbon stocks under burned and unburned sugarcane in Brazil. **Geoderma**, Amsterdam, v.153, p. 347-352, 2009. doi: 10.1016/j.geoderma.2009.08.025.

GALLOWAY, J. N. et al. Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. **Science,** New York, v. 320, n. 5878, p. 889–892, 2008.

HE, Z. et al. GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. **The ISME Journal**. London. v. 4, p. 1167–1179. 2010a. doi: 10.1038/ismej.2010.46.

HE, Z. et al. Metagenomic analysis reveals a marked divergence in the structure of belowground microbial communities at elevated CO_2. **Ecology Letters**, New Jersey, v. 13, p. 564 –575. 2010b. doi.org/10.1111/j.1461-0248.2010.01453.x.

HE, Z. et al. Microbial functional gene diversity predict groundwater contamination and ecosystem functioning. **MBio,** Washington, DC, v.9, p. e02435-17. 2018. doi.org/10.1128/mBio.02435-17.

HEIMANN, M.; REICHSTEIN, M. Terrestrial ecosystem carbon dynamics and climate feedbacks. **Nature**, London, v. 451, p. 289–292, 2008. doi.org/10.1038/nature06591.

HINK, L.; NICOL, G.W., PROSSER, J.I. Archaea produce lower yields of N_2O than bacteria during aerobic ammonia oxidation in soil. **Environmental Microbiology**, Baltimore, v. 19, p 4829–37, 2017. doi.org/10.1111/1462-2920.13282.

HONG, H.X. et al. Critical role of bioinformatics in translating huge amounts of next-generation sequencing data into personalized medicine. **Science China Life Sciences**, London, v. 56, p. 110–8, 2013. doi: 10.1007/s11427-013-4439-7.

JESUS, E. et al. Changes in land use alter the structure of bacterial communities in Western Amazon soils. **The ISME Journal**, London, v.3, p. 1004 –11, 2009. doi:10.1038/ismej.2009.47.

JUNGKUNST, H.F., FIEDLER S. Latitudinal differentiated water table control of carbon dioxide, methane and nitrous oxide fluxes from hydromorphic soils: feedbacks to climate change. **Global Change Biology**, New Jersey, v.13 pp. 2668-2683. 2007.

KIMES, N.E. et al. Microbial functional structure of Montastraea faveolata, an important Caribbean reef building coral, differs between healthy and yellow band diseased colonies. **Environmental Microbiology**, Oxford, v.12, p. 541–556. 2010. doi.org/10.1111/j.1462-2920.2009.02113.x.

KONNEKE, M. et al. Isolation of an autotrophic ammonia-oxidizing marine archaeon. **Nature,** London, v. 437, p. 543–546, 2005. doi.org/10.1038/nature03911.

KUYPERS, M., MARCHANT, H., KARTAL, B. The microbial nitrogen-cycling network. **Nature Review Microbiology,** London, v. 16, p. 263–276, 2018. doi.org/10.1038/nrmicro.2018.9.

LAND, M. et al. Insights from 20 years of bacterial genome sequencing. **Functional and Integrative Genomics**, London, v.15, p.141–61. 2015. doi: 10.1007/s10142-015-0433-4.

LEHTOVIRTA-MORLEY, L. E. Mini Review – Environmental Microbiology Ammonia oxidation: Ecology, physiology, biochemistry and why they must all come together. **FEMS Microbiology Letters,** Amsterdam, v. 365, n. 9, p. fny058, 2018.

LEMOS, L. N. et al. Bioinformatics for microbiome research: concepts, strategies, and advances. In: PYLRO, V.; ROESCH, L. **The Brazilian microbiome**. New York: Springer, 2017. p. 111-123.

LEO, V.C. et al. Use of next-generation sequencing and candidate gene analysis to identify underlying defects in patients with inherited platelet function disorders. **Journal of Thrombosis and Haemostasis,** New York, v. 13, p. 643–50. 2015. doi: 10.1111/jth.12836.

LEVY, S. et al. The diploid genome sequence of an individual human. **PLOS Biology**, v.5, n. 10, p. e254. 2007. doi.org/10.1371/journal.pbio.0050254.

LI, X. H. et al. Soil chemical and biological properties affected by 21-year application of composted manure with chemical fertilizers in a Chinese Mollisol. **Canadian Journal of Soil Science**. Otawa. v. 92, n. 3, p. 419-428. 2012.

LIANG, Y. et al. Functional gene diversity of soil microbial communities from five oil-contaminated fields in China. **The ISME Journal**, London, v.5, p. 403– 413. 2011. doi.org/10.1038/ismej.2010.142.

LIN, W. et al. The effects of chemical and organic fertilizer usage on rhizosphere soil in tea orchards. **PLoS ONE**, San Francisco, v.14, n. 5, p. e0217018. 2019.

LOURENÇO, K. S. et al. Dominance of bacterial ammonium oxidizers and fungal denitrify in the complex nitrogen cycle pathways related to nitrous oxide emission. **Global Change Biology Bioenergy**, Orlando, v. 10, p. 645–660. 2018. doi: 10.1111/gcbb.12519

MAHATO, N. K. et al. Microbial taxonomy in the era of OMICS: application of DNA sequences, computational tools and techniques. **Antoine van Leeuwenhoek,** London, v. 110, n. 10, p. 1357–1371, 2017.

MARDIS, E.R. A decade's perspective on DNA sequencing technology. **Nature**, London, v. 470, p. 198–203, 2011. doi: 10.1038/nature09796.

MASON, O.U. First investigation of the microbiology of the deepest layer of ocean crust. **PLoS One,** San Francisco, v. 5, p. e15399. 2010. doi.org/10.1371/journal.pone.0015399.

MENDES, L. W. et al. Land-use system shapes soil bacterial communities in Southeastern Amazon region. **Applied Soil Ecology**, Amsterdam, v. 95, p. 151–60, 2015a. doi: 10.1016/j.apsoil.2015.06.005.

NAVARRETE, A.A et al. Land-use systems affect archaeal community structure and functional diversity in western Amazon soils. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 35, n. 5, p. 1527–1540, 2011.

NAVARRETE, A. A. et al. Multi-analytical approach reveals potential microbial indicators in soil for sugarcane model systems. **PLoS ONE,** São Francisco, v.10, 6, p. e0129765. 2015a. doi:10.1371/journal.pone.0129765.

NITZSCHE, K.N. et al. Visualizing land-use and management complexity within biogeochemical cycles of an agricultural landscape. Ecosphere, Washington, v. 7, p. e01282, 2016. doi.org/10.1002/ecs2.1282.

PAULA, F.S. et al. Land use change alter functional gene diversity, composition and abundance in Amazon forest soil microbial communities. **Molecular Ecology**, Amsterdam, v. 23, p. 2988 – 99, 2014. doi.org/10.1111/mec.12786.

PEDRINHO, A. et al. Forest-to-pasture conversion, and recovery based on the assessment of microbial communities in the eastern Amazon rainforest. **FEMS Microbiology Ecology,** Amsterdam, v. 95, n. 3, fiy236, 2019.

PEOPLES, M. et al. The contributions of legumes to reducing the environmental risk of agricultural production. In: Lemaire, G., Carvalho, P. C. F., Kronberg, S., Recous, S. **Agroecosystem diversity**. London: Elsevier, 2019.p. 123-143. doi.org/10.1016/B978-0-12-811050-8.00008-X.

PHILIPPOT, L. Denitrifying genes in bacterial and archaeal genomes. **Biochimica et Biophysica Acta,** Amsterdam, v. 1577, n. 3, p. 191–192, 2002.

PINTO, A.J, RASKIN, L. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. **PLoS One,** San Francisco, v.7, p. e43093. 2012. doi.org/10.1371/journal.pone.0043093.

PROSSER J. I. Nitrification. In: Huang PM, Li Y, Summer M. **Handbook of soil sciences, properties and processes. Soil Biology and Biochemistry: soil biology in its second age.** Boca Raton, CRC Press: pp 27.19– 27.32. 2012.

QUINTON, J. N. et al. The impact of agricultural soil erosion on biogeochemical cycling. **Nature Geoscience**, London, v. 3, n. 5, p. 311–314, 2010.

RABBINI, B., TEKIN M, MAHDIEH N. The promise of whole-exome sequencing in medical genetics. **Journal of Human Genetics,** London, v. 59, p. 5–15, 2014. doi: 10.1038/jhg.2013.114.

RODRIGUES, J. L. et al. Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 110, p. 988–93, 2013. doi: 10.1073/pnas.1220608110.

ROESCH, L. F. W. et al. Pyrosequencing enumerates and contrasts soil microbial diversity. **The ISME Journal**, London, v.1, n. 4, p. 283–290, 2007.

ROMALDE, J.L.; BALBOA, S.; VENTOSA, A. Editorial: Microbial taxonomy, phylogeny and biodiversity. **Frontiers in Microbiology**, Lausanne, v.10, p. 1324, 2019. doi: 10.3389/fmicb.2019.01324.

ROSSELLÓ-MÓRA, R., WHITMAN, W. B. Dialogue on the nomenclature and classification of prokaryotes. **Systematic and Applied Microbiology**, Amsterdam, v. 42, p. 5–14, 2019. doi.org/10.1016/j.syapm.2018.07.002.

SANGER, F.; NICKLEN, S.; COULSON, A.R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Science of the USA**, Washington, DC, v. 74, n. 12, p. 5463–5467.

SCAVINO, A. F. et al. Structure and function of the methanogenic microbial communities in Uruguayan soils shifted between pasture and irrigated rice fields. **Environmental Microbiology,** Washington, DC, v.15, n. 9, p. 2588–2602, 2013.

SCHLOTER, M. et al. Microbial indicators for soil quality. **Biology and Fertility of Soils**, Berlim, v. 54, p. 1–10, 2018. doi.org/10.1007/s00374-017-1248-3.

SCHULZE, E.D. et al. Importance of methane and nitrous oxide for Europës terrestrial greenhouse-gas balance. **Nature Geoscience,** London, v. 2, pp. 842-850. 2009.

SHARMA, U. et al. Effects of continuous use of chemical fertilizers and manure on soil fertility and productivity of maize–wheat under rainfed conditions of the western Himalayas. **Communications in Soil Science and Plant Analysis,** London. v**.** 45, n. 20, p. 2647-2659, 2014. doi: 10.1080/00103624.2014.941854.

SHI, Z. et al. Functional gene array-based ultrasensitive and quantitative detection of microbial populations in complex communities. **mSystems,** Washington DC, v. 4, e00296-19. 2019. doi.org/10.1128/mSystems.00296-19.

SILVA, D.J. et al. Organic and nitrogen fertilization of soil under 'Syrah' grapevine: effects on soil chemical properties and nitrate concentration. **Revista Brasileira de Ciência do Solo**. Viçosa, v. 40, p. e0150073, 2016.

SIMPSON, G. G. **Principles of Animal Taxonomy.** New York, Columbia Univ. Press, 1961, pp. 247.

STEIN, L. Y.; KLOTZ, M. G. The nitrogen cycle. **Current Biology**, Cambridge, v. 26, R94–R98, 2016.doi.org/10.1016/j.cub.2015.12.021.

STEUDLER, P.A. et al. Microbial controls of methane oxidation in temperate forest and agricultural soils. In: Murrell J.C., Kelly D.P. **Microbiology of Atmospheric Trace Gases**. Berlin: Springer, 1996. P. 69-84.

SULEIMAN, A. K. A. et al. Recycling organic residues in agriculture impacts soil-borne microbial community structure, function and N_2O emissions. **Science of the Total Environment,** Amsterdam, v. 63, p. 1089–1099, 2018. doi.org/10.1016/j.scitotenv.2018.03.116.

TANG, J. et al. Effects of soil nitrogen (N) deficiency on photosynthetic N-use efficiency in N-fixing and non-N-fixing tree seedlings in subtropical China. **Scientific Reports**, London, v.9, p. 4604, 2019. doi.org/10.1038/s41598-019-41035-1.

TEELING, H.; GLOCKNER, F. O. Current opportunities and challenges in Microbial metagenome analysis bioinformatics perspective. **Briefings in Bioinformatics**, Oxford, v. 13, n.6, p. 728 – 742, 2012.

THOMPSON, J.F.; MILOS, P.M. The properties and applications of single-molecule DNA sequencing. **Genome Biology**, Oxford, v.12, n. 2, p. 217, 2011.

TU, Q. et al. GeoChip 4: a functional gene array based high throughput environmental technology for microbial community analysis. **Molecular Ecology Resource**s, New Jersey, v. 14, p. 914 – 928. 2014. doi.org/10.1111/1755-0998.12239.

TU, Q. et al. HuMiChip2 for strain level identification and functional profiling of human microbiomes. **Applied Microbiology and Biotechnology**, New Jersey, v.101, p. 423. 2017. doi.org/10.1007/s00253-016-7910-0.

VALDES, A.M., GLASS, D., SPECTOR, T.D. Omics technologies and the study of human ageing. **Nature Review Genetics**, London, v. 14, p. 601– 607. 2013. doi.org/10.1038/ nrg3553.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, London, v.10, p. 57–63, 2009. doi: 10.1038/nrg2484.

WHEELER, D. A. et al: The complete genome of an individual by massively parallel DNA sequencing. **Nature**, London, v. 452, p. 872–6, 2008. doi: 10.1038/ nature06884.

WILKE, A. et al. The MG-RAST metagenomics database and portal in 2015. **Nucleic Acid Research**, Oxford, v. 44, p. D590–D594, 2016. doi: 10.1093/nar/gkv1322.

WOODCROFT, B. J. et al. Genome-centric view of carbon processing in thawing permafrost. **Nature**, London, v. 560, n. 7716, p. 49–54, 2018.

WU, Y. W.; SIMMONS, B. A.; SINGER, S. W. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. **Bioinformatics**, Oxford, v.32, n. 4, p. 605-607, 2016.

WU, L. Microbial functional trait of rRNA operon copy numbers increases with organic levels in anaerobic digesters. **The ISME Journal**, London, v.11, p. 2874. 2017. doi.org/10.1038/ismej.2017.135.

XIAO, J. et al. Remote sensing of the terrestrial carbon cycle: a review of advances over 50 years. **Remote Sensing of Environment,** Amsterdan, v.233, p. 111383, 2019. doi.org/10.1016/j.rse.2019.111383.

XU, M. et al. Responses of microbial community functional structures to pilot-scale uranium in situ bioremediation. **The ISME Journal**, London, v.4, p. 1060-1070. 2010. doi.org/10.1038/ismej.2010.31.

XUE, K. et al. Tundra soil carbon is vulnerable to rapid microbial decomposition under climate warming. **Nature Climate Change,** London, v.6, p. 595. 2016a. doi.org/10.1038/nclimate2940.

XUE, K. et al. Annual removal of aboveground plant biomass alters soil microbial responses to warming. **MBio**, Washington, DC, v.7, p. e00976-16. 2016b. doi.org/10.1128/mBio.00976-16.

ZANI, C.F. et al. Vinasse application and cessation of burning in sugarcane management can have positive impact on soil carbon stocks. **Peer J**., San Diego, v.6, e5398, 2018. doi.org/10.7717/peerj.5398.

ZHANG J. Bacteriophage–prokaryote dynamics and interaction within anaerobic digestion processes across time and space. **Microbiome**, London, v.5, p. 57. 2017. doi.org/10.1186/s40168-017-0272-8.

ZHOU, J.Z. et al. Reproducibility and quantitation of amplicon sequencing-based detection. **The ISME Journal,** London, v.5, p. 1303–1313. 2011. doi.org/10.1038/ismej.2011.11.

ZHOU, J. et al. Microbial mediation of carbon-cycle feedbacks to climate warming. **Nature Climate Change**, London, v.2, p. 106 –110. 2012. doi.org/10.1038/nclimate1331.

ZHOU, J. et al. High-throughput metagenomic technologies for complex microbial community analysis: open and closed formats. **MBio**, Washington, DC, v.6, p. e02288-14. 2015. doi.org/10.1128/mBio.02288-14.

2. Chapter 1. Responses of Archaea and Bacteria N-cycling genes and functional processes in sugarcane soils under vinasse and nitrogen fertilization as indicated by metagenome and GeoChip v.5.0M

Abstract

A common agricultural practice to combine organic fertilizer vinasse (a liquid residue of ethanol production from sugarcane) with mineral nitrogen (N) fertilizer may promote N losses such as greenhouse gases (GHG), due to the effects of physicochemical changes on soil microbiota. We applied microarray GeoChip v.5.0M technology to identify and quantify the N functional genes families associated with the N processes in treatments without N fertilizer (N0), with urea at 60 kg ha^{-1} (N60), and with vinasse combined with urea (NV). Soil samples were collected at 7 (T7) and 150 (T150) days after the beginning of the greenhouse experiment, based on the maximum and minimum nitrous oxide (N_2O) emissions obtained in previous studies, for molecular and physicochemical analysis. The metagenomes from these DNA samples, previously deposited in the MG-RAST server, were accessed to investigate the functions and taxonomic groups associated with selected gene families. The results revealed that 87% of the select gene families were significantly responsive to the fertilizer combined treatment (NV) in the first days after the application. The main genes and processes responsive were nitrification (with the *amo*A gene from ammonia-oxidizing Bacteria (AOB) and Archaea (AOA) and *hao* from Bacteria), ammonification (with *gdh* and *ure*C gene from Bacteria and Archaea), and denitrification (with *p450nor*). The AOA, *Nitrosopumillus*, and AOB, *Nitrosomonas*, were the groups with the greatest functions associated with nitrification, as well as a pathogenic *Mycobacterium*, with denitrification. The results also revealed that under N fertilizers and decreased O_2 in soil, the increases in K and P nutrients can promote the growth of the halophile Archaea *Natronomonas* and the Bacteria *Anaeromyxobacter*,which can reduce N_2O. In conclusion, this typical fertilizer agricultural management may favor the growth of archaeal and bacterial groups associated with N processes in tropical soil agroecosystems that have potential sustainable and biotechnological roles.

Keywords: AOA, AOB, ammonification, nitrification, denitrification, *Anaeromyxobacter*, *Natronomomas*, *p450nor*.

2.1. Introduction

Vinasse is a byproduct of ethanol from sugarcane industry, generated during the distillation process (Glória, 1992; Dias et al., 2009), with a production rate of 10 to 18L per liter (Gasparotto et al., 2014). The chemical composition of vinasse is generally 93% water and 7% organic materials and minerals (Hidalgo et al., 2009; Christofoletti et al., 2013). This residue has high levels of organic matter and a high C:N ratio (Moran-Salazar et al., 2016), but low levels of nitrogen (N) (0.97 to 4.75 g L^{-1}) and phosphorus (P) (1 to 190 mg L^{-1}). The main non-aqueous components in vinasse are organic matter such as glycerol, organic acids, and yeasts (Christofoletti et al., 2013), and the volatile solids are present at high levels (Moran-Salazar et al., 2016). When applied to the soil, the amounts of organic material, nitrate (NO_3^-), ammonium (NH_4^+), potassium (K^+), calcium (Ca^{+2}), magnesium (Mg^{+2}), sodium (Na^+), and metals increase (Christofoletti et al., 2013; Moran-Salazar et al., 2016). An acidic pH (3.9 to 5.1), high chemical oxygen demand (COD) (50.000 to 95.000 mg L^{-1}), and biological oxygen demand (DBO) (18.900 to 78.300 mg.L^{-1}) occur (Moran-Salazar et al., 2016). If vinasse is poured into the environment in an unplanned way, it can cause pollution in the soil surface and underground due to the excessive accumulation of salts, which results in a toxic condition (Moran-Salazar et al., 2016).

Vinasse application as a liquid fertilizer can be applied with mineral N sources in the sugarcane crops to minimize the ecological problem of its residue disposal into the environment (Moraes et al., 2014; Penatti et al., 1988). However, even though this practice increases sugarcane productivity, it also causes physical, chemical, and biochemical changes in the soil environment (Madejón et al., 2001; Tejada and Gonzalez, 2006). There is a lack of information about the impacts of this agricultural management on the structure and functionality of soil microbial communities in tropical soils, but it is known that their changes are often correlated with chemical factors (Nilsson et al., 2007; Lauber et al., 2009; Jenkins et al., 2009). This has been demonstrated for Bacteria (Jesus et al., 2009; Faoro et al., 2010; Navarrete et al., 2015), but not for Archaea.

Studies have shown that the addition of vegetable organic residues commonly used in the fallow season such as vinasse or sugarcane straw (Wang et al., 2015), increase the emissions of nitrous oxide (N_2O), a greenhouse gas (GHG) (Wang et al., 2015; Lourenço et al., 2018), especially when the vinasse is combined with N addition (Wang et al., 2015; Lourenço et al., 2018). Lourenço et al. (2018) concluded that the *amo*A from ammonia-oxidizing Bacteria (AOB), active in the nitrification process, was the most functional gene related to N_2O

emissions in sugarcane crops under vinasse fertilization. Regarding the ammonia-oxidizing Archaea (AOA), which also have the *amo*A gene (Lehtovirta-Morley, 2018), some studies have demonstrated its relevant role in the N biogeochemical cycle (Tourna et al., 2011; Sauder et al., 2017). However, experts warn about the imminent need to associate genomic data from molecular methods in Microbial Ecology with biochemical data, to close the "gap" between physiological mechanisms by AOA and N metabolism (Lehtovirta-Morley, 2018).

Cloutier et al. (2019) described the parallel quantification of N functional genes such as one emerging technology that could be combined with high-throughput-sequencing (HTS) of microbial DNA to study N metabolism in soil. The microarray technology GeoChip v.5.0M (Agilent Technologies Inc., Santa Clara, CA, United States) has become a powerful and high-performance tool to monitor environmental processes with high-sensitivity detection of microbial functional genes (Varon-Lopez et al., 2014; Cong et al., 2015; Shi et al., 2019; Zhao et al., 2020). The chip contains 167,044 distinct probes belonging to different categories of genes involved in carbon (C) and N (ammonification, nitrification, fixation, among others) metabolism (Shi et al., 2019). The probes were originated from gene sequences of microorganisms correspond to a unique sequence of a gene family (Ex: *ace*A) and encode the same class of proteins. Several studies have concluded that this technique can be used mainly to link data from microbial communities to the processes and functions to which they are related (Yu and Laber, 2015) in a broad sense. Although recent studies have been able to obtain relevant information about the N metabolism genes in soil microbiota with molecular techniques (Smith et al., 2015; Lourenço et al., 2018), GeoChip v.5.0M technology offers the advantage of analyzing many functional genes at once, providing a broader view of the functional response of the microbiota under a given condition (Carter et al., 2012).

Under the assumption that changes in soil properties, due to the organic fertilization combined with N mineral, significantly affect the abundance of functions from Archaea and Bacteria related to N metabolism, we aim to assess gene families, processes, and groups with a significant response to these environmental conditions. Since the microbial N metabolism in soils is associated with N losses, such as GHGs (Wang et al., 2015), better understanding the functionality of Archaea and Bacteria can help mitigate this economic and environmental problem.

2.2. Material and Methods

2.2.1. Experimental design, soil treatments, and soil sampling

This study was carried out from the implementation of an experiment in a greenhouse, in which sugarcane seedlings were grown over 150 days between April and December 2013 in nine 100 L pots, as described by Navarrete et al. (2015). The sugarcane seedlings of *Saccharum* spp. CTC-02 were obtained from in vitro fertilization and tissue culture. The red clayey podzolic soil was obtained from the 0-20 cm layer in the Areão farm, belonging to the Luiz de Queiroz School of Agriculture, University of São Paulo (ESALQ-USP), in Piracicaba, São Paulo, Brazil (22° 42' 30" S e 47° 38' 00" W). All the pots received, in addition to the sugarcane seedlings, 90 kg of soil, and basic fertilization (150 kg ha-1 of triple P_2O_5 superphosphate and 80 kg ha-1 of potassium chloride (KCl)). Three treatments were established containing three replicates each: the control treatment (N0) received only the basic fertilization, without a nitrogen fertilizer; the second treatment (N60) received, in addition to the basic fertilization, inorganic (urea) nitrogen fertilizer (at a dose of 60 kg ha-1), which was immediately mixed with the soil to avoid losses due to volatilization; the third treatment (NV) received urea (60 kg ha^{-1}) supplemented with vinasse (at a dose of 0.06 L kg^{-1} (120 m^3 ha^{-1}). The K content of the KCl added in the NV treatment was calculated excluding the amount already present in the vinasse (defined above), in order to introduce into the system the same amount of this nutrient in all the pots. Water was added in the same volume as the vinasse, and in the treatments that did not receive this organic residue, the soil moisture content in the pots was maintained at 80% of the field capacity throughout the experiment using a humidity sensor (Extech MO750, Nashua, NH, United States).

The soil samples were collected at 7 and 150 days after the addition of the nitrogen fertilizers and vinasse to perform molecular analyses and for the determination of chemical factors. These collection periods were determined according to the maximum and minimum carbon dioxide (CO_2) and N_2O emissions from the soil, as shown in Navarrete et al. (2015). Three bulk soil samples (approximately 100 g in each one) were collected from the topsoil (0 to 10 cm) in equidistant positions within an equilateral triangle (radius equivalent to 1/3 of the circular surface of the pot) and pooled together, forming only one sample with 300 g. For molecular analysis, soil subsamples were separated from the sample with 300 g obtained from each mesocosm. The subsamples with 250 mg were stored at -20 °C until processing

within 72 h at the Cell and Molecular Biology Laboratory from the Center for Nuclear Energy in Agriculture (CENA-USP).

2.2.2. Analysis of Soil Chemical Factors

Several soil chemical properties were analyzed according to the methodology used by the Brazilian Agricultural Research Corporation (2011). After the soil samples were sieved, in order to make a homogeneous mixture, a LECO CN 2000 elementary analyzer (Perkin Elmer, Waltham, MA, United States) was used to determine the C and total N by dry combustion. The pH was determined using a $CaCl_2$ 0.01 M solution. The content of B (boron) was obtained by extraction with hot water, while Al (aluminum), Ca, and Mg were extracted with KCl 1 M solution. After extraction, Ca and Mg were determined by atomic absorption spectrometry and Al by acid-base titration. P and K were extracted with ion exchange resin and determined using colorimetric and atomic emission spectrometry, respectively.

To calculate the sum of exchangeable bases (SB), the soil contents of Ca, Mg, and K were combined. The values of Ca, Mg, K, Al, and H were used to calculate the CEC (cation exchange capacity). Base saturation (BS%) was obtained from the ratio between SB and CEC. Finally, for the potential acidity (H + Al), the equation based on the Shoemaker-McLean-Pratt (SMP) buffer method was used.

2.2.3. DNA extraction and microarray probes analysis

Using the Power Lyzer Power Soil kit (MoBio Laboratories, Carlsbad, CA, United States), the genomic DNA was extracted from the subsamples of each mesocosm. After this procedure, the subsamples with DNA extracted were mixed, totaling 18 samples (3 treatments x 3 replicates x 2 collection periods). The quality and quantity of each one of 18 DNA samples were evaluated by spectrophotometry in a NanoDrop device (NanoDrop RND-1000 NanoDrop Technologies, Inc., Wilmington, DE, United States). DNA was stored at -20ºC until use.

In order to prepare the DNA for GeoChip v.5.0M analysis, each sample was purified and marked with the fluorescent dye Cy-3 (Wu et al., 2006). The DNA (600 ng each) was mixed with random primers (300 ng mL^{-1}), denatured at 99.9 °C for 5 min and immediately cooled on ice. A solution containing 5 mM of dAGC-TP, 2.5 mM of dTTP, 40 U of the Klenow fragment, and 25 nM of Cy-3 dUTP was added to the denatured DNA, and the reaction volume was adjusted to 50 mL with H_2O. The labeling solution was incubated at 37 °C for 6 h followed by

3 min at 95 °C. The labeled DNA was purified with the QIAquick Kit (Qiagen, Valencia, CA, United States), and dye incorporation was confirmed with a NanoDrop spectrophotometer (NanoDropR ND-1000 Technologies, Inc., Wilmington, DE, United States) using the absorption spectra of the standard solution for Cy-3. DNA samples were dried under vacuum and stored at -20°C until hybridization.

The tagged DNA was resuspended in DNase/RNase free distilled water and subsequently mixed with the solution for hybridization of the GeoChip v.5.0M probes, according to the procedure described in de Chaves et al. (2019). After hybridization, the slides containing the GeoChip microarrays were washed with a buffer solution (Agilent Wash Buffers 1 and 2 - Agilent Technologies Inc., Santa Clara, CA, USA) following the manufacturer's protocol. Using a NimbleGen MS200 scanner (Roche NimbleGen, Madison, WI, United States), the microarrays were digitalized, and the image data were extracted using Agilent Feature Extraction v. 2.6.

To determine the luminous intensity resulting from the hybridization of the genes marked with the corresponding probes, the software Agilent Feature Extraction v.11.5 was used (Agilent Technologies Inc., Santa Clara, CA, United States). The data were submitted to the Microarray Data Manager, available from the Institute for Environmental Genomics (website http: //www.ou.edu / content/ ieg / tools / data-analysis-pipeline.html), and analyzed according to the parameters described in de Chaves et al. (2019).

2.2.4. Shotgun sequencing of total soil genomic DNA and data processing

Eighteen DNA sequencing libraries were prepared using the Illumina Nextera sample preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Equal concentrations of libraries were loaded on the MiSeq Reagent v2 sequencing reagent kit (Illumina, San Diego, CA, USA). The equipment used for shotgun metagenomic sequencing was a MiSeq Personal Sequencing System by Illumina (Illumina, San Diego, CA, USA), operated in Rapid Run Mode to generate 2 x 250 bp paired-end reads. In summary, the sequencing resulted in an average of 105.5 MB of sequences per sample.

Regarding data processing, first, paired-end reads were merged using *FLASH* v. 1.2.5 (Magoč and Salzberg, 2011) to produce consensus sequences. Quality control of the consensus sequences was performed using the *Phred* quality score (20) to each base call (Ewing and Green, 1998) with an executable default script in *SeqClean* (http://www.bioinformatics.org/), and the low-quality bases were removed. Shotgun sequencing of the soil DNA samples resulted

in approximately 547,415 merged reads and 375,429 unmerged reads after the quality-based filtering procedure. Merged and unmerged trimmed sequences were concatenated into a single file for the metagenomic dataset, which is available on the Metagenomics Rapid Annotation (MG-RAST) server (https://www.mg- rast.org) under the "Metagenome of sugarcane soil - CENA/USP" project with accession numbers 4582104.3 to 4582153.3.

2.2.5. Selection of gene families, functions, and groups responsive to treatments

To select the gene families responsive to the treatments and their associated functions, the values obtained from the hybridization of the genomic DNA marked with the microarray probes corresponding to the category “N metabolism” were submitted to statistical analysis of variance (ANOVA followed by Tukey/LSD, $p < 0.05$) using the program Sisvar v.5.6. Gene families that showed statistically significant changes in absolute abundance in relation to the treatment N0 were used as references for the functional and taxonomic investigation analysis in the MG-RAST server.

The curated databases Refseq and Subsystems (SEED base annotation) (https://www.theseed.org/wiki/Home_of_the_SEED) were accessed for taxonomic and functional annotation, respectively. Functions investigated were selected according to the uniprot.org portal (https://www.uniprot.org/uniprot/P00370) and its previous description in the literature: i. each metagenome was analyzed in the Refseq database to determine the abundance of each group level; and ii. the sequences classified from Archaea or Bacteria were subjected to the functional filter selected previously in the Subsystems.

2.2.6. Statistical correlation analyzes

To verify which soil factors can affect the abundance of the N metabolism gene family, we performed a Spearman correlation using Sigma Plot software v.12.5 SYSTAT (Software Inc., CA, United States) and assumed a $p < 0.05$ threshold level for acceptance, between two datasets: soil chemical parameters and each previously significant selected gene family. All dataset combinations (positive and negative) were used for the network construction to illustrate the relationships between soil factors and gene families from N metabolism in soil with vinasse and urea fertilization. The constructed network was analyzed using *Gephi* software version 0.9.2 (Bastian et al., 2009), which has a graphical interface. The nodes in the network represent the soil factor and the selected gene family, and the edges represent significative ($p < 0.05$)

positive or negative correlations between the nodes. Network graphs were based on a set of statistical measures and parameters, such as the number of nodes and edges, modularity, clusterization coefficient, degree and betweenness centrality distribution, training algorithm parameters, and activation functions (Brandes, 2001; Blondel et al., 2008).

Spearman rank correlation coefficients were also calculated to generate a heatmap correlation graph in order to illustrate the connection degree of these datasets. This analysis was performed in the R environment v.4.0.2 with the "multtest" package: *i*. the absolute abundance of the selected gene families among the GeoChip hybridization signals and *ii*. the soil physicochemical factors (Pollard et al., 2005).

2.3. Results

2.3.1. Analysis of soil chemical factors

The majority (~ 84%) of the soil chemical properties analyzed were significantly responsive ($p < 0.05$) to the addition of vinasse, regardless of the period of collection of the soil samples (T7 or T150) (Table 1). However, in the T150 period, an increase of 10% in the contents of these attributes was observed. S and K (both with an ~90% increases in NV compared to N0) were highlighted, presenting the highest levels. Zn and Mn showed significant increase in the T7 period (25 and 65%, respectively), while Mg and SB, increased significantly only in T150 (23 and 30%). C and N showed significant increases in both sample collection periods, though, only in the treatment that received vinasse (15 and 5%, respectively, concerning treatment N0). However, these soil attributes showed no significant differences between the two periods (Table 1). Corresponding to the increases in C and N in the NV treatment, the C:N ratio showed a significant reduction of 18% in this treatment compared with the N0 treatment. Still considering the soil factors responsive to the addition of vinasse, the pH increased by ~6% in the two periods about to the control treatment.

When only urea was added to the soil (N60 treatment), the pH decreased significantly (8% compared to the N0 control treatment), but only in the T150 period. P was the only soil attribute that increased significantly in the treatment N60 (30% compared to N0) but only in T7. Considering the values of these soil factors between the two periods (T7 and T150), a smaller number of factors in the N60 treatment showed significant differences in their levels to those analyzed in the NV treatment (6 in N60 and 11 in NV, of a total of 18). In NV, Zn and K showed the highest rates of increase, with 95 and 70%, respectively.

Table 1. Chemical compounds present in the soil treatments without nitrogen fertilization (N0), with nitrogen fertilization (N60) and nitrogen fertilization combined with vinasse (NV) collected at seven (T7) and one hundred and fifty (T150) days after the start of the experiment

Soil properties		T7			T150			T7 vs. T150		
	Unit	N0	N60	NV	N0	N60	NV	N0	N60	NV
C	g. kg^{-1}	2.24 ± 0.04	2.29 ± 0.03	2.59 ± 0.07**	2.24 ±0.05	2.29 ± 0.03	2.59± 0.07**			
N	g. kg^{-1}	0.18 ± 0.00	0.19 ± 0.00	0.26± 0.02**	0.18 ± 0.00	0.19 ± 0.07	0.26 ± 0.02**			
pH		5.23 ± 0.06	5.23 ± 0.15	5.60± 1.09**	5.50 ± 0.10	5.10 ± 0.00**	5.80± 0.10*	0.016		0.025
O.M.	$g.dm^{-3}$	38.3 ± 1.53	36.3 ± 0.58	39.7 ± 3.79	32.3 ± 1.53	32.0 ± 1.00	35.7 ± 0.57*	0.008	0.003	
P	$mg.dm^{-3}$	61.0 ± 17.7	99.6*± 12.6	90.0 ± 12.1	259 ± 94.9	248 ± 83.5	213 ± 0.57	0.023	0.038	
S	$mg.dm^{-3}$	6.33 ± 1.53	11.7 ± 0.58	184 ± 16.5**	18.3 ± 7.64	17.3 ± 3.51	237± 22.3**			0.028
K	$mmolc.dm^{-3}$	1.23 ± 0.06	1.13 ± 0.06	11.4± 0.92**	3.60 ± 0.50	9.93 ± 2.72	39.0 ± 12.0**	0.001	0.005	0.014
Ca	$mmolc.dm^{-3}$	54.0 ± 2.64	55.0 ± 1.00	48.7 ± 3.50	71.0 ± 2.64	56.0 ± 16.6	66.3 ± 10.1	0.002		0.046
Mg	$mmolc.dm^{-3}$	16.7 ± 1.52	17.3 ± 0.57	18.7 ± 0.58	17.3 ± 0.60	14.0 ± 1.73	22.3± 2.08*		0.034	0.043
H+Al	$mmolc.dm^{-3}$	43.7 ± 2.88	44.0 ± 5.19	34.0± 0.00*	32.0 ± 1.73	39.3± 2.30*	25.0± 3.00*	0.004		0.007
CEC	$mmolc.dm^{-3}$	115 ± 6.22	117 ± 3.42	113 ± 4.90	124 ± 4.53	119 ± 20.3	153 ± 6.75			0.001
C:N	%	12.0 ± 0.30	11.9 ± 0.55	9.93± 0.52**	12.0 ± 0.37	12.0 ± 0.55	9.93 ± 0.52**			
SB	%	61.7 ± 1.15	62.3 ± 3.21	70.0 ± 1.00**	74.0 ± 1.00	66.7 ± 4.93*	84.0 ± 1.50	0.000		0.000
B	$mg.dm^{-3}$	0.24 ± 0.04	0.22 ± 0.00	0.14 ± 0.03*	0.29 ± 0.03	0.28 ± 0.02	0.37 ± 0.07		0.022	0.007
Cu	$mg.dm^{-3}$	0.93 ± 0.06	0.93 ± 0.06	0.93 ± 0.12	0.83 ± 0.06	0.93 ± 0.12	0.90 ± 0.20			
Fe	$mg.dm^{-3}$	37.0 ± 3.60	37.0 ± 4.00	48.3 ± 19.6	27.7 ± 2.08	38.0 ± 4.36	29.3 ± 7.51	0.018		
Mn	$mg.dm^{-3}$	7.40 ± 0.50	7.47 ± 1.12	21.0 ± 3.78**	6.27 ± 0.81	9.37 ± 1.04	21.0 ± 11.0			
Zn	$mg.dm^{-3}$	2.00 ± 0.50	1.70 ± 0.10	2.26 ± 1.48	10.9 ± 4.67	10.8 ± 4.90	47.0 ± 26.3**	0.030	0.032	0.042

N0. Treatment without the addition of nitrogen fertilizer; N60. Treatment with nitrogen at the dose of 60 kg ha^{-1}; and NV. Treatment with nitrogen at the dose of 60 kg $há^{-1}$ and the addition of vinasse at a rate of 0.06 L kg^{-1}. a = average for each of three replicates soil. b = standard deviations of the average for each of three replicates soil. The analysis were perform at seven (T7) and one hundred and fifty days (T150) after the beginning of the experiment. Analysis of variance and Turkey's HSD test two-pair combination with a significance level of * ($p < 0.05$) and ** ($p < 0.001$) about to N0. Blank spaces indicate that the values were not significant.

2.3.2. Selection of gene families responsive to treatments by GeoChip v.5.0M and their presence in Archaea and Bacteria

GeoChip v.5.0M has 30 gene families within the nitrogen category, which correspond to the following processes or functions, associated with nitrogen metabolism: nitrification, denitrification, ammonification, anammox, dissimilatory N reduction, assimilatory N, assimilatory N reduction, and respiration. In all gene families available in this version of GeoChip v.5.0, there were hybridizations between the corresponding probes and the DNA samples from the treatments.

For each replica of the samples, the values resulting from the quantification of the hybridization between each gene from the N metabolism and the corresponding probes were added. These values were used to represent the replicas in the statistical analysis. In a total of 30 gene families in the GeoChip v.5.0M (nitrogen category), ten (*amo*A, *hao*, *gdh*, *ure*C, *hzs*A, *nap*A, *nir*K, nitrate/nitrite_transporter, and *p450nor*) (Table 2) showed significant values ($p < 0.05$) in the variance test (Tukey/LSD). These gene families belong to the following N processes: *amoA* and hao to nitrification; *gdh* and *ure*C to ammonification; *hzs*A to annamox, *napA* to dissimilatory N; *nir*K to denitrification; nitrate and nitrite_transporter to assimilatory N; and *p450nor* to respiration in denitrification. In 90% of the gene families selected through statistical analysis, the highest number of hybridizations between the probes and the corresponding genes occurred in the T7 period. Only for the *hao* gene, the highest number of hybridizations occurred at T150 (Table 2). The highest abundance values of hybridizations occurred in 50% in of the samples corresponding to probes from Archaea (45% of the total available in GeoChip v.5.0M) (Table 2). The *gdh* gene, involved in the ammonification process, presented 6% of the total probes available from Archaea and 83% from Bacteria. *ure*C, *nap*A, and *nir*K presented 95 to 100% from Bacteria. The *hao* and *hzs*A genes had only probes from Bacteria available in GeoChip v.5.0M, as nitrate and nitrite transporter and *p450nor* with 100% from Eucarya (Table 2).

Interestingly, the *gdh* and *ure*C genes, associated with the ammonification process, were more abundant in the N60 treatment and the T7 sampling period, as well as *nir*K (denitrification). The number of hybridizations by the *p450nor* gene (associated with the respiration in the denitrification process) (Higgins et al., 2018), followed by the *hao* gene, presented the highest significance values in the statistical analysis ($p = 0.008$ and $p = 0.005$ respectively). Coincidentally, *p450nor,* and *hao* also showed the highest hybridization values in the treatment that received vinasse.

Table 2. GeoChip v.5.0M nitrogen family genes significance ($p < 0.05$) in Tukey LSD test, your function N cycle correspondence, treatment (N60 or NV), and collection period (T7 or T150) with more potential occurrence and origin of probes (Bacteria, Archaea and Eukarya)

N-Cycle Functions	Family genic in GeoChip v.5.0M	Treatment with the higher number of hybridizations	Period of greatest abundance of hybridizations	GeoChip v.5.0M probes from Bacteria	GeoChip v.5.0M probes from Archaea	GeoChip v.5.0M probes from Eukarya	*p* value
				---------- % ----------			
Nitrification	*hao*	NV	T150	100	0	0	0.005
	*amo*A	NV	T7	55	45	0	0.01
Ammonification	*gdh*	N60	T7	83	6	11	0.01
	ureC	N60	T7	95	3	2	0.01
Anammox	*hzsA*	NV	T7	100	0	0	0.02
Dissimilatory N	*napA*	N60	T7	99	1	0	0.02
Denitrification	*nirK*	N60	T7	96	1	3	0.02
Assimilatory N	nitrate_transporter	NV	T7	0	0	100	0.01
	nitrite_transporter	N60	T7	0	0	100	0.02
Respiration	*p450nor*	NV	T7	0	0	100	0.008

N0 treatment without the addition of nitrogen fertilizer; N60, treatment with nitrogen addition at the dose of 60 kg ha^{-1}; and NV, treatment with nitrogen fertilization at the dose of 60 kg ha^{-1} and the addition of vinasse at a rate of 0.06 L kg^{-1}. The analysis were perform at seven (T7) and one hundred and fifty days (T150) after the beginning of the experiment in a greenhouse. Analysis of variance and Tukey/ LSD test was realized in Sisvar® v. 5.6. Software in a significance level ($p < 0.05$) in relation of N0 treatment. GeoChip v.5.0M has probes constructed based on DNA of Archaea, Bacteria, or Eucaria and here we detailed it in percentage (%).

2.3.3. Genes, functions, and groups associated with the N metabolism analyzed in the MG-RAST server

Only three of the eight functions analyzed in MG-RAST were present in Archaea (Table 3). Ammonia monooxygenase, glutamate dehydrogenase, and urease are enzymes associated with *amo*A, *gdh,* and *ure*C genes (Tables 2 and 3). These genes correspond to the nitrification (*amo*A) and ammonification (*gdh* and *ure*C) processes (Table 2). The genus *Nitrosopumillus,* from the phylum Thaumarchaeota, presented the greatest number of sequences of ammonia monooxygenase among Archaea in treatment NV and period T7 (Table 3). *Sulfolobus* (from Crenarchaeota) presented the greatest number of sequences of glutamate dehydrogenase and *Natronomonas* (from Euryarchaeota) of urease, both in period T150. The glutamate dehydrogenase in treatment NV and urease in N60. However, only the value from urease in *Natronomonas* was statistically significant ($p < 0.05$ in Tukey/LSD). Ammonia monooxygenase sequences were analyzed at domain and genus levels. From total Archaea, this function presented a four fold decrease between NVT7 and NV150 in numbers of sequences. From Bacteria, between NVT7 and NVT150, it increased twice.

All functions investigated were found in Bacteria. Proteobacteria from the genus *Anaeromyxobacter* stood out because they exhibited the greatest number of sequences of glutamate dehydrogenases, urease, and nitrate/nitrite transporter. *Anaeromyxobacter* was the most frequent genus among Bacteria with the selected functions (40% of the total), and NVT7 included all these sequences (Table 3). The oxidoreductase enzyme, associated with the *hao* gene family, was found only in Bacteria, in which the treatment NVT7 had the greatest number of sequences. The abundances of functions were statistically significant to nitrate/nitrite transporter in *Anaeromyxobacter,* a Proteobacteria ($p < 0.001$).

The treatment that received vinasse combined with urea, NV, presented the majority of groups with the selected functions (~70%). Conversely, N60 presented only one urease function in *Natronomonas* Archaea. The gene family *hzs*A was not found in the subsystems by MG-RAST (Table 3).

Table 3. Archaea and bacteria frequency of functions in metagenomes corresponding of N cycle gene families selected by GeoChip v.5.0M

Gene family (GeoChip)	Function in Subsystems (MG-RAST)	Bacteria (genus)	Sequences inside total of Bacteria in treatment (%)	Treatment with greater abundance	Archaea (genus)	Sequences inside total of Archaea in treatment (%)	Treatment with greater abundance
amoA	Ammonia_monoxigenase	Nitrosomonas	0.000114	NVT150	Nitrosopumillus	0.1057	NVT7
hao	Oxidorreductase	Nitrosomonas	0.000151	NVT7	-		
gdh	Glutamate dehydrogenase	Anaeromyxobacter	0.000102	NVT7	Sulfolobus	0.0394	NVT150
ureC	Urease	Anaeromyxobacter	0.001021	NVT7	Natronomonas	0.0129**	N60T150
napA	Nitrate reductase cytocromo c550 type subunit	Bradyrhizobium	0.000472	N0T7	-		
nirK	Nitrite reductase	Acidovorax	0.000225	N0T7	-		
Nitrate/nitrite transporter	Nitrate/nitrite transporter	Anaeromyxobacter	0.000392**	NVT7	-		
p450nor	Nitric oxide reductase activation	Mycobacterium	0.000510	NVT150	-		
hzsA	-				-		

N0, treatment without the addition of nitrogen fertilizer; N60, treatment with nitrogen addition at the dose of 60 kg ha^{-1}; and NV, treatment with nitrogen fertilization at the dose of 60 kg ha^{-1} and the addition of vinasse at a rate of 0.06 L kg^{-1}. The analysis were performed at seven (T7) and one hundred and fifty days (T150) after the beginning of the experiment in a greenhouse. Analysis of variance and Tukey/ LSD test were realized in Sisvar® v. 5.6. Software in a significance level ($p < 0.05$) in relation of N0 treatment. Tukey/LSD * $p < 0.05$; ** $p < 0.001$

2.3.4. Correlation analyses between selected gene families and physicochemical soil factors

Heatmaps and network correlation tests between GeoChip v.5.0M gene families responsive to N treatments and soil physicochemical properties (Figs. 1 and 2) revealed that T7 featured two times more correlations than T150, where the majority in T7 were positives (85% of the network) and in T150, negatives. The soil factor P was involved in 80% of the positive correlations in heatmap for T7, and in the network, T7 was the soil factor that was most correlated (12% of the total). Particularly in the network, all correlations involving P were positive. P was especially involved with the ammonification gene family in both analyses, but only in T7. In the heatmap for T7 (Fig. 1), OM was involved in 75% of the negative correlations, but in the network T150 was the soil factor that established the most correlations (1/4), with all being positive.

The *p450nor* gene family established most of the correlations in the network T7 (1/4 of total), and the majority of then (85%) were positive (Fig. 2). In T150, the *hao* gene family was the most present gene family in correlations, involved in 60% of them in the heatmap and 75% in the network, with all being positives, especially with N (Figs. 1 and 2). Curiously, in the heatmap T7 (Fig. 1), N had positive correlations with the *p450nor* gene, and in T150, the correlations were positive with the *hao* gene. Instead, C:N had negative correlations with the *p450nor* gene in T7 and negative correlations in T150 with the *hao* gene. Another interesting result is that *hao* did not exibit any correlations in the T7 network analysis (Fig. 2).

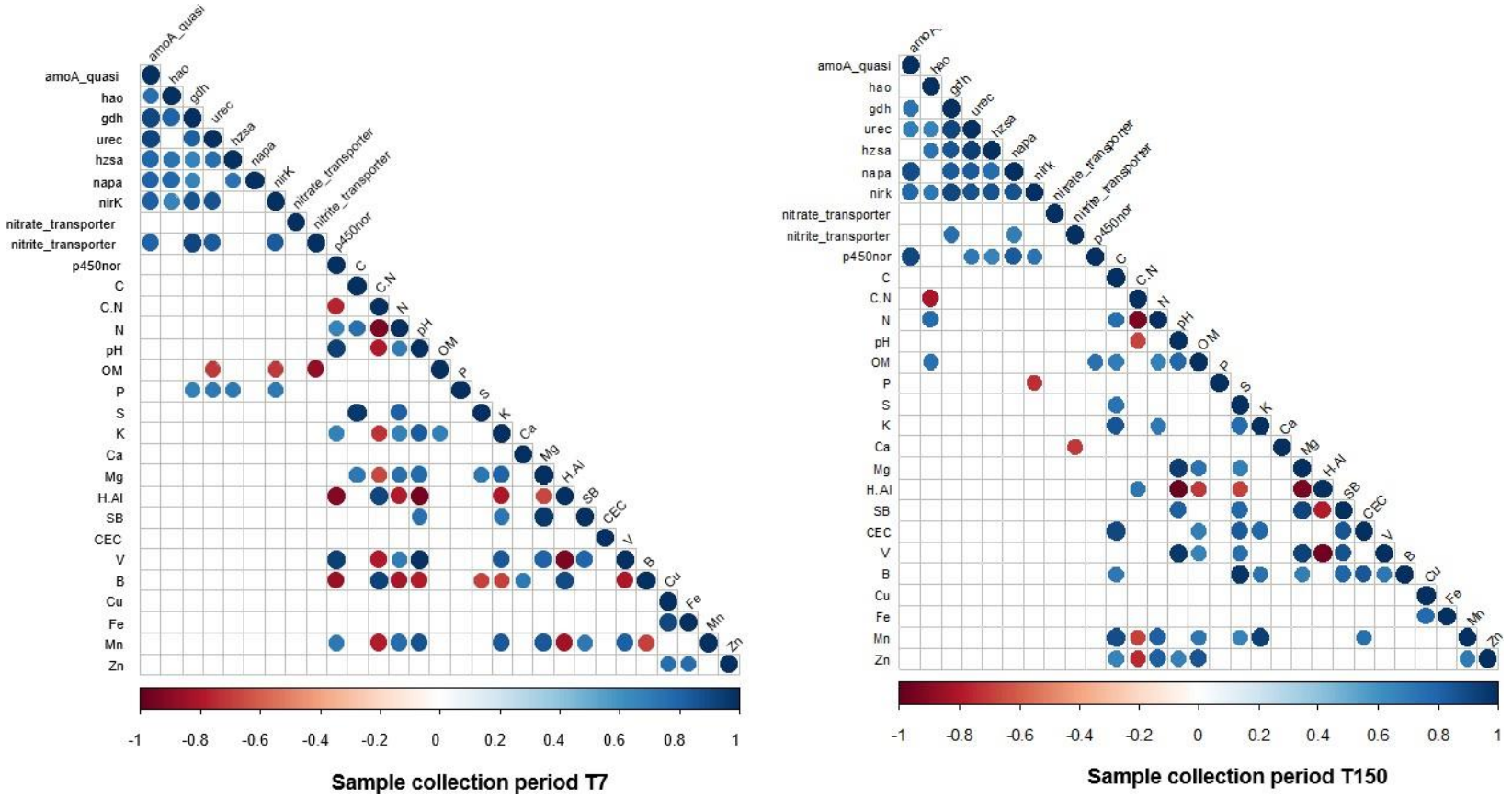


Figure 1. Heatmaps correlation test between GeoChip v.5.0M gene family responsive to N treatments and physicochemical soil properties, according to significant values ($p < 0.05$) of Spearman rank correlation coefficient performed in R environment with "Hmisk and CorrPlot" package. Positive values (blue circle) represents positive correlation and negative values (red circle) represents negative correlation.

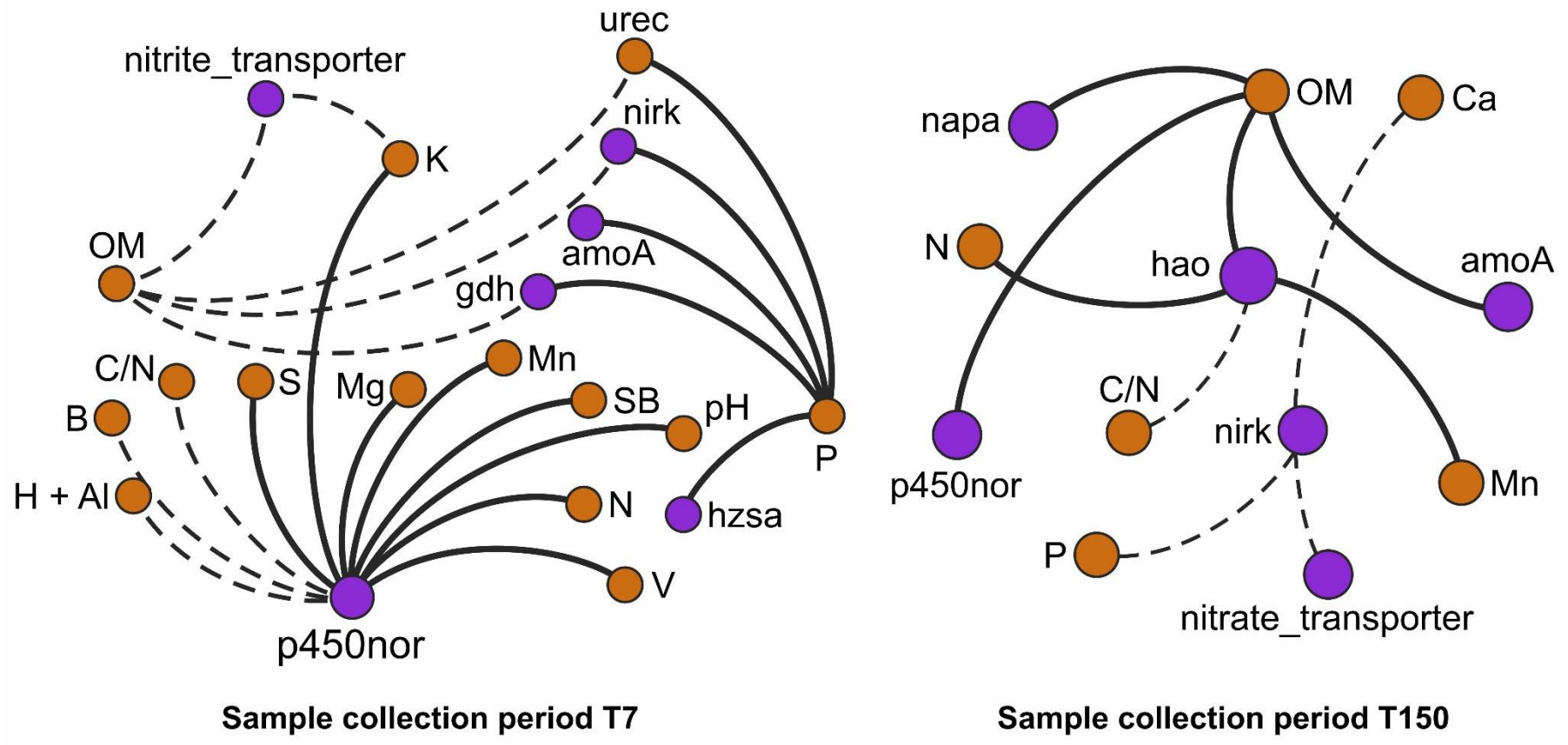


Figure 2. Networks correlation analysis of the GeoChip v.5.0M gene families responsive to N treatments and physicochemical soil properties based on significant ($p < 0.05$) Spearman rank correlation coefficient values. The nodes represent the physicochemical soil factors (light brown) and gene family (purple). The thick edges represent positive correlations, and the dashed edges represent negative correlations

2.4. Discussion

Although the use of urea as a nitrogen fertilizer combined with vinasse is widely recommended, researchers and government agencies have carried out studies in the field and laboratories to evaluate their roles as emission sources of GHGs, such as N_2O, which results from this practice. Understanding the microbial processes involved in the production of this gas has become relevant for the sugar and alcohol sector since N losses in the sugarcane systems have environmental and economic impacts on society. Our results using GeoChip v.5.0M technology showed that nearly 87% of the functions related to the N metabolism might have been active in the first days after the addition of fertilizers. There was a prevalence of correlations between the two datasets, contents of soil properties, and gene families responsive to treatments, in the T7 period, suggesting the occurrence of great microbial metabolic activities in this period. The rapid increases in the availability of N and C sources in the system resulting from the addition of fertilizers (Table 1) would have led the microbiota quickly accessing these less complex sources of nutrients and performing the immobilization and mineralization of N (Han et al., 2004). Consequently, the nitrification process must have been quickly activated, which justified the great abundance of *amo*A genes detected in the NV treatment (Table 2). Other evidence, of nitrification occurrence on a large scale was the increase of the ammonia monooxygenase sequences from AOA *Nitrosopumillus* (Table 3) and the reduction of the C:N ratio in NVT7 (Table 1), which suggests losses of N in the form of NO_3 (Han et al., 2004; Schroder et al., 2011). Moreover, in the same treatment, Navarrete et al. (2015) detected the greatest production of N_2O. It is known that the addition of this organic residue results in the availability of nutrients in sugarcane soils that are able to promote changes in the structure and functionality of the microbiota associate with N and C cycles (Navarrete et al., 2015; Chaves et al., 2019).

Three nutrients analyzed here deserve to be highlighted due to their relevance in response to treatments: Zn, K, and P, which, especially in the T150 period, showed a great increase (Table 1). A study evaluating the number of copies of the *amo*A gene from Archaea and Bacteria in an Australian soil submitted to different doses of Zn concluded that high doses of this nutrient could have, inhibited the nitrification process of ammonia related to nitrification (Mertens et al., 2009). In the same study, after only two years, the population of ammonia-oxidizing Bacteria (AOB) became resilient to high doses of Zn and was able to realize the process. The same inhibitory effect promoted by Zn on the nitrification can be observed in other studies (Mertens et al., 2007; Liu et al., 2015). Since Zn is one of the most abundant

micronutrients and heavy metals in vinasse (Wilkie et al., 2000; Prasad and Shih, 2016), in our study, it was the soil chemical factor that showed the greatest increase between periods T7 and T150 in the NV treatment (Table 1). High levels of this micronutrient in the NV treatment could justify the reduction of ammonia monooxygenase in T150 associated with the *amo*A gene from *Nitrosopumillus*. In addition, it is possible that ammonia monooxygenase sequences from *Nitrosomonas* (AOB) increase only in T150 because of their resilience to the Zn inhibition effect.

The K was another nutrient that vinasse residue supplies in large quantities (Moraes et al., 2015; Parsaee et al., 2019), which was already demonstrated to be a good nitrification reducer in alkaline soils (Cai et al., 2018). In an incubation experiment with potassium thiosulfate (KTS), N_2O production decrease after twelve days application (Cai et al., 2018). About P, it has already been demonstrated that in agricultural alkaline soils, the high availability of this nutrient can reduce the nitrification process (Shi et al., 2016). Here, although the pH values remained between 5.2 and 5.8 (acidic), there was a significant increase in this factor ($p < 0.05$) in the NVT150 that could favor the availability of P in soil with vinasse. In addition, our results obtained from correlations analysis showed that P was strongly associated with gene families in the first days after the application of N fertilizers, suggesting its influence on N metabolism from microbiota (Fig. 1 and 2). The gene family *hzs*A, associated with the anammox process had the greatest abundance in NVT7 and also showed positive correlations with P in this study and other agricultural soils (Zhou et al., 2017). This result suggests an influence from this nutrient on the anammox process, besides of nitrification. Once in this N process, the NO_2 is used as an electron acceptor to oxidase NH_3 due to anaerobic conditions (Narita et al., 2017) and is possible that the higher consumption of O_2 in the first days of application of the N fertilizer with vinasse may promote oxygen-limiting conditions.

In the NV treatment, while the pH had the lowest values in T7 in relation to the T150 period, OM and P were higher under these same conditions (Table 1). If in alkaline soils, the inhibitory effect of P can be efficient, in acidic soils P would be adsorbed to OM and clay (Shi et al., 2016), which would neutralize it. Interestingly, the only soil propertie that showed correlations with the *amo*A gene in the networks were OM and P (Figs. 1 and 2). The acid pH itself could reduce the nitrification function, as evidenced in previous studies (Schroder et al., 2011), but it inhibits the function, only below pH 4.5 (Maier et al., 2009). In NV treatment, the pH was higher in the T150 period, which would help nitrification (Maier et al., 2009; Schroder et al., 2011). In NVT150, the ammonia monooxygenase function

from *Nitrosomonas*, AOB, had the highest abundance. However, if the potential inhibitory effect from P on nitrification could not occur due to this acid pH in NVT7, another N processes, such as denitrification, could be occurring on a large scale and increasing the N_2O production, as observed in NVT7 by Navarrete et al. (2015).

The *p450nor* gene family stood out in our analysis as the most correlated gene family with soil properties (Figs. 1 and 2). It presented the greatest prevalence in NV (p = 0.008) in the T7 period according to the GeoChip analysis (Table 2). However, considering the nitric oxide reductase activation function from Bacteria associated with *p450nor*, the prevalence was in T150 (Table 3). The *p450nor* gene (cytochrome P450 nitric oxide (NO) reductase (Nor)) is related to the reduction of NO_2 to N_2O (Novinscak et al., 2016; Higgins et al., 2018). The system is located in the mitochondria and is activated during anaerobic respiration (Shoun et al., 2012), which is carried out by denitrifying fungi in agricultural soils (Shoun et al., 2012; Mothapo et al., 2015; Wu et al., 2017), which cause plant diseases (Shoun et al., 2012; Mendes et al., 2019). In Bacteria, this denitrification gene has been demonstrated to enhance virulence against the host because it uses an alternative electron acceptor, which is favorable to their lifestyle under anaerobic conditions (Fritz et al., 2002). According to our results, there were a significant increases in *p450nor* in NV treatment and in the T7 period and, regarding its associated functions with *Mycobacterium* (Table 3), in the T150 period. Suleiman et al. (2018) obtained increases in the sporulation of fungi concomitantly with the highest volumes of N_2O during the first days after the addition of vinasse in sugarcane soils. According to Higgins et al. (2016), environments with regular N inputs impose the utilization of N-oxides by the microbiota. These observations suggest that fungi denitrification by *p450nor* could have occurred and increased under NVT7 conditions. Lourenço et al. (2018) associated the production of N_2O in sugarcane soils under vinasse with N fertilizers with the *nir*K gene from denitrifying fungi; however, this study did not analyze the *p450nor* denitrification gene. Instead, the *nir*K gene seems not to be assessed here on a large scale by the microbiota in NVT7 since it presented the greatest abundance in N60 and sequences of nitrite reductase in N0.

Another soil factor that may benefits from the increase of *p450no*r in NV is the high availability of S (Table 1), which can also intensify the denitrification process, with a reduction in sulfate for the decomposition of organic matter (Lau et al., 2006) and an increase in N_2O production (Simek et al., 2011) This observation could contribute to justify the positive correlations among the *p450nor* gene, OM, and S in correlation analyses (Figs. 1 and 2). According to Schlüter et al. (2018), the lack of O_2 and the availability of NO_3 as an electron acceptor are basic requirements for heterotrophic denitrification in soil. It is possible that the

rapid immobilization of N is readily available to microorganisms in the first days after fertilization, while the sugarcane seedlings, which do not require large amounts of this nutrient (Mariano et al., 2016), favored the nitrification processes related to *amo*A and denitrification related to *p450nor* in NV treatment.

In this study, one of the genes that showed the greatest response to NV was *hao*, which encodes the oxidoreductase enzyme in the nitrification process (HAO) (Whittaker et al., 2000; Koch et al., 2019). Zhao et al. (2019) quantified the *hao* gene associated with nitrifying Bacteria in a lake that received domestic sewage. The lake regions with the highest amounts of NO_3 and NH_3 showed the highest abundance of *hao*, suggesting the association of this gene with the eutrophication process. Here, the low C:N ratio (< 10) in the NV treatment indicates losses of N in the system by leaching, which could be in the form of NO_3. Is it possible that the physicochemical conditions of the soil submitted to the application of vinasse in this period were similar to a eutrophication process such as that described in Zhao et al. (2019)? Analyzing the oxidoreductase enzymes that correspond to the *hao* gene family, *Nitrosomonas* (AOB), which is related to N_2O emissions by nitrification (White and Lehnert, 2016), had the greatest number of sequences in NVT7. Wu et al. (2019) concluded that *Nitrosomonas* and *Nitrosopumillus* tended to dominate in moderately eutrophic sediment with greater ammonium input (2.86 mM). In addition, the higher prevalence of ammonia monooxygenase and oxidoreductase sequences in NVT7 suggests a high rate of nitrification, which would also lead to the formation of NO_3 and N_2O. Smith et al. (2015) concluded that the *napA* gene, associated with the reduction of nitrate to nitrite in denitrification, is more important in environments with lower concentrations of NO_3. In our microarray analyses, *nap*A was more abundant in the N60 treatment and, in the MG-RAST analysis, the nitrate reductase in N0 (Tables 2 and 3), suggesting that NO_3 could be found in higher concentrations in NV.

The gene families *ure*C and *gdh* are associate with the ammonification process, in which the formation of NH_4 from the organic matter by microorganisms occurs (Dannenmann et al., 2009). Once a great number of correlations between *ure*C and *gdh* with P were identified, we also considered the potential effect of this element on the ammonification process. Previous studies revealed the inhibitory effect of KH_2PO_4 on the ammonification (Ryan et al., 1972). In our study, the contents of K increase in N60 and NV treatments at the T150 period (Table 1). This could justify the increases of *ure*C and *gdh* in N60T7, and urease and glutamate dehydrogenase from Bacteria in NVT7 and their decrease in NVT150.

The results also revealed the greatest number of nitrate and nitrite transporter functions (related to ammonification), from genus *Anaeromyxobacter*, a Proteobacteria (Table 3). Due to its versatile metabolism, this genus is composed of ecologically competitive Bacteria that can survive in several ecosystems, microaerophilic or anaerobic, including agricultural soils (Sanford et al., 2012). This genus can reduce NO_3 to NH_4 via NO_2 as an electron acceptor, and to reduce N_2O (Yoon et al., 2016). A specific clade from the *nos*Z operon, *A. dehalogenans*, has a higher affinity for N_2O, suggesting that they are more competitive under the environmental levels of N_2O (Yoon et al., 2016). *Anaeromyxobacter* type *nos*Z sequences were found to be the most abundant *nos*Z genes in agricultural soils from Illinois (USA) and appear to compose a significant portion of the N_2O-reducing community (Sanford et al., 2012; Orellana et al., 2014), suggesting an important role in the N cycling. On the other hand, Archaea associated with the ammonification process was frequent in the T150 period, where the salt concentrations were larger than in T7. While the T7 nonsaline soil conditions may favor AOA *Nitrosopumillus,* Archaea with organic matter recycling capabilities, such as *Natronomonas* (from order Halobacteriales), was found in the saline soil condition in T150 (Navarro-Noya et al., 2015). *Natronomonas* is a haloalkaline Archaea with the capacity to survive in soils with higher salt contents (Navarro-Noya et al., 2015), suggesting that in the 150 period the treatments with N fertilizer have more salts than in T7 and selected Archaea with specialized functions. According to our results, the highest number of urease sequences in Archaea is from *Natronomonas*. Some studies have shown that urease is sensitive to low pH (Longo and Melo, 2005), which may justify the presence of halophiles in soils with more salts.

GeoChip v.5.0M is a highly specific, sensitive, and quantitative method based on both computational and experimental assays (Shi et al., 2019). This technology can reveal information about the functionality of the microbial community in several environments, not yet analyzed through other molecular techniques (Carter et al., 2012; Shi et al., 2019). In this study, the analyses of metagenomes combined with GeoChip v.5.0M allowed the identification of which gene families associated with the N cycle significantly respond to the physicochemical changes resulting from typical agricultural management in tropical soils. Moreover, it was possible to indicate groups of Archaea and Bacteria associated with the N process, with potential increases in higher salt contents and N_2O levels, however, in low O_2 sources. This information suggests their specific metabolic capabilities in restricted conditions. The results also call attention to the increase of the *p450nor* gene family in these conditions that were associated with the denitrification process by Fungi pathogenic to plants, and *Mycobacterium* pathogenic to animals. Finally, the Bacteria and Archaea domains are certainly ubiquitous in

soil but have developed very specialized functions depending on the physical and chemical characteristics of their environments (Navarro-Noya et al., 2015).

2.5. Conclusion

N losses in sugarcane crops soils by management practices promote economic and environmental impacts that need attention. Our results suggest "key" genes and processes in the nitrogen cycle that may be changed due to the new composition of soil factors under vinasse and urea fertilization. Our DNA analysis from the soil samples using microarray technology and metagenome annotation revealed that the genes families related to ammonification (*gdh* and *ure*C from Bacteria and Archaea*)*, nitrification (*amo*A from AOA and AOB), and denitrification (*p450nor*) are the main processes responsive to this typical fertilizer management in the first days after the application. Ammonia oxidizers and halophile Archaea were responsive to the increase of salts of K and nutrients such as P in soil under vinasse application. This common practice for cultivated sugarcane has the potential to promote the selection of pathogenic Bacteria and fungi, but also may favor an increase of *Anaeromyxobacter*, which can reduce the GHG N_2O gas.

References

Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. BT - International AAAI Conference on Weblogs and Social. Int. AAAI Conf. Weblogs Soc. Media, 361–362.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. 2008, 1–12. doi:10.1088/1742-5468/2008/10/P10008.

Brandes, U. (2001). A faster algorithm for betweenness centrality, The Journal of Mathematical Sociology, 25:2, 163-177. doi:10.1080/0022250X.2001.9990249.

Cai, Z., Gao, S., Xu, M., and Hanson, B. D. (2018). Evaluation of potassium thiosulfate as a nitrification inhibitor to reduce nitrous oxide emissions. Sci. Total Environ. 618, 243–249. doi:10.1016/j.scitotenv.2017.10.274.

Carter, M. Q., Xue, K., Brandl, M. T., Liu, F., Wu, L., Louie, J. W., et al. (2012). Functional Metagenomics of Escherichia coli O157:H7 Interactions with Spinach Indigenous Microorganisms during Biofilm Formation. PLoS One 7. doi:10.1371/journal.pone.0044186.

Christofoletti, C. A., Escher, J. P., Correia, J. E., Marinho, J. F. U., and Fontanetti, C. S. (2013). Sugarcane vinasse: Environmental implications of its use. Waste Manag. 33, 2752–2761. doi:10.1016/j.wasman.2013.09.005.

Cloutier, M. L., Bhowmik, A., Bell, T. H., and Bruns, M. A. (2019). Innovative Technologies Can Improve Understanding of Microbial Nitrogen Dynamics in Agricultural Soils. Agric. Environ. Lett. 4, 1–6. doi:10.2134/ael2019.08.0032.

Cong, J., Liu, X., Lu, H., Xu, H., Li, Y., Deng, Y., et al. (2015). Analyses of the influencing factors of soil microbial functional gene diversity in tropical rainforest based on GeoChip 5.0. Genomics Data 5, 397–398. doi:10.1016/j.gdata.2015.07.010.

Dannenmann, M., Simon, J., Gasche, R., Holst, J., Naumann, P. S., Kögel-Knabner, I., et al. (2009). Tree girdling provides insight into the role of labile carbon in nitrogen partitioning between soil microorganisms and adult European beech. Soil Biol. Biochem. 41, 1622–1631. doi:10.1016/j.soilbio.2009.04.024.

De Chaves, M. G., Silva, G. G., Rossetto, R., Edwards, R. A., Tsai, S. M., and Navarrete, A. A. (2019). Acidobacteria subgroups and their metabolic potential for carbon degradation in sugarcane soil amended with vinasse and nitrogen fertilizers. Front. Microbiol. 10. doi:10.3389/fmicb.2019.01680.

Dias, M. O. S., Ensinas, A. V., Nebra, S. A., Maciel Filho, R., Rossell, C. E. V., and Maciel, M. R. W. (2009). Production of bioethanol and other bio-based materials from sugarcane bagasse: Integration to the conventional bioethanol production process. Chem. Eng. Res. Des. 87, 1206–1216. doi:10.1016/j.cherd.2009.06.020.

Ewing, B., and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8, 186–194. doi:10.1101/gr.8.3.186.

Faoro, H., Alves, A. C., Souza, E. M., Rigo, L. U., Cruz, L. M., Al-Janabi, S. M., et al. (2010). Influence of soil characteristics on the diversity of bacteria in the southern brazilian atlantic forest. Appl. Environ. Microbiol. 76, 4744–4749. doi:10.1128/AEM.03025-09.

Fritz, C., Maass, S., Kreft, A., and Bange, F. C. (2002). Dependence of Mycobacterium bovis BCG on anaerobic nitrate reductase for persistence is tissue specific. Infect. Immun. 70, 286–291. doi:10.1128/IAI.70.1.286-291.2002.

Gasparotto, F., Rodrigues, F. S., Seratoo, C. D., and da costa, T. R. (2014). Cadeias Produtivas da Cana-de-açúcar, do Algodão e de Frutas. Maringá: UniCesumar.

Glória, N.A. Uso Agronômico de Resíduos. in: Reunião Brasileira de Fertilidade do Solo e Nutrição de Plantas, 20. Piracicaba, 1992. p.195-212.

Han, K. H., Choi, W. J., Han, G. H., Yun, S. I., Yoo, S. H., and Ro, H. M. (2004). Urea-nitrogen transformation and compost-nitrogen mineralization in three different soils as affected by the interaction between both nitrogen inputs. Biol. Fertil. Soils 39, 193–199. doi:10.1007/s00374-003-0704-4.

Hidalgo, K., Rodríguez, B., and Febles, M. V. M. (2009). Utilización de la vinaza de destilería como aditivo para pollos en ceba. Rev. Cuba. Cienc. Agrícola 43, 281–284.

Higgins, S. A., Schadt, C. W., Matheny, P. B., and Löffler, F. E. (2018). Phylogenomics reveal the dynamic evolution of fungal nitric oxide reductases and their relationship to secondary metabolism. Genome Biol. Evol. 10, 2474–2489. doi:10.1093/gbe/evy187.

Higgins, S. A., Welsh, A., Orellana, L. H., Konstantinidis, K. T., Chee-sanford, J. C., Sanford, R. A., et al. (2016). Detection and Diversity of Fungal Nitric Oxide Reductase Genes (p450nor) in Agricultural Soils. Appl. Env. Microbiol. 82, 2919–2928. doi:10.1128/AEM.00243-16.

Jenkins, S. N., Waite, I. S., Blackburn, A., Husband, R., Rushton, S. P., Manning, D. C., et al. (2009). Actinobacterial community dynamics in long term managed grasslands. Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol. 95, 319–334. doi:10.1007/s10482-009-9317-8.

Jesus, E. D. C., Marsh, T. L., Tiedje, J. M., and Moreira, F. M. D. S. (2009). Changes in land use alter the structure of bacterial communities in Western Amazon soils. ISME J. 3, 1004–1011. doi:10.1038/ismej.2009.47.

Koch, H., van Kessel, M. A. H. J., and Lücker, S. (2019). Complete nitrification: insights into the ecophysiology of comammox Nitrospira. Appl. Microbiol. Biotechnol. 103, 177–189. doi:10.1007/s00253-018-9486-3.

Lau, G. N., Sharma, K. R., Chen, G. H., and van Loosdrecht, M. C. M. (2006). Integration of sulphate reduction, autotrophic denitrification and nitrification to achieve low-cost excess sludge minimisation for Hong Kong sewage. Water Sci. Technol. 53, 227–235. doi:10.2166/wst.2006.101.

Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl. Environ. Microbiol. 75, 5111–5120. doi:10.1128/AEM.00335-09.

Lehtovirta-Morley, L. E. (2018). Ammonia oxidation: Ecology, physiology, biochemistry and why they must all come together. FEMS Microbiol. Lett. 365, 1–9. doi:10.1093/femsle/fny058.

Liu, A., Fang, D., Wang, C., Li, M., and Young, R. B. (2015). Recovery of soil nitrification after long-term zinc exposure and its co-tolerance to Cu in different soils. Environ. Sci. Pollut. Res. 22, 314–319. doi:10.1007/s11356-014-3338-1.

Longo, R. M., and Melo, W. J. de (2005). Hidrólise da uréia em latossolos: efeito da concentração de uréia, temperatura, pH, armazenamento e tempo de incubação. Rev. Bras. Ciência do Solo 29, 651–657. doi:10.1590/s0100-06832005000400018.

Lourenço, K. S., Dimitrov, M. R., Pijl, A., Soares, J. R., do Carmo, J. B., van Veen, J. A., et al. (2018). Dominance of bacterial ammonium oxidizers and fungal denitrifiers in the complex nitrogen cycle pathways related to nitrous oxide emission. doi:10.1111/gcbb.12519.

Madejón, E., Burgos, P., López, R., and Cabrera, F. (2001). Soil enzymatic response to addition of heavy metals with organic residues. Biol. Fertil. Soils 34, 144–150. doi:10.1007/s003740100379.

Maier, R. M. Pepper, I. L., Gerba, C. P., eds. Environmental microbiology, 2nd ed, Elsevier/Academic Press, Amsterdam ; Boston, 2009.

Magoč, T., and Salzberg, S. L. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27, 2957–2963. doi:10.1093/bioinformatics/btr507.

Mariano, E., Leite, J. M., Vieira-Megda, M. X., Ciampitti, I. A., Vitti, A. C., Faroni, C. E., et al. (2016). Biomass and nutrient content by sugarcane as affected by fertilizer nitrogen sources. Crop Sci. 56, 1234–1244. doi:10.2135/cropsci2015.06.0349.

Mendes, L. W., de Chaves, M. G., Fonseca, M. de C., Mendes, R., Raaijmakers, J. M., and Tsai, S. M. (2019). Resistance Breeding of Common Bean Shapes the Physiology of the Rhizosphere Microbiome. Front. Microbiol. 10, 1–10. doi:10.3389/fmicb.2019.02252.

Mertens, J., Broos, K., Wakelin, S. A., Kowalchuk, G. A., Springael, D., and Smolders, E. (2009). Bacteria, not archaea, restore nitrification in a zinc-contaminated soil. ISME J. 3, 916–923. doi:10.1038/ismej.2009.39.

Mertens, J., Degryse, F., Springael, D., and Smolders, E. (2007). Zinc toxicity to nitrification in soil and soilless culture can be predicted with the same biotic ligand model. Environ. Sci. Technol. 41, 2992–2997. doi:10.1021/es061995+.

Moraes, B. S., Junqueira, T. L., Pavanello, L. G., Cavalett, O., Mantelatto, P. E., Bonomi, A., et al. (2014). Anaerobic digestion of vinasse from sugarcane biorefineries in Brazil from energy, environmental, and economic perspectives: Profit or expense? Appl. Energy 113, 825–835. doi:10.1016/j.apenergy.2013.07.018.

Moraes, B. S., Zaiat, M., and Bonomi, A. (2015). Anaerobic digestion of vinasse from sugarcane ethanol production in Brazil: Challenges and perspectives. Renew. Sustain. Energy Rev. 44, 888–903. doi:10.1016/j.rser.2015.01.023.

Moran-Salazar, R. G., Sanchez-Lizarraga, A. L., Rodriguez-Campos, J., Davila-Vazquez, G., Marino-Marmolejo, E. N., Dendooven, L., et al. (2016). Utilization of vinasses as soil amendment: consequences and perspectives. Springerplus 5. doi:10.1186/s40064-016-2410-3.

Mothapo, N., Chen, H., Cubeta, M. A., Grossman, J. M., Fuller, F., and Shi, W. (2015). Phylogenetic, taxonomic and functional diversity of fungal denitrifiers and associated N2O production efficacy. Soil Biol. Biochem. 83, 160–175. doi:10.1016/j.soilbio.2015.02.001.

Narita, Y., Zhang, L., Kimura, Z. ichiro, Ali, M., Fujii, T., and Okabe, S. (2017). Enrichment and physiological characterization of an anaerobic ammonium-oxidizing bacterium '*Candidatus Brocadia sapporoensis*.' Syst. Appl. Microbiol. 40, 448–457. doi:10.1016/j.syapm.2017.07.004.

Navarrete, A. A., Diniz, T. R., Braga, L. P. P., Silva, G. G. Z., Franchini, J. C., Rossetto, R., et al. (2015). Multi-analytical approach reveals potential microbial indicators in soil for sugarcane model systems. PLoS One 10, 1–19. doi:10.1371/journal.pone.0129765.

Navarro-Noya, Y. E., Valenzuela-Encinas, C., Sandoval-Yuriar, A., Jiménez-Bueno, N. G., Marsch, R., and Dendooven, L. (2015). Archaeal communities in a heterogeneous Hypersaline-Alkaline soil. Archaea 2015. doi:10.1155/2015/646820.

Nilsson, L. O., Bååth, E., Falkengren-Grerup, U., and Wallander, H. (2007). Growth of ectomycorrhizal mycelia and composition of soil microbial communities in oak forest soils along a nitrogen deposition gradient. Oecologia 153, 375–384. doi:10.1007/s00442-007-0735-x.

Novinscak, A., Goyer, C., Zebarth, B. J., Burton, D. L., Chantigny, M. H., and Filion, M. (2016). Novel P450nor gene detection assay used to characterize the prevalence and diversity of soil fungal denitrifiers. Appl. Environ. Microbiol. 82, 4560–4569. doi:10.1128/AEM.00231-16.

Orellana, L.H., Rodriguez-R, L.M., Higgins, Chee-Sanford, S. J.C., Sanford, R.A., Ritalahti, K.M., Löffler, F.E., Konstantinidis, K.T. (2014). Detecting Nitrous Oxide Reductase (nosZ) Genes in Soil Metagenomes: Method Development and Implications for the Nitrogen Cycle, MBio. 5 e01193-14. doi:10.1128/mBio.01193-14.

Parsaee, M., Kiani Deh Kiani, M., and Karimi, K. (2019). A review of biogas production from sugarcane vinasse. Biomass and Bioenergy 122, 117–125. doi:10.1016/j.biombioe.2019.01.034.

Penatti CP, Cambria S, Boni PS, Arruda FCO, Manoel LA. Efeitos da aplicação de vinhaça e nitrogênio na soqueira da cana-de-açúcar. Boletim Técnico Copersucar, v.44; 1988. pp. 32–38.

Pollard, K. S., Sandrine, D., van der Laan, M.J. Multiple testing procedures: the multtest package and applications to genomics. In: Bioinformatics and computational biology solutions using R and bioconductor. New York: Springer, 2005. pp. 52.

Prasad, M.N.V., Shih, K. eds., Environmental materials and waste: resource recovery and pollution prevention, Elsevier/AP, Academic Press is an imprint of Elsevier, Amsterdam; Boston, 2016.

Ryan, J. A., Sims, J. L., and Peaslee, D. E. (1972). Effect of Phosphate and Chloride Salts on Ammonification in Waterlogged Soils. Soil Sci. Soc. Am. J. 36, 915–917. doi:10.2136/sssaj1972.03615995003600060023x.

Sanford, R. A., Wagner, D. D., Wu, Q., Chee-Sanford, J. C., Thomas, S. H., Cruz-García, C., et al. (2012). Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. Proc. Natl. Acad. Sci. U. S. A. 109, 19709–19714. doi:10.1073/pnas.1211238109.

Sauder, L. A., Albertsen, M., Engel, K., Schwarz, J., Nielsen, P. H., Wagner, M., et al. (2017). Cultivation and characterization of Candidatus Nitrosocosmicus exaquare, an ammonia-oxidizing archaeon from a municipal wastewater treatment system. ISME J. 11, 1142–1157. doi:10.1038/ismej.2016.192.

Schlüter, S., Henjes, S., Zawallich, J., Bergaust, L., Horn, M., Ippisch, O., et al. (2018). Denitrification in soil aggregate analogues-effect of aggregate size and oxygen diffusion. Front. Environ. Sci. 6, 1–10. doi:10.3389/fenvs.2018.00017.

Schroder, J. L., Zhang, H., Girma, K., Raun, W. R., Penn, C. J., and Payton, M. E. (2011). Soil Acidification from Long-Term Use of Nitrogen Fertilizers on Winter Wheat. Soil Sci. Soc. Am. J. 75, 957–964. doi:10.2136/sssaj2010.0187.

Shi, X., Hu, H. W., Müller, C., He, J. Z., Chen, D., and Suter, H. C. (2016). Effects of the nitrification inhibitor 3,4-dimethylpyrazole phosphate on nitrification and nitrifiers in two contrasting agricultural soils. Appl. Environ. Microbiol. 82, 5236–5248. doi:10.1128/AEM.01031-16.

Shi, Z., Yin, H., Van Nostrand, J. D., Voordeckers, J. W., Tu, Q., Deng, Y., et al. (2019). Functional Gene Array-Based Ultrasensitive and Quantitative Detection of Microbial Populations in Complex Communities. mSystems 4, 1–16. doi:10.1128/msystems.00296-19.

Shoun, H., Fushinobu, S., Jiang, L., Kim, S. W., and Wakagi, T. (2012). Fungal denitrification and nitric oxide reductase cytochrome P450nor. Philos. Trans. R. Soc. B Biol. Sci. 367, 1186–1194. doi:10.1098/rstb.2011.0335.

Šimek, M., Virtanen, S., KrištUfek, V., Simojoki, A., and Yli-Halla, M. (2011). Evidence of rich microbial communities in the subsoil of a boreal acid sulphate soil conducive to greenhouse gas emissions. Agric. Ecosyst. Environ. 140, 113–122. doi:10.1016/j.agee.2010.11.018.

Smith, C. J., Dong, L. F., Wilson, J., Stott, A., Osborn, A. M., and Nedwell, D. B. (2015). Seasonal variation in denitrification and dissimilatory nitrate reduction to ammonia process rates and corresponding key functional genes along an estuarine nitrate gradient. Front. Microbiol. 6, 1–11. doi:10.3389/fmicb.2015.00542.

Suleiman, A. K. A., Lourenço, K. S., Pitombo, L. M., Mendes, L. W., Roesch, L. F. W., Pijl, A., et al. (2018). Recycling organic residues in agriculture impacts soil-borne microbial community structure, function and N2O emissions. Sci. Total Environ. 631–632, 1089–1099. doi:10.1016/j.scitotenv.2018.03.116.

Tejada, M., and Gonzalez, J. L. (2006). Effects of two beet vinasse forms on soil physical properties and soil loss. Catena 68, 41–50. doi:10.1016/j.catena.2006.04.025.

Tourna, M., Stieglmeier, M., Spang, A., Könneke, M., Schintlmeister, A., Urich, T., et al. (2011). Nitrososphaera viennensis, an ammonia oxidizing archaeon from soil. Proc. Natl. Acad. Sci. U. S. A. 108, 8420–8425. doi:10.1073/pnas.1013488108.

Varon-Lopez, M., Dias, A. C. F., Fasanella, C. C., Durrer, A., Melo, I. S., Kuramae, E. E., et al. (2014). Sulphur-oxidizing and sulphate-reducing communities in Brazilian mangrove sediments. Environ. Microbiol. 16, 845–855. doi:10.1111/1462-2920.12237.

Wang, J., Chen, Z., Xiong, Z., Chen, C., Xu, X., Zhou, Q., et al. (2015). Effects of biochar amendment on greenhouse gas emissions, net ecosystem carbon budget and properties of an acidic soil under intensive vegetable production. Soil Use Manag. 31, 375–383. doi:10.1111/sum.12202.

White, C. J., and Lehnert, N. (2016). Is there a pathway for N2O production from hydroxylamine oxidoreductase in ammonia-oxidizing bacteria? Proc. Natl. Acad. Sci. U. S. A. 113, 14474–14476. doi:10.1073/pnas.1617953114.

Whittaker, M., Bergmann, D., Arciero, D., Hooper, A. B. (2000). Electron transfer during the oxidation of ammonia by the chemolithotrophic bacterium Nitrosomonas europaea, Biochimica et Biophysica Acta (BBA) - Bioenergetics. 1459 346–355. doi:10.1016/S0005-2728(00)00171-7.

Wilkie, A.C., Riedesel, K.J., Owens, J.M. (2000). Stillage characterization and anaerobic treatment of ethanol stillage from conventional and cellulosic feedstocks, Biomass and Bioenergy. 19, 63–102. doi:10.1016/S0961-9534(00)00017-9.

Wu, D., Senbayram, M., Well, R., Brüggemann, N., Pfeiffer, B., Loick, N., et al. (2017). Nitrification inhibitors mitigate N2O emissions more effectively under straw-induced conditions favoring denitrification. Soil Biol. Biochem. 104, 197–207. doi:10.1016/j.soilbio.2016.10.022.

Wu, L., Han, C., Zhu, G., and Zhong, W. (2019). Responses of active ammonia oxidizers and nitrification activity in eutrophic lake sediments to nitrogen and temperature. Appl. Environ. Microbiol. 85. doi:10.1128/AEM.00258-19.

Wu, L., Liu, X., Schadt, C. W., and Zhou, J. (2006). Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. Appl. Environ. Microbiol. 72, 4931–4941. doi:10.1128/AEM.02738-05.

Yoon, S., Nissen, S., Park, D., Sanford, R. A., and Löffler, E. (2016). Clade I NosZ from Those Harboring Clade II NosZ. Appl. Envioronmental Microbiol. 82, 3793–3800. doi:10.1109/ROSE.2014.6952980.

Yu, X., and Labaer, J. (2015). High-throughput identification of proteins with AMPylation using self-assembled human protein (NAPPA) microarrays. Nat. Protoc. 10, 756–767. doi:10.1038/nprot.2015.044.

Zhao, L., Liu, Y., Wang, Z., Yuan, S., Qi, J., Zhang, W., et al. (2020). Bacteria and fungi differentially contribute to carbon and nitrogen cycles during biological soil crust succession in arid ecosystems. Plant Soil 447, 379–392. doi:10.1007/s11104-019-04391-5.

Zhao, R., Chen, Y., Qu, J., Jin, P., Zheng, Z., and Cui, Z. (2019). Insights into the variations of hao-dependent nitrifying and nir-dependent denitrifying microbial communities in ammonium-graduated lake environments. Appl. Sci. 9, 1–15. doi:10.3390/app9163229.

Zhou, H., Zhang, Q., Gu, C., Jabeen, S., Li, J., and Di, H. (2017). Soil anammox community structure in different land use soils treatment with 13C urea as determined by analysis of phospholipid fatty acids. Appl. Microbiol. Biotechnol. 101, 6659–6669. doi:10.1007/s00253-017-8404-4.

3. Chapter 2. Acidobacteria subgroups and their metabolic potential for carbon degradation in sugarcane soil amended with vinasse and nitrogen fertilizers

Abstract

Acidobacteria is a predominant bacterial phylum in tropical agricultural soils, including sugarcane cultivated soils. The increased need for fertilizers due to the expansion of sugarcane production is a threat to the ability of the soil to maintain its potential for self-regulation in the long term, in which carbon degradation has an essential role. In this study, a culture-independent approach based on high-throughput DNA sequencing and microarray technology was used to perform taxonomic and functional profiling of the Acidobacteria community in a tropical soil under sugarcane (*Saccharum* spp.) that was supplemented with nitrogen (N) combined with vinasse. These analyses were conducted to identify the subgroup-level responses to chemical changes and the carbon (C) degradation potential of the different Acidobacteria subgroups. Eighteen Acidobacteria subgroups from a total of 26 phylogenetically distinct subgroups were detected based on high-throughput DNA sequencing, and 16 gene families associated with C degradation were quantified using Acidobacteria-derived DNA microarray probes. The subgroups Gp13 and Gp18 presented the most positive correlations with the gene families associated with C degradation, especially those involved in hemicellulose degradation. However, both subgroups presented low abundance in the treatment containing vinasse. In turn, the Gp4 subgroup was the most abundant in the treatment that received vinasse but did not present positive correlations with the gene families for C degradation analyzed in this study. The metabolic potential for C degradation of the different Acidobacteria subgroups in sugarcane soil amended with N and vinasse can be driven in part through the increase in soil nutrient availability, especially calcium (Ca), magnesium (Mg), potassium (K), aluminum (Al), boron (B) and zinc (Zn). This soil management practice reduces the abundance of Acidobacteria subgroups, including those potentially involved with C degradation in this agricultural soil.

Keywords: soil metagenome, DNA microarray, mineral and organic fertilizers, carbon cycling, microbe mediated process in soil

3.1. Introduction

Acidobacteria is among the most widespread bacterial phyla that occur in soils around the world, including the tropical soils under sugarcane *Saccharum* spp. (Rachid et al., 2013; Navarrete et al., 2015a; Val-Moraes et al., 2016). The presence of membrane transporters and the use of carbon (C) sources ranging from simple sugars to more complex substrates, such as hemicellulose, cellulose, and chitin, are among the genomic and physiological characteristics that may contribute to the survival and growth of Acidobacteria in soil (Ward et al., 2009;

Rawat et al., 2012). Kielak et al. (2016) recently reviewed the genomic and physiological characteristics of Acidobacteria and showed that there are still many gaps in understanding the functional role of this bacterial phylum in the soil C degradation process. Despite this lack of biological and ecological information for Acidobacteria, previous studies in agricultural soils have shown that both microbial C degradation processes and acidobacterial community can be affected by soil management (Craine et al., 2007; Navarrete et al., 2015a; Omori et al., 2016; Wang et al., 2018; Lian et al., 2019).

The soil management practices used in sugarcane cultivation require synthetic mineral fertilizers nitrogen/phosphorus/potassium-NPK (Heffer and Prud'homme, 2008), micronutrients and complete recycling of byproducts of the ethanol and sugar production in sugarcane production fields in the form of organic fertilizer (Mutton et al., 2010). Vinasse is a byproduct of the ethanol industry produced at a ratio of ten to eighteen liters for each liter of ethanol produced (Gasparotto et al., 2014). The chemical composition of vinasse varies with the ethanol plant in which it was generated and the distillation process, although it generally consists of water (93%) and organic and mineral compounds (7%) (Christofoletti et al., 2013). Vinasse has high levels of organic matter but low concentrations of N (0.97 to 4.75 g L^{-1}) and P (1 to 190 mg L^{-1}) and high C:N ratio (Moran-Salazar et al., 2016). The main non-aqueous component of vinasse is organic matter in the form of glycerol, organic acids, and yeast (Christofoletti et al., 2013). Depending on the most used in the sugarmill fermentation process, vinasse has also high concentration of potassium (K), calcium (Ca), and sulphur (S), medium concentration of magnesium (Mg) and micronutrients (Christofoletti et al., 2013). Since the 1960s, vinasse has been used as a fertilizer in sugarcane production fields to solve the ecological problem of its disposal in water sources like rivers and lagoons. Studies from the 1980s have recommended the use of N fertilizer in combination with vinasse in sugarcane fields with high productivity.

The addition of vinasse in soils under sugarcane causes changes in the soil microbial community and chemical processes (Barros et al., 2010; Ribeiro et al., 2012; Jiang et al., 2012; Montenegro et al., 2009), including organic matter decomposition (Resende et al., 2006). Omori et al. (2016) reported increased bacterial diversity after the application of vinasse to the soil under sugarcane and showed that the Acidobacteria subgroups Gp3 and Gp4 were more abundant in soil fertilized with vinasse. In turn, Navarrete et al. (2015a) showed that the Acidobacteria subgroups Gp4, Gp11, Gp17, Gp21, and Gp25 were positively related to chemical factors of the soil fertilized with N and vinasse compared with soils fertilized with N

alone and soils without N fertilizer and vinasse. Soils fertilized with N and vinasse usually present high levels of sulfur S, K, and total C and increased pH (Glória and Orlando Filho, 1983). The addition of N as fertilizer may decrease recalcitrant C decomposition (Craine et al., 2007; Wang et al., 2018), which may affect members of the bacterial community that act as decomposers and plays a vital role in the C cycle.

Because of the substantial effects that soil agricultural management has on carbon-degrading microorganisms in soils, we would like to obtain better insight into the Acidobacteria community in sugarcane cultivated soil. For this purpose, the present study was designed to evaluate the response of Acidobacteria subgroups to the addition of N and vinasse in tropical soil under sugarcane and the metabolic potential of the subgroups in the degradation of C in these soils. For this purpose, soil genomic DNA shotgun sequencing was performed to identify the Acidobacteria subgroups. A high-throughput functional gene array termed 'GeoChip v. 5.0M' was used for the large-scale quantification of the functional microbial genes associated with C degradation. While soil genomic DNA shotgun sequencing-based approach is useful the detection of Acidobacteria subgroups based on the taxonomic identification of soil metagenome sequences (Navarrete et al., 2015a), high-throughput functional gene array is a powerful and high-performance tool to analyze specific genes associated with microbe-mediated processes in different habitats (Zhou et al., 2008). Soil chemical analyses were performed, and the results were correlated with the abundance of Acidobacteria subgroups and functional genes associated with soil C degradation using statistical and computational methods.

3.2. Material and Methods

3.2.1. Mesocosm experiment and soil sampling

A greenhouse experiment with sugarcane *Saccharum* spp. variety CTC-02 was performed over 150 days between April and December 2013 to normalize certain environmental parameters, such as the moisture regime and soil type. This sugarcane variety has medium to late maturity and presents high productivity and longevity, and the seedlings used in this study were obtained by *in vitro* fertilization and tissue culture. The soil used in the experiment was a clayey-loamy dark red podzolic soil according to the Brazilian Soil Classification collected in the 0-20 cm layer of the Areião Farm, which belongs to the Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo ESALQ-USP in Piracicaba, São Paulo, Brazil 22 42' 30" S; 47 38' 00" W. On a 15 cm layer of washed crushed stones, 90 kg of soil was added to

each of the nine mesocosms 100 L pots (Navarrete et al., 2015a), which received mineral fertilization at planting consisting of 150 kg ha^{-1} of P_2O_5 triple superphosphate and 80 kg ha^{-1} of KCl potassium chloride. These mineral fertilizers were incorporated within the soil using a paddle mixer. Three treatments with three replicates each were established according to the N fertilizer source: N0 Control, no N fertilizer; N60, N fertilizer in the form of urea; and NV, urea supplemented with vinasse. Urea (450 g of N kg^{-1}) was added in the 5-10 cm soil layer at a dose of 60 kg ha^{-1}, and it was immediately mixed into the soil to prevent volatilization. Vinasse was used to irrigate the soil surface at a dose of 0.06 L kg^{-1} (120 m^3 ha^{-1}) as a source of K in addition to organic matter and other nutrients. The treatments that did not receive V received an equivalent volume of water. The KCl dosage was calculated minus the equivalent input of K in the case of V treatment according to previous measurements of K content in vinasse samples. The soil moisture content in the pots was maintained at 80% of field capacity throughout the experiment using a moisture sensor (Extech MO750, Nashua, NH, USA). Initially, ten sugarcane seedlings were placed in each mesocosm to ensure a rapid influence of the plants on the soil microbiota, and they were periodically removed in pairs to keep the root system below the pot capacity limit, with bulk soil and rhizosphere detachable.

For each mesocosm, three bulk soil samples (about 100 g each sample) were collected from topsoil layer (0 to 10 cm) at equidistant positions within an equilateral triangle (ray equivalent to 1/3 of the circular surface area of the pot) using sterile PVC tubes (15 cm in length and 5 cm in diameter) for both chemical and molecular analyses. These samples were collected on the 7th and 150th days after fertilizer application based on the maximum and minimum carbon dioxide (CO_2) and nitrous oxide (N_2O) emissions from soil (Navarrete et al., 2015a). The samples were immediately processed after collection for the chemical analyses. For molecular analysis, one subsample was taken from each of the three bulk soil samples for each mesocosm after undeform each one in a plastic bag separately, and they were transported to the Laboratory of Cellular and Molecular Biology of Centro de Energia Nuclear na Agricultura CENA-USP, stored at -20 °C and processed within 72 h.

3.2.2. Analysis of soil chemical factors

The soil samples were air dried at room temperature and sieved through a 0.149-mm sieve to determine the total C and N by dry combustion using a LECO CN 2000 elemental analyzer (PerkinElmer, Waltham, MA, USA). The soil fertility factors analyzed were as

follows: pH, potential acidity hydrogen (H + aluminum (Al)), Ca, K, Mg, P, S, micronutrients - iron (Fe), manganese (Mn), zinc (Zn), copper (Cu) and boron (B), exchangeable bases, cation exchange capacity (CEC) and base saturation (BS).

The chemical factors of each soil sample were determined according to Embrapa (2011). Soil pH was determined in a soil/0.01 M $CaCl_2$ 1:5 suspension. Boron was obtained by hot water extraction. Al, Ca and Mg were extracted with 1 M potassium chloride. Ca and Mg were determined by spectrometric atomic absorption, whereas Al was determined by acid-base titration. Available P and K were extracted by ion-exchange resin and determined by colorimetry and atomic emission spectroscopy, respectively. The combined results were used to calculate the sum of exchangeable bases of Ca, Mg and K, CEC sum of Ca, Mg, K, Al and H, BS, percentage ratio between BS and CEC and potential acidity H + Al using an equation based on the Shoemaker-McLean-Pratt SMP pH-buffer method.

3.2.3. Extraction and sequencing of total soil genomic DNA

Genomic DNA was extracted from 250 mg of soil obtained from each subsample (avoiding small pieces of roots) using the PowerLyzer® PowerSoil® DNA Isolation kit (MoBio Laboratories, Carlsbad, CA, USA) following the manufacturer's instructions. The DNA isolated from each of the three soil subsamples per mesocosm was then pooled and concentrated using the Genomic DNA Clean and Concentrator kit (Zymo Research Corporation, Irvine, CA, USA), constituting a single DNA sample per mesocosm. The concentrated DNA was resuspended in 20 μL of PCR water and both purity and quality of the genomic DNA were assessed via spectrophotometry on a NanoDrop apparatus (NanoDrop® ND-1000 NanoDrop Technologies, Inc., Wilmington, DE, USA) to determine the absorbance at the following wavelengths: 230, 260, 280 and 320 nm. The DNA concentration was determined with the Quant-iT PicoGreen kit (Molecular Probes/Invitrogen, Carlsbad, CA, USA). DNA was stored at -20ºC until use.

Eighteen DNA sequencing libraries were prepared using the Illumina Nextera sample preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. The libraries were evaluated on 2100 Bioanalyzer using High Sensitivity DNA kit (Agilent, Santa Clara, CA, USA) to estimate the library size. Libraries were quantified using Qubit dsDNA HS kit on a Qubit 2.0 fluorometer (Life technologies, Carlsbad, CA, USA) and KAPA SYBR FAST qPCR Master mix and Illumina standards and primer premix (KAPA Biosystems, Wilmington, MA, USA) according to the Illumina suggested protocol. The resulting DNA libraries were denatured with NaOH, diluted to 8 pM in Illumina's HT1 buffer, and spiked with

1 % PhiX. Equal concentration of libraries was loaded on MiSeq Reagent v2 sequencing reagent kit (Illumina, San Diego, CA, USA). The equipment used for shotgun metagenomic sequencing was a MiSeq Personal Sequencing System by Illumina (Illumina, San Diego, CA, USA) operated in Rapid Run Mode to generate 2 x 250 pb paired-end reads. In summary, we captured an average of 105.5 MB of genomic sequences per sample.

3.2.4. Shotgun metagenomic data processing and taxonomic identification of Acidobacteria subgroups

First, paired-end reads were merged using *FLASH* v. 1.2.5 (Magoč and Salzberg, 2011) to produce consensus sequences. Quality control of the consensus sequences was performed using the *Phred* quality score 20 to each base call (Ewing and Green, 1998) with an executable default script in *SeqClean* (http://www.bioinformatics.org/), and the low-quality bases were removed. Shotgun sequencing of soil DNA from samples resulted in approximately 547.415 thousand merged sequence reads and 375.429 thousand not merged sequence reads after the quality-based filtering procedure. The unmatched trimmed sequences were concatenated in a single file for the metagenomic dataset, which is available on the Metagenomics Rapid Annotation (MG-RAST) server under the "Metagenome of sugarcane soil - CENA/USP" project with accession numbers from 4582104.3 to 4582153.3.

Sequencing data were analyzed using MG-RAST software version 3.2 (http://metagenomics.anl.gov) (Meyer et al. 2008) with default settings to identify the sequences belonging to the Archaea and Bacteria domains and the fungi group. The taxonomic identification of the Acidobacteria community at the class taxonomic level subgroups was carried out using FOCUS software (Silva et al., 2014) with default settings, which uses a database of bacterial genomes. Initially, a table of relative abundance of bacteria occurrence was generated for each dataset at the phylum taxonomic level. Next, the 18 metagenomes were aligned using high scoring pairs in BLASTN 2.2.28 (Altschul et al., 1990) and an *E*-value threshold of 10^{-5} similarity value gene sequences against a database of Acidobacteria 16S rRNA extracted from the Ribosomal Database Project (RDP) release 11 (http://rdp.cme.msu.edu/) (Cole et al., 2014) for the identification of Acidobacteria subgroups. Only the best hit for each query sequence was used in the count. Relative abundances of Acidobacteria subgroups were estimated by dividing the number of sequences classified as the Acidobacteria subgroups by the total number of sequences classified as Acidobacteria per sample.

3.2.5. Purified DNA labeling and image scanning for GeoChip analysis

Each of the 18 samples of the concentrated DNA was purified and labeled with Cy-3 fluorescent dye according to Wu et al. (2006). Briefly, the DNA 600 ng was mixed with random primers 300 ng mL^{-1}, denatured at 99.9°C for 5 min and immediately cooled on ice. A solution containing 5 mM dAGC-TP, 2.5 mM dTTP, Klenow fragment 40 U and Cy-3 dUTP 25 nM was added to the denatured DNA, and the reaction volume was adjusted to 50 µL with H_2O. The labeling solution was incubated at 37°C for 6 h followed by 3 min at 95°C. The labeled DNA was purified with the QIAquick Kit Qiagen, Valencia, CA, USA, and dye incorporation was confirmed with a NanoDrop spectrophotometer (NanoDrop® ND-1000 NanoDrop Technologies, Inc., Wilmington, DE, USA) using the absorption spectra of the standard solution for Cy-3. The labeled DNA samples were dried under vacuum and stored at -20°C until hybridization.

GeoChip v. 5.0M, which is manufactured by Agilent (Agilent Technologies Inc., Santa Clara, CA, USA), in the 4x180 K format was used (Cong et al., 2015). The chip contains 167,044 distinct probes belonging to different gene categories involved in C metabolism e.g., C degradation, C fixation, methane production, N metabolism ammonification, nitrification, N fixation, S metabolism, P cycling, metal homeostasis Zn transport, secondary metabolism antibiotic, pigments, stress responses oxidative, virulence infection and other microbial genes of known function. Here, the gene sequence refers to each unique sequence targeted by GeoChip 5.0M, and a gene family consists of all gene sequences that are assigned the same name e.g., called AceA and encode the same class of proteins. The probes originated from bacteria, archaea, fungi and viruses (both bacteriophages and eukaryotic viruses). Prior to hybridization, the labeled DNA was resuspended in 27.5 mL of DNase/RNase-free distilled water and mixed with 99.4 mL of the hybridization solution containing 2 x Hi-RPM hybridization buffer (Agilent Technologies Inc., Santa Clara, CA), 10 x aCGH blocking agent (Agilent Technologies Inc., Santa Clara, CA, USA), 10% formamide (JT Baker, Philipsburg, NJ, USA), 0.05 mg/mL of Cot^{-1} DNA (Agilent, Technologies Inc., Santa Clara, CA, USA) and universal standard DNA labeled with Cy5 dye 10 pM for standardization (Liang et al., 2010). The solution was denatured at 95°C for 3 min, incubated at 37°C for 30 min and hybridized with GeoChip v. 5.0M probes. The hybridization was carried out at 67°C in the presence of 10% formamide in a hybridization oven (Agilent Technologies Inc., Santa Clara, CA, USA) for 24h. After hybridization, the slides were washed using Agilent Wash Buffers 1 and 2 following the manufacturer's protocol. Next, the gene arrays were scanned with

100% laser power and 75% photomultiplier tube using a NimbleGen MS200 microarray scanner (Roche NimbleGen, Madison, WI, USA). The image data were extracted using Agilent Feature Extraction software v. 2.6. The alignment of points and determination of signal intensity was performed with Agilent Feature Extraction software v. 11.5 (Agilent Technologies Inc., Santa Clara, CA, USA). Data were submitted to the Microarray Data Manager available at the Institute for Environmental Genomics website http://www.ou.edu/content/ieg/tools/data-analysis-pipeline.html and analyzed using the following parameters: i. removal of points with a signal intensity lower than 1.3 or with a signal-to-noise ratio SNR below 2.0 SNR = spot signal intensity - background mean / background SD; ii. removal of nonrepresentative data i.e., singletons, or positive probes in only one sample from an experimental group using the cutoff of group 1 by default; iii. normalization performed by dividing the signal intensity of each point by the mean of the universal standard spots and then by the mean signal intensities in each sample i.e., mean ratio approach.

3.2.6. Quantification of gene families encoding for C degradation using Acidobacteria-derived probes

Quantification of gene families encoding for C degradation using Acidobacteria-derived probes was performed based on data obtained with GeoChip analysis for each of the 18 DNA samples. Among the 167,044 probes available in GeoChip v.5.0M, 7,592 probes are belonging to the gene families with the "carbon degradation" function. Of this total, 100 probes with the "carbon degradation" function originated from Acidobacteria (Examples: *Solibacter usitatus* Ellin6076, *Acidobacterium capsulatum* ATCC51196 and *Granulicella mallensis* MP5ACTX8, among others). These 100 probes hybridize to 16 gene families xylanase, ara (alpha - L-arabinofuranosidase), xyla (xylose isomerase), endoglucanase, cellobiase, RgaE (acetylesterase), pme (pectinesterase), pectinase, cda (alpha amylase), glucoamylase, pula (alpha-1,6-glucosidade), amyA (alpha amylase), AceA (isocitrate liase), AceB (malato synthase), chitinase and acetylglucosaminidase from the category "carbon cycle", subcategory 1 "carbon degradation", subcategory 2 "hemicellulose", "cellulose", "pectin", "starch", "glyoxalate cycle" and "chitin". In this study, only the hybridization signals with these 100 probes were considered for quantification of gene families encoding for C degradation using Acidobacteria-derived probes in the different DNA samples from the three different treatments.

3.2.7. Statistical analyses of taxonomic and functional data and soil chemical factors

Analysis of variance ANOVA and Tukey's test were applied to the taxonomic Acidobacteria subgroups and functional gene families associated with C degradation data and soil chemical factors using STATISTICA software v. 13 (StatSoft Inc., Palo Alto, CA, USA). The explicit relationships among these variables were examined by constrained ordination generated by redundancy analysis RDA. The RDA was performed using CANOCO Software version 4.5. (Ter Braak and Smilauer, 1998) based on the Monte Carlo permutation test at 5% significance.

3.2.8. Construction of integration networks of taxonomic and functional data and soil chemical factors

To visualize connections among the different Acidobacteria subgroups, functional genes related to carbon degradation, and soil chemical factors, co-occurrence network was constructed. *Spearman* rank's correlation coefficients were calculated ($P < 0.05$) between: *i*. relative abundance of Acidobacteria subgroups and soil chemical factors, *ii*. relative abundance of Acidobacteria and functional genes, and *iii*. functional genes and soil chemical factors using SigmaPlot software v.14.0 SYSTAT (Software Inc., California, USA). *Spearman* rank's correlation coefficients from all dataset combinations with $P < 0.05$ were used to network construction. The topology number of layers, units in each layer, training algorithm parameters and activation functions were determined (Brandes, 2001; Blondel, 2008). The network constructed was integrated to the different application modules and analyzed using *Gephi* software (Bastian et al., 2009), which has a graphical interface.

3.3. Results

3.3.1. Soil chemical factors

The addition of vinasse in combination with N fertilizer at the doses used in this study promoted significant changes ($P \leq 0.05$) in the levels of most soil chemical factors (Table 1). In particular, non-acid cations provided by the organic residue, K, Mg and the values of SB, CEC and BS were higher at 150 days in comparison with 7 days after the addition of fertilizers in soil amended with N and vinasse. In turn, the H + Al value decreased and pH increased in

this circumstance. The total N, total C, and OM values also increased in NV treatment at 7 days to 150 days of the experiment. However, OM revealed lower values for all treatments at 150 days than at 7 days after the addition of fertilizers (Table 1). The S, B, Mn, and Zn levels were high to that of the treatments that did not receive vinasse at 150 days after the addition of fertilizer.

Table 1. Chemical compounds present in the soil treatments without nitrogen fertilization (N0), with nitrogen fertilization (N60) and nitrogen fertilization combined with vinasse (NV) collected at seven (T7) and one hundred and fifty (T150) days after the start of the experiment

Soil properties		T7			T150			T7 vs. T150		
	Unit	N0	N60	NV	N0	N60	NV	N0	N60	NV
C	g. kg^{-1}	2.24 ± 0.04	2.29 ± 0.03	2.59 ± 0.07**	2.24 ±0.05	2.29 ± 0.03	2.59± 0.07**			
N	g. kg^{-1}	0.18 ± 0.00	0.19 ± 0.00	0.26± 0.02**	0.18 ± 0.00	0.19 ± 0.07	0.26 ± 0.02**			
pH		5.23 ± 0.06	5.23 ± 0.15	5.60± 1.09**	5.50 ± 0.10	5.10 ± 0.00**	5.80± 0.10*	0.016		0.025
O.M.	$g.dm^{-3}$	38.3 ± 1.53	36.3 ± 0.58	39.7 ± 3.79	32.3 ± 1.53	32.0 ± 1.00	35.7 ± 0.57*	0.008	0.003	
P	$mg.dm^{-3}$	61.0 ± 17.7	99.6*± 12.6	90.0 ± 12.1	259 ± 94.9	248 ± 83.5	35.7 ± 0.57	0.023	0.038	
S	$mg.dm^{-3}$	6.33 ± 1.53	11.7 ± 0.58	184 ± 16.5**	18.3 ± 7.64	17.3 ± 3.51	237± 22.3**			0.028
K	$mmolc.dm^{-3}$	1.23 ± 0.06	1.13 ± 0.06	11.4± 0.92**	3.60 ± 0.50	9.93 ± 2.72	39.0 ± 12.0**	0.001	0.005	0.014
Ca	$mmolc.dm^{-3}$	54.0 ± 2.64	55.0 ± 1.00	48.7 ± 3.50	71.0 ± 2.64	56.0 ± 16.6	66.3 ± 10.1	0.002		0.046
Mg	$mmolc.dm^{-3}$	16.7 ± 1.52	17.3 ± 0.57	18.7 ± 0.58	17.3 ± 0.60	14.0 ± 1.73	22.3± 2.08*		0.034	0.043
H+Al	$mmolc.dm^{-3}$	43.7 ± 2.88	44.0 ± 5.19	34.0± 0.00*	32.0 ± 1.73	39.3± 2.30*	25.0± 3.00*	0.004		0.007
CEC	$mmolc.dm^{-3}$	115 ± 6.22	117 ± 3.42	113 ± 4.90	124 ± 4.53	119 ± 20.3	153 ± 6.75			0.001
C:N	%	12.0 ± 0.30	11.9 ± 0.55	9.93± 0.52**	12.0 ± 0.37	12.0 ± 0.55	9.93 ± 0.52**			
SB	%	61.7 ± 1.15	62.3 ± 3.21	70.0 ± 1.00**	74.0 ± 1.00	66.7 ± 4.93*	84.0 ± 1.50	0.000		0.000
B	$mg.dm^{-3}$	0.24 ± 0.04	0.22 ± 0.00	0.14 ± 0.03*	0.29 ± 0.03	0.28 ± 0.02	0.37 ± 0.07		0.022	0.007
Cu	$mg.dm^{-3}$	0.93 ± 0.06	0.93 ± 0.06	0.93 ± 0.12	0.83 ± 0.06	0.93 ± 0.12	0.90 ± 0.20			
Fe	$mg.dm^{-3}$	37.0 ± 3.60	37.0 ± 4.00	48.3 ± 19.6	27.7 ± 2.08	38.0 ± 4.36	29.3 ± 7.51	0.018		
Mn	$mg.dm^{-3}$	7.40 ± 0.50	7.47 ± 1.12	21.0 ± 3.78**	6.27 ± 0.81	9.37 ± 1.04	21.0 ± 11.0			
Zn	$mg.dm^{-3}$	2.00 ± 0.50	1.70 ± 0.10	2.26 ± 1.48**	10.9 ± 4.67	10.8 ± 4.90	47.0 ± 26.3	0.030	0.032	0.042

N0. Treatment without the addition of nitrogen fertilizer; N60. Treatment with nitrogen at the dose of 60 kg ha^{-1}; and NV. Treatment with nitrogen at the dose of 60 kg $há^{-1}$ and the addition of vinasse at a rate of 0.06 L kg^{-1}. a = average for each of three replicates soil. b = standard deviations of the average for each of three replicates soil. The analysis were perform at seven (T7) and one hundred and fifty days (T150) after the beginning of the experiment. Analysis of variance and Tukey's HSD test two-pair combination with a significance level of * ($p < 0.05$) and ** ($p < 0.001$) in relation of N0. Blank spaces indicate that the values were not significative.

3.3.2. Relative abundance of Acidobacteria subgroups

The taxonomic identification of soil metagenome sequences resulted in the detection of 18 different Acidobacteria classes (Table 2) from a total of 26 subgroups – 1 to 8 according to Hugenholtz et al. (1998), subgroups 9 to 11 according to Zimmermann et al. (2005), and subgroups 12 to 26 according to Barns et al. (2007) – which demonstrates good representativeness of Acidobacteria in the soil used in this study in comparison with other tropical soils from previous studies (Navarrete et al., 2013; Catão et al., 2014; Navarrete et al., 2015ab). Most of the subgroups showed a decrease in abundance in the NV treatment; however, the differences between sampling times T7 and T150 were not statistically significant (Table 2). Consistent with previous studies, Gp4 was the most abundant class in the treatment that received vinasse while Gp9 and Gp18 were the least represented in the NV treatment.

In treatment C, the Gp6 and Gp17 subgroups were the most abundant. However, Gp6 showed a decrease in abundance of more than 50% in treatment NV, whereas Gp17 significantly increased its abundance in the treatments that received N fertilization N and NV. For subgroup Gp2, N fertilization does not seem to favor the occurrence of representatives of this Acidobacteria subgroup in the soil because their abundance was reduced in the N and NV treatments. The Gp13 subgroup showed low representativeness in the C treatment but was responsive to N fertilization and showed increased abundance in the N treatment but not in the NV treatment.

Table 2. Abundance of Acidobacteria subgroups relative to total Acidobacteria community in the soil mesocosms in the treatments without nitrogen fertilization (C), with nitrogen fertilization (N) and with nitrogen fertilization combined with vinasse (NV) over 7 days (T7) and 150 days (T150)

Acidobacteria subgroups	T7			T150			Statistics								
							T7			T150			T7 *vs* T150		
	C	N	NV	C	N	NV	C *vs* N	C *vs* NV	N *vs* NV	C *vs* N	C *vs* NV	N *vs* NV	C	N	NV
Gp1	1.80[a] ± 1.5[b]	4.63 ± 1.8	3.69 ± 0.3	2.04 ± 1.9	3.83 ± 0.1	3.37 ± 0.7									
Gp2	6.43 ± 1.7	2.21 ± 0.6	3.42 ± 1.1	2.37 ± 2.3	3.04 ± 0.8	1.73 ± 1.1	0.014								
Gp3	6.24 ± 3.0	5.37 ± 0.6	5.42 ± 1.0	6.65 ± 2.7	5.77 ± 0.5	5.25 ± 2.2									
Gp4	9.08 ± 2.3	11.1 ± 4.0	12.1 ± 2.2	10.7 ± 2.8	6.44 ± 0.9	13.1 ± 0.3				0.05		0.007			
Gp5	1.82 ± 0.5	1.37 ± 0.5	1.92 ± 1.9	1.04 ± 1.8	1.58 ± 0.5	1.45 ± 0.9									
Gp6	13.5 ± 1.5	11.1 ± 2.1	11.5 ± 2.1	14.3 ± 5.6	12.6 ± 1.6	6.94 ± 2.3									
Gp7	7.41 ± 2.6	8.85 ± 2.1	6.45 ± 2.3	6.23 ± 2.2	4.79 ± 2.5	6.13 ± 1.2									
Gp9	1.01 ± 0.3	0.64 ± 1.1	0.12 ± 0.2	0.52 ± 0.9	1.10 ± 0.4	0.35 ± 0.3									
Gp10	7.50 ± 1.8	6.82 ± 1.4	7.98 ± 2.6	7.43 ± 4.3	9.24 ± 0.5	6.81 ± 0.9								0.05	
Gp11	0.55 ± 0.5	1.58 ± 0.5	1.06 ± 1.3	1.01 ± 1.3	0.61 ± 0.5	0.61 ± 0.6									
Gp13	0.29 ± 0.5	0.84 ± 0.2	0.51 ± 0.4	0.17 ± 0.3	1.22 ± 0.9	0.41 ± 0.4									
Gp17	9.26 ± 2.0	8.11 ± 1.7	9.95 ± 1.4	14.4 ± 5.0	13.1 ± 2.3	11.6 ± 3.3								0.04	
Gp18	0.29 ± 0.5	0.94 ± 0.9	0.88 ± 0.2	0.35 ± 0.6	0.85 ± 0.8	0.35 ± 0.3									
Gp21	1.05 ± 0.9	1.26 ± 0.5	0.42 ± 0.4	0.00 ± 0.0	1.34 ± 1.2	1.44 ± 1.1									
Gp22	7.23 ± 2.4	8.23 ± 3.0	7.26 ± 2.1	8.38 ± 3.8	7.47 ± 2.1	6.77 ± 1.9									
Gp23	4.25 ± 1.7	4.01 ± 2.0	3.78 ± 1.2	1.70 ± 1.5	4.81 ± 1.0	3.26 ± 0.6									
Gp25	4.69 ± 2.7	3.99 ± 1.0	6.73 ± 0.4	5.66 ± 1.4	5.41 ± 0.7	10.1 ± 5.4									
Gp26	4.43 ± 0.7	4.95 ± 2.2	3.22 ± 0.7	3.13 ± 3.8	2.80 ± 2.1	4.23 ± 1.7									
Others	2.16 ± 2.0	2.26 ± 0.3	2.74 ± 0.2	2.79 ± 0.7	3.33 ± 0.2	1.76 ± 0.9								0.004	

C = treatment without the addition of nitrogen fertilizer; N = treatment with nitrogen addition at the dose of 60 kg ha^{-1}; and NV = treatment with nitrogen fertilization at the dose of 60 kg ha^{-1} and the addition of vinasse at a rate of 0.06 L kg^{-1}. [a]Average for each of three replicates soil. Values are expressed as percentage. [b]Standard deviations of the average for each of three replicates soil. Analysis of variance and Tukey's HSD test two-pair combination at the significance level of $P \leq 0.05$. Blank spaces indicate that the values were not significant.

3.3.3. Gene families associated with carbon degradation

The functional subcategories with the highest number of Acidobacteria-derived probes are "starch" and "hemicellulose" (Table 3). The 18 soil genomic DNA samples from this study hybridized with 16 gene families covered by the Acidobacteria-derived probes available in GeoChip v.5.0M – xylanase, ara, xyla, endoglucanase, cellobiase, RgaE, pme, pectinase, cda, glucoamylase, pula, amya, AceA, AceB, chitinase and acetylglucosaminidase. The highest number of hybridizations occurred with the subcategory "starch", followed by the subcategory "chitin" (Table 4), and the lowest number occurred with the subcategory "pectin".

A significant decrease ($P \leq 0.05$) in the signal intensity of the hybridizations between the soil genomic DNA and Acidobacteria derived probes belonging to the gene families associated with C degradation was observed in the NV treatment when comparing T7 and T150. In the N treatment, most gene families showed a significant decrease ($P \leq 0.05$) in the hybridization signal intensity in the same period (T7 and T150) (Table 4). In the C treatment, significant differences were not observed in the hybridization signal intensity.

Table 3. Total number of probes related to the "carbon degradation" function and the subcategories in the treatments and corresponding number of probes derived only from Acidobacteria

Carbon degradation	T7			T150			Probes	
	C	N60	NV	C	N60	NV	Total	Acidobacteria
Total	5002.0[a] ± 386.3[b]	5472.6 ± 95.3	5379.1 ± 147.5	4738.2 ± 848.5	4896.0 ± 398.3	4936.7 ± 219.5	7592	103
Gene subcategory 2								
Starch	1717.9 ± 27.4	1877.9 ± 13.4	1847.6 ± 65.9	1635.9 ± 39.7	1692.9 ± 34.4	1732.4 ± 24.3	2674	37
Camphor	7.4 ± 0.6	8.0 ± 0.1	8.0 ± 0.1	7.2 ± 1.1	7.5 ± 0.5	7.4 ± 0.6	7	0
Cellulose	367.6 ± 7.4	411.4 ± 2.4	402.3 ± 18.1	340.9 ± 6.7	355.6 ± 4.0	353.0 ± 5.3	562	3
Chitin	741.4 ± 12.9	816.0 ± 5.9	808.4 ± 35.9	697.3 ± 19.0	719.3 ± 15.7	728.9 ± 12.9	1195	18
Cyanide	16.3 ± 0.3	18.1 ± 0.3	18.1 ± 0.8	14.6 ± 0.5	16.0 ± 0.3	15.7 ± 0.7	23	0
Cutin	134.6 ± 5.5	142.5 ± 3.5	143.3 ± 2.0	128.7 ± 16.3	128.6 ± 4.3	129.3 ± 3.7	157	0
Phospholipids	55.5 ± 0.9	61.9 ± 1.6	62.6 ± 4.0	50.7 ± 1.5	54.0 ± 0.4	52.4 ± 1.3	94	0
Glyoxylate cycle	405.0 ± 6.7	445.4 ± 3.7	435.5 ± 13.9	389.2 ± 7.6	405.4 ± 2.4	402.5 ± 4.4	595	3
Hemicellulose	620.5 ± 5.9	670.2 ± 3.9	663.5 ± 23.5	587.3 ± 16.1	608.1 ± 8.9	611.0 ± 7.9	947	23
Inulin	8.2 ± 0.2	9.2 ± 0.2	9.4 ± 0.6	7.7 ± 0.3	8.3 ± 0.0	8.3 ± 0.2	11	0
Lactose	11.6 ± 0.2	12.9 ± 0.6	10.6 ± 0.2	9.9 ± 0.3	10.3 ± 0.2	10.4 ± 0.3	25	0
Lignin	182.8 ± 3.3	202.5 ± 3.6	193.0 ± 5.8	171.4 ± 3.5	176.2 ± 1.6	174.0 ± 2.1	273	0
Others	10.7 ± 0.3	12.5 ± 0.7	11.5 ± 0.9	8.4 ± 0.7	9.6 ± 0.2	9.4 ± 0.5	23	0
Pectin	447.0 ± 6.3	482.0 ± 5.6	470.7 ± 13.0	429.9 ± 11.2	431.9 ± 8.7	440.9 ± 4.2	645	19
Protein	19.9 ± 0.4	22.6 ± 1.1	20.6 ± 0.4	17.9 ± 0.3	19.3 ± 0.7	18.2 ± 0.6	36	0
Tannins	20.4 ± 0.3	22.4 ± 0.4	21.0 ± 0.9	17.5 ± 0.2	19.7 ± 0.3	18.1 ± 0.1	29	0
Terpenes	90.5 ± 1.4	99.9 ± 0.6	98.6 ± 3.7	86.4 ± 1.8	90.2 ± 0.9	87.5 ± 0.6	112	0
Valin/Lignin	144.4 ± 1.6	157.0 ± 1.0	154.3 ± 5.4	137.1 ± 3.3	143.2 ± 1.1	137.2 ± 2.3	184	0

C = treatment without the addition of nitrogen fertilizer; N = treatment with nitrogen addition at the dose of 60 kg ha^{-1}; and NV = treatment with nitrogen fertilization at the dose of 60 kg ha^{-1} and the addition of vinasse at a rate of 0.06 L kg^{-1}. [a]Average for each of three replicates soil. [b]Standard deviations of the average for each of three replicates soil.

Table 4. Signal intensities of the families of carbon degradation genes obtained using GeoChip v. 5.0M from Acidobacteria and hybridized with DNA sampled at seven (T7) and one hundred and fifty (T150) days after vinasse application

Carbon degradation genes	T7			T150			Statistic								
							T7 *vs* T150			T7			T150		
	C	N	NV	C	N	NV	C	N	NV	C *vs* NV	C *vs* NV	N *vs* NV	C *vs* NV	C *vs* NV	N *vs* NV
Glyoxalato cycle															
AceA	198.3[a] ± 10.6[b]	210.9 ± 4.3	205.7 ± 4.6	190.1 ± 21.3	192.9 ± 6.0	191.1 ± 1.5									
AceB	313.3 ± 16.6	331.1 ± 5.7	329.8 ± 2.4	305.9 ± 28.3	311.0 ± 7.5	303.6 ± 4.7									
Chitin															
acetylglucosaminidase	347.3 ± 22.8	373.3 ± 4.1	367.6 ± 3.9	340.9 ± 36.9	343.6 ± 14.7	338.1 ± 5.4									
chitinase	722.1 ± 45.2	764.9 ± 5.8	758.6 ± 7.4	700.0 ± 69.0	712.0 ± 17.8	696.8 ± 4.1									
Starch															
amyA	1953.0 ± 116.5	2084.1 ± 15.2	2049.0 ± 24.8	1885.1 ± 200.7	1901. ± 70.3	1895.7 ± 13.6									
cda	148.6 ± 11.1	158.0 ± 2.5	156.3 ± 3.1	146.5 ± 16.1	147.0 ± 3.9	140.5 ± 3.2									
glucoamylase	107.8 ± 5.8	113.1 ± 0.4	111. 7 ± 0.4	103.67 ± 9.7	103.0 ± 7.1	104.4 ± 1.9									
pula	120.4 ± 7.8	127.3 ± 2.6	126.6 ± 2.4	116.2 ± 13.4	115.4 ± 7.5	115.6 ± 2.0									
Hemicellulose															
ara	273.5 ± 14.7	289.1 ± 2.2	285.7 ± 1.9	270.8 ± 19.9	272.2 ± 8.0	267.2 ± 3.6									
xyla	176.0 ± 10.1	185.5 ± 1.3	183.8 ± 4.7	173.5 ± 16.9	171.2 ± 6.1	170.3 ± 2.5									
xylanase	273.4 ± 18.3	291.2 ± 2.8	286.1 ± 2.2	266.1 ± 25.0	270.4 ± 12.8	263.0 ± 2.6									
Cellulose															
cellobiase	229.6 ± 14.3	245.2 ± 2.1	239.0 3 ± 1.3	221.3 ± 24.9	223.1 ± 8.3	223.1 ± 3.0									
endoglucanase	167.0 ± 11.1	177.5 ± 1.6	175.0 ± 1.7	164.0 ± 14.7	165.5 ± 6.0	161.6 ± 2.3									
Pectin															
pectinase pectate_liase	96.27 ± 6.4	101.9 ± 1.3	101. 3 ± 2.4	92.91 ± 8.7	94.17 ± 4.3	94.4 ± 2.1									
pmE	95.20 ± 5.0	100.2 ± 2.9	100. 3 ± 1.4	94.26 ± 7.9	95.09 ± 5.3	94.9 ± 1.8									
RgaE	118.2 ± 6.5	124.1 ± 1.5	122.3 ± 3.0	114.7 ± 10.2	116.5 ± 1.6	116.3 ± 1.0									

C = treatment without the addition of nitrogen fertilizer; N = treatment with nitrogen addition at the dose of 60 kg ha^{-1}; and NV = treatment with nitrogen fertilization at the dose of 60 kg ha^{-1} and the addition of vinasse at a rate of 0.06 L kg^{-1}. [a]Average for each of three replicates soil. [b]Standard deviations of the average for each of three replicates soil. Analysis of variance and Tukey's HSD test two-pair combination at a significance level of $P \leq 0.05$. Blank spaces indicate that the values were not significant.

3.3.4. Relationship among the relative abundance of Acidobacteria subgroups, gene families associated with C degradation and soil chemical characteristics

The RDA revealed two distinct groups according to the sampling times (Figure 1). The abundance of the Gp3, Gp6, Gp17, Gp21 and Gp25 subgroups was positively related to the chemical characteristics of the soil samples collected at T150. The abundance of most of the subgroups identified in this study and all the gene families analyzed were positively related to the chemical characteristics of soil samples collected at T7.

The combination of the three datasets taxonomic data, functional data and soil chemical factors for the construction of the neural network based on all correlations significant or not in *Spearman's* nonparametric test resulted in a network with a total of 42 nodes and 85 edges and modularity of -0.174 (Figure 2). The constructed network improved the visualization and interpretation of the theoretical relationships among the Acidobacteria subgroups, gene families and soil chemical factors beyond those revealed by the RDA.

Based on the network, the greatest number of observed interactions occurred between the Acidobacteria subgroups and the soil chemical factors, most of which were negative. The chemical factors with the greatest number of interactions with the gene families and the subgroups were CEC, B and Zn. Only the Gp3 and Gp6 subgroups presented positive interactions with Zn and CEC. *Spearman's* correlation analysis (Table S1) showed a significantly positive correlation between Gp6 and the C:N ratio. The Gp2 subgroup presented the greatest number of interactions in the network and was the only subgroup to interact positively with H + Al. *Spearman's* correlation analysis showed that this subgroup was negatively correlated with S and BS (Table S1). The only subgroup that showed significant negative correlations with all C degradation gene families was Gp17 (Table S1). However, this subgroup was significantly positively correlated with P, BS, V, B and Zn (Figure 1 and Table S1). The only class that showed positive correlations with the gene families involved with C degradation in the network was Gp13, which was correlated with "xylanase" and "pme" belonging to the subcategories "hemicellulose" and "pectin", respectively. However, *Spearman's* correlation analysis indicated that this subgroup was significantly positively correlated with 11 of the 16 gene families evaluated (Table S2). With a similar profile based on *Spearman's* correlation analysis, Gp18 also showed a positive correlation with most gene families (Table S2); however, it showed a negative correlation with Ca and CEC (Table S1). Subgroup Gp1 showed a significant positive correlation with Cu and Mn (Figure 1), and *Spearman's* correlation analysis indicated that this subgroup was correlated with

Fe (Table S1). The Acidobacteria subgroups Gp4 and Gp6 were positively correlated with the soil pH (Table S1 and Figure 1). Nine Acidobacteria subgroups revealed a significant correlation among them (Table S3).

Figure 1. Constrained ordination diagram for sample plots in the first two-redundancy analysis (RDA) axes based on the soil physicochemical factors of the different soil treatments and their relationship with the Acidobacteria subgroups and carbon degradation gene families

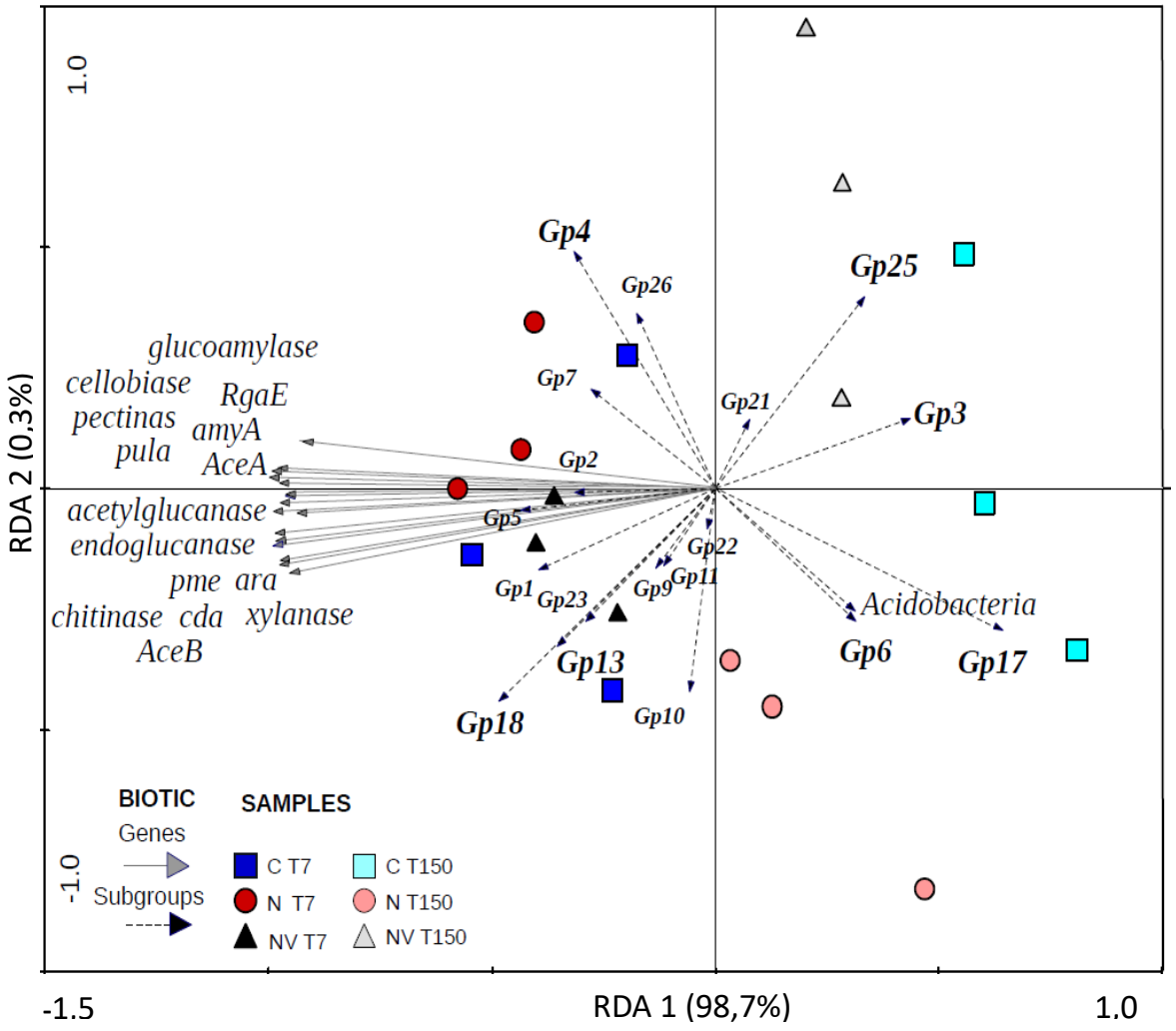
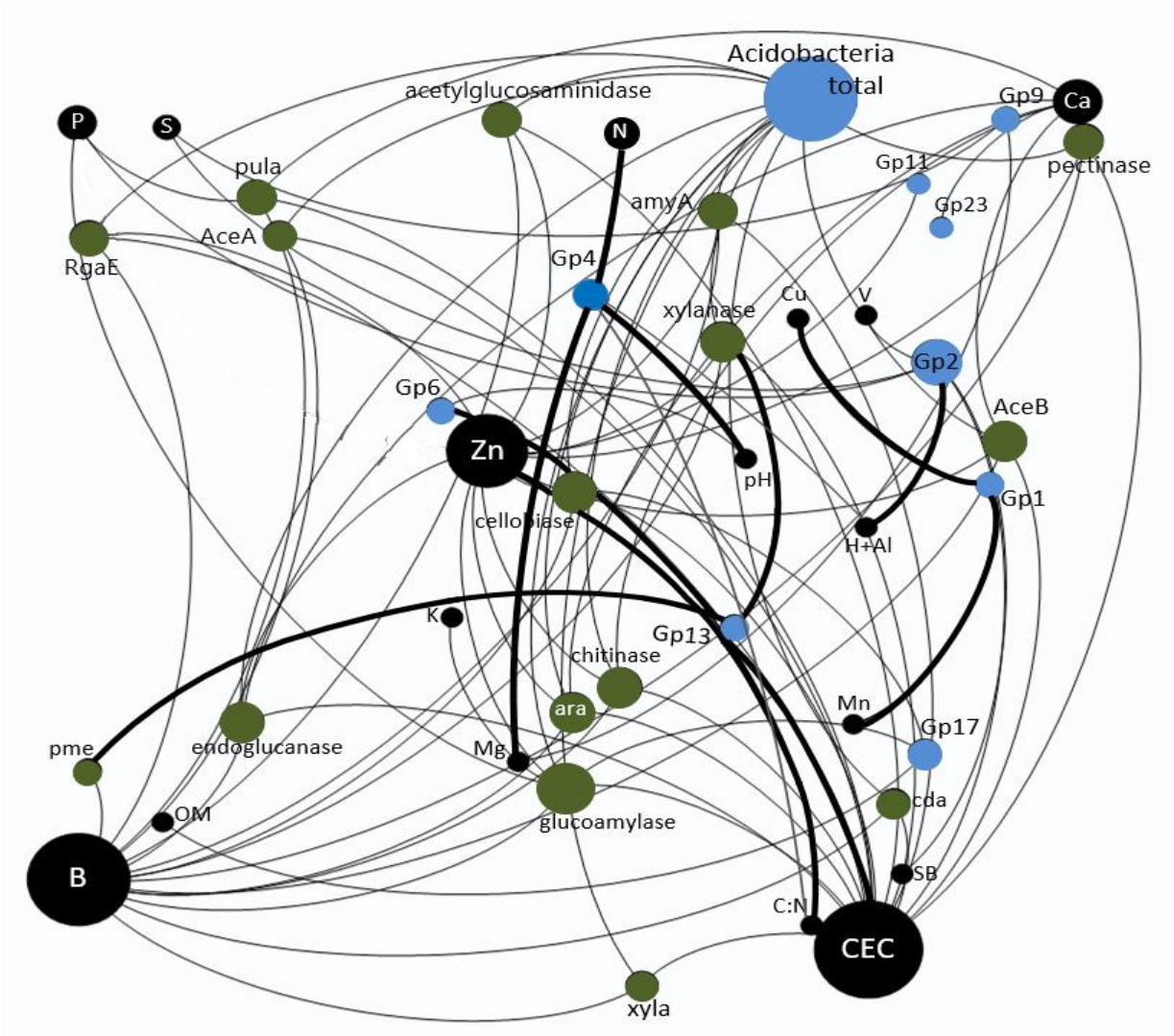


Figure 2. Correlation among the soil chemical factors, different Acidobacteria subgroups, and quantified carbon degradation gene families in the treatments with nitrogen fertilization and vinasse application. The size of the circles corresponds to the number of interactions. The thicker lines correspond to positive interactions, and the thinner lines correspond to negative interactions.



3.4. Discussion

One of the most ubiquitous and abundant bacterial phyla in the soil environment still has large knowledge gaps regarding its role in soil. Analyses based on genomic DNA from environmental samples have provided relevant information on the metabolic potential of Acidobacteria (Lee et al., 2008; Liles et al., 2013; Faoro et al., 2012; Lin et al., 2019) and led to alternative methods of culture-dependent techniques, which can provide a better understanding of the ecology of this recalcitrant bacterial phylum in a culture medium. Such analyses based on culture-independent techniques have provided insights into the role of Acidobacteria in the soil environment. This study combines genomic, chemical, statistical, and computational analyses to obtain information about the dynamics of Acidobacteria subgroups and their potential for C degradation in soils under sugarcane in a tropical region. Previous studies on tropical Amazonian soils with Acidobacteria are consistent with the results presented here, in which different lifestyles were observed for the different subgroups (Navarrete et al., 2010, 2013, 2015) except for subgroups Gp6 and Gp17. However, the functional aspects of the Acidobacteria subgroups have not been studied yet in tropical soils.

Navarrete et al. (2010) showed that a high abundance of the subgroup Gp6 occurs in Amazonian Dark Earth (ADE), which presents high levels of OM and nutrients and a pH of about 5.0. The Gp6 subgroup was also highly abundant in 27 pasture soils in Germany, where the pH, N, P and temperature values were high and the C:N ratio was low compared with that of forest soils in the region (Naether et al., 2012). Although the Gp6 subgroup showed a positive correlation with the Zn contents, CEC and C:N ratio, a decrease in the abundance of this subgroup was observed in the NV treatment, which presented chemically similar conditions as ADE and European pasture soils. These observations indicate that the Gp6 subgroup has a preference for soils with pH below 5.5 (Lin et al., 2019), which justifies its negative correlation with the pH in this study as well as its high abundance in the C treatment pH 5.2 to 5.5 and decreased abundance in the NV treatment pH 5.8. Authors have suggested that this subgroup may be positively or negatively correlated with soil pH (Chan et al., 2006; Mukherjee et al., 2014), which shows that the behavior of the same Acidobacteria subgroup can vary in different types of soil (Naether et al., 2012). Additionally, studies have explored the interactions of the Gp6 subgroup with bacteria from other phyla and even protozoa (Naether et al., 2012; Spring et al., 2000). However, these studies have not provided conclusive results for possible ecological interactions.

The abundance of subgroup Gp17 has been positively correlated with pH in pasture soils (Naether et al., 2012; Navarrete et al., 2015b). Although the abundance of subgroup Gp17 was not correlated with the soil pH in this study, it was correlated with several chemical factors in soil treatment that received N fertilizer combined with vinasse. In ADE with high N levels and pH values approximately 5.0, this group was not very abundant (Navarrete et al., 2010), suggesting that the occurrence of Gp17 can be modulated by other chemical factors, such as P, BS, V, B, and Zn. Although the Gp17 subgroup has the potential to become one of the most abundant Acidobacteria subgroups in soils under sugarcane supplemented with vinasse, it probably does not have genes related to the degradation of hemicellulose, cellulose, pectin, starch and chitin and/or participation in the glyoxalate cycle according to the results of this study. This result suggests that C degradation metabolic pathways different from those analyzed here or ecological interactions with other microorganisms must occur to obtain the required C sources for their development.

The predominance of the Gp4 subgroup in soil under sugarcane supplemented with vinasse is corroborated by a previous study by Omori et al. (2016). The results of this study also showed that a longer period after vinasse application contributes to the increased abundance of the Gp4 subgroup. Navarrete et al. (2010) also showed a predominance of the Gp4 subgroup when ADE is supplemented with biochar or biocarbon. Considering that biocarbon constitutes a hotspot and presents high Ca, Mg and Zn amounts, pH values that reach 10.0, microaerophilic conditions and higher temperatures than the adjacent soil, the increased abundance of this subgroup in the vinasse treatment suggests that soil microhabitats in fields treated with this residue provide adequate physical and chemical conditions for the development of these bacteria (Kuzyakov and Blagodatskaya, 2015). Studies have indicated that the application of vinasse for a short period in the soil can reduce the water infiltration rate and total porosity due to the cementing effect of OM (Dalri et al., 2010; dos Santos et al., 2017), suggesting that the addition of vinasse may alter the soil aggregate structure. *Chloracidobacterium thermophilum*, a representative of subgroup Gp4 isolated from hot springs, is microaerophilic and presents optimal growth at a temperature of 5ºC and pH ranging from 5.0 to 7.0 (Tank and Bryant, 2015). *Pyrinomonas methylaliphatogenes* K22T is another isolate from New Zealand hot springs that have been grown in a low nutrient culture medium, and it is moderately acidophilic with an optimal growth pH of 6.5 (Lee et al., 2015). Wust et al. (2016) detected exopolysaccharides (EPS) production in Gp4 isolates from African savanna soil, which is interesting because EPS can be used by microorganisms to remain attached to soil particles. Although the GP4 subgroup has been described as one of the most versatile in the use

of carbohydrates because it can use C from cellobiose, sucrose, maltose, chitins, carboxymethyl cellulose and microcrystalline cellulose (Foesel et al., 2013; Huber et al., 2014; Lee et al., 2015), the analyses performed in this study did not reveal interactions of this subgroup with the families involved in C degradation. A large number of hybridizations between the Acidobacteria-derived probes and the gene families associated with "chitin" degradation can be explained by the fact that the Gp4 subgroup, which is abundant in treatments with NV and containing hydrolase (Huber et al., 2014), has the potential to use the C found in the cell wall of fungi used in the fermentative process that results in the vinasse. Nakamura (2014) performed a prospective study of isolates that degrade aromatic hydrocarbons in ADE and enriched culture media with C from lignocellulose but did not obtain any Acidobacteria isolates. Kielak et al. (2016) warned against premature conclusions about the ability of Acidobacteria to degrade these polysaccharides since there are discrepancies between the information contained in the genomes of cultured representatives and the result of physiological tests with such representatives.

The subgroups Gp13 and Gp18 presented the highest number of positive correlations with the gene families related to C degradation, especially those involved in hemicellulose degradation. However, both subgroups presented low abundance in the treatment containing vinasse, which indicates that the physical and chemical conditions imposed by the addition of this organic residue to the soil has a negative effect on the occurrence of these subgroups. The negative correlations between the Gp18 subgroup and the Ca and CEC levels indicate that representatives of these subgroups may benefit from a lower availability of nutrients, which is inconsistent with the conditions generated via the addition of vinasse to soil. Navarrete et al. (2015b) observed that the abundance of subgroup Gp18 increased in pasture soil, which presents higher pH and nutrient availability relative to forest soils. Although the soil subjected to vinasse application in this study had chemical characteristics very similar to that of pasture soils (Navarrete et al., 2015b; Naether et al., 2012), subgroup Gp18 was the least abundant in the NV treatment. Physical factors that can be altered in the presence of vinasse and the pasture, such as soil moisture and aeration (dos Santos et al., 2017; Silva Filho et al., 2010; Dedecek et al., 2000), may have influenced the occurrence of the Gp18 subgroup. In the conversion of forest to pasture, Gp13 decreases significantly (Navarrete et al., 2015b), suggesting that the ecological role of these microorganisms in natural environments is important. Unfortunately, anthropogenic impacts on the soil environment, such as the addition of agroindustrial residues and pasture establishment, may promote the reduced abundance of subgroups with relevant functions in C degradation.

This study also provided additional evidence on the metabolic potential of Gp2 regarding Al. Similar to observations in soil from the Amazon forest, Atlantic forest and Cerrado Brazilian Savanna (Navarrete et al., 2013b; Cato et al., 2014; de Carvalho et al., 2016), this subgroup was positively associated with H + Al.

Recent studies on the phylum Acidobacteria have suggested that its subgroups differ with respect to their lifestyle, with some showing oligotrophic behavior and others showing copiotrophic behavior (Kielak et al., 2016; Yao et al., 2017), and the behavior associated with the lifestyle may even vary for the same subgroup in different soil types as mentioned above. In this study, the gene families associated with C degradation and most of the Acidobacteria subgroups were positively correlated with the soil chemical characteristics at T7, in which the lowest N, C, Ca, Mg, K levels were obtained (Figure 1). Olander and Vitousek (2000) evaluated the chitinase enzyme activity in Hawaiian soils with different ages and N levels, and the authors concluded that enzymatic activity was inhibited in older soils with higher N levels while chitinase production was activated in younger soils with lower total N levels. Chitinase is one of the most abundant enzymes in tropical soils because these soils contain a large amount of fungi with chitinous cell wall and invertebrates (Olander and Vitousek, 2000). In soils supplemented with vinasse, a residue resulting from the fermentation process, the amount of yeast is high. In this study, the high availability of N in the treatments observed in T150 may have altered the enzymatic metabolism associated with C in Acidobacteria, with carbon dioxide and nitrous oxide emissions increased in the first days after application of the fertilizers as revealed by Navarrete et al. (2015a) when assessing the same experiment and sugarcane soil. Access to OM occurred before T150, when the stocks are significantly reduced (Table 1). If the addition of N as fertilizer can reduce the decomposition of recalcitrant C (Craine et al., 2007), then the members of the bacterial community that act as decomposers of this C fraction would be promoted, which suggests that Acidobacteria members that increase in abundance in soils that receive N fertilization in combination with vinasse such as the Gp4 subgroup may be able to use different sources of C and adapt to variations in the physicochemical conditions of the environment. In this sense, Gp4 would be able to change its lifestyle from oligotrophic to copiotrophic depending on the selective pressure of the environment. Morphologically, the Gp4 isolates from African savanna soil showed cytoplasmic extensions, indicating adaptations for nutrient uptake in nutrient-poor environments (Wust et al., 2016). In addition, these isolates contain genes encoding thermophilic hydrolase enzymes (Huber et al., 2014).

In soils fertilized with vinasse, the pH seems to have little influence on the general distribution of Acidobacteria subgroups. Recent studies involving Acidobacteria have

questioned the strong correlation between these bacteria and pH because whether this correlation is based on a direct causal relationship or the covariation among different soil chemical factors remains unclear (Kielak et al., 2016). Suleiman et al. (2018) also suggest that the application of organic residues to soil results in an increase in the abundance of microbial groups related to the N cycle. Physical factors altered by the application of vinasse may influence the occurrence and frequency of Acidobacteria subgroups, and such factors include temperature and/or organisms that possibly interact with Acidobacteria. Future studies should include analyses of soil physical factors in addition to chemical factors as well as of aspects related to the interactions between Acidobacteria subgroups and the biotic factors of their environment. Moreover, the lack of isolates of most Acidobacteria subgroups increases the difficulty of conducting physiological tests to validate genetic predictions, which is a limiting factor for advancing our understanding of the role of Acidobacteria in ecosystems. Genomic and physiological studies with isolates of the Gp4 subgroup can contribute to a better understanding of the role of Acidobacteria in C degradation in the soil and could provide new possibilities for the application of the biotechnological potential of these bacteria favored in sugarcane soil enriched with N and vinasse.

3.5. Conclusion

The addition of N and vinasse to soil under sugarcane can increase the availability of nutrients in this environment, especially Ca, Mg, K, Al, B and Zn, whose increased levels in the soil were related to decrease in the abundance of Acidobacteria subgroups and gene families associated with C degradation. The Gp13 and Gp18 subgroups, which are positively associated with C degradation, did not show adaptive success to the physical-chemical conditions imposed by the addition of N and vinasse in the soil, suggesting that changes resulting from this agricultural soil management practice can affect C metabolism in Acidobacteria.

References

Altschul, S.F., Gish, W., Miller, W. et al. (1990). Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Barros, R.P., Viegas, P.R.A., da Silva, T.L. et al. (2010). Alterações em atributos químicos de solo cultivado com cana-de-açúcar e adição de vinhaça. Pesq. Agropec. Trop. 40, 341-346.

Barns, S.M., Cain, E.C., Sommerville, L., Koske, C.R. (2007). Acidobacteria phylum sequence sin uranium - contaminated sub surface sediments greatly expand the known diversity within the phylum. Appl. Environ. Microbiol. 73, 3113–3116.

Bastian, M., Heymann, S., Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. In: International AAAI conference on weblogs and social media. San Jose, California.

Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre E. (2008). Fast unfolding of communities in large networks. J. Stat. Mech. P10008. doi: 10.1088/1742-5468/2008/10/P10008

Brandes, U. (2001). A faster algorithm for betweenness centrality. J. Math. Sociol. 25, 163-177.

Catão, E.C.P., Lopes, F.A.C., Araújo, J.F. et al. (2014). Soil acidobacterial 16S rRNA gene sequences reveal subgroup level differences between savanna-like cerrado and atlantic forest brazilian biomes. Int. J. Microbiol. 2014, 156341. doi: 10.1155/2014/156341

Christofoletti, C. A., Escher, J. P., Correia, J. E. et al. (2013). Sugarcane vinasse: environmental implications of its use. Waste Manag. 33, 2752-61.

Chan, O.C., Yang, X.D., Fu, Y. et al. (2006). 16S rRNA gene analyses of bacterial community structures in the soils of evergreen broad-leaved forests in south-west China. FEMS Microbiol. Ecol. 58, 247–259.

Craine, J.M., Carl, M., Noah, F. (2007). Microbial nitrogen limitation increases decomposition. Ecology. 88, 2105-2113.

Cole, J.R., Wang, Q., Fish. J.A., et al. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 42, D633–D642.

Cong, J., Liu, X., Lu, H., et al. (2015). Analyses of the influencing factors of soil microbial functional gene diversity in tropical rainforest based on GeoChip 5.0. Gen. Data. 5, 397-398.

De Carvalho, T.S., Jesus, E.C, Barlow, J. et al. (2016). Land use intensification in the humid tropics increased both alpha and beta diversity of soil bacteria. Ecology. 1-1.

Dedecek, R.A., Mengol, O., Bellote, A.F. (2000). Avaliação da compactação do solo em plantios jovens de *Pinus taeda*, com diferentes sistemas de preparo do solo. Bol. Pesq. Fl. 40, 5-21.

Dos Santos, C.H., Claudinei, A.C., Creste, J.E. et al. (2017). Propriedades físicas de um argissolo após fertirrigação com vinhaça e sistemas de colheita da cana de açúcar. Col. Agr. 13, 58-66.

Empresa Brasileira de Agropecuária - Embrapa, 2011. Manual de Métodos de Análise do Solo. Rio de Janeiro, Brasil.

Ewing, B., Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8(3), 186–194.

Faoro, H., Glogauer, A., Couto, G.H., et al. (2012). Characterization of a new Acidobacteria-derived moderately thermostable lipase from a Brazilian Atlantic Forest soil metagenome. FEMS Microbiol Ecol. 81, 386–394.

Foesel, B.U., Rohde, M., Overmann, J. (2013). *Blastocatella fastidiosa* gen. nov., isolated from semiarid savana soil–the first described species of Acidobacteria subdivision 4. Syst. Appl. Microbiol. 36, 82–89.

Gascho, G.J., Ascho, G.J., Shih, S.F. (1983). Sugarcane. In Teare ID, Peet MM (ed) Crop-water relations. Wiley-Interscience, New York, pp 445-479.

Gasparotto, F., Rodrigues, F.S., Seratoo, C.D. et al. (2014). Cadeias produtivas da cana-de-açúcar, do algodão e de frutas. Maringá, Paraná.

Glória, N.A., Filho, O.J. (1983). Boletim Técnico Planalsucar, Piracicaba, São Paulo.

Heffer, P. and Prud'homme M. (2008). Medium-term outlook for global fertilizer demand, supply and trade 2008-2012. Presented at the 76th IFA Annual Conference.

Hugenholtz, P., Goebel, B.M., Pace, N.R. (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J. Bacterial. 180, 4765–4774.

Huber, K.J., Wust, P.K., Rohde, M. et al. (2014). *Aridibacterfamidurans* gen. nov., spnovand *Aridibacter kavangonensis* sp nov., two novel members of subdivision 4 of the Acidobacteria isolated from semiarid savannah soil. Int. J. Syst. Evol. Microbiol. 64, 1866–1875.

Jiang, Z.P., Li, Y.R., Wei, G.P. et al. (2012). Effect of long-term vinasse application on physico-chemical properties of sugarcane field soils. Sugar Tech. 14, 412–417.

Kielak, A.M., Barreto, C.C., Kowalchuk, G.A. et al. (2016). The Ecology of Acidobacteria: moving beyond genes and genomes. Front. Microbiol. 7-744.

Kuzyakov Y., Blagodatskaya E. (2015). Microbial hotspots and hot moments in soil: Concept & review. Soil Biol. Biochem. 83, 184-199.

Lee, S.H., Ka, J.O., Cho, J.C. (2008). Members of the phylum Acidobacteria are dominant and metabolically active in rhizosphere soil. FEMS Microbiol. Lett. 285, 263–269.

Lee, K.C., Morgan, X.C., Power, J.F. et al. (2015). Complete genome sequence of the thermophilic Acidobacteria, *Pyrinomonas methylaliphatogenes* type strain K22T. Stand. Genomic Sci. 10,101.

Lian, T., Yu, Z; Li, Y. et al. (2019). The shift of bacterial community composition magnifies over time in response to different sources of soybean residues. Appl. Soil Ecol. 136, 163-167.

Liang, Y., He, Z., Liyou, W. et al. (2010). Development of common oligonucleotide reference standard for microarray data normalization and comparison across different microbial communities. Appl. Env. Microbiol. 76, 1088-1094.

Liles, M.R., Manske, B.F., Bactrim, S.B. et al. (2003). A census of rRNA genes and linked genomic sequences within a soil metagenomic library. Appl. Environ. Microbiol. 69, 2684–2691.

Lin Yu-Te, Lin Yu-Fei 1, Tsai Isheng J. et al. (2019). Structure and diversity of soil bacterial communities in Ofshore Islands. Nat. Sci. Rep. 9:4689

Magoc, T., Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinf. 27, 2957-2963.

Montenegro, S.J.C., Menjivar, C.R., Bonilla, R. et al. (2009). Influencia de la aplicación de vinaza en actividad y biomasa microbiana en un entic dystropept y un fluventic haplustoll del Valle del Cauca, Colombia. Acta. Agron. 58, 41-45.

Moran-Salazar, R. G., Sanchez-Lizarraga, A. L., Rodriguez-Campos, J. et al. (2016). Utilization of vinasses as soil amendment: consequences and perspectives. Springer Plus. 5, 1007.

Meyer, F., Pearmann, D., D'Souza, M. et al. (2008). The metagenomics RAST server — a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinf. 9, 386.

Mutton, M.A., Rossetto, R., Mutton, M.J.R. (2010). Agricultural use of stillage. In: Sugar-cane Bioethanol. R&D for Produc. and Sustain. 1, 423–440.

Mukherjee, S., Juottonen, H., Siivonen, P. et al. (2014). Spatial patterns of microbial diversity and activity in an aged creosote-contaminated site. ISME J. 8, 2131–2142.

Navarrete, A.A., Cannavan, F.S., Taketani, R.G., Tsai, S.M. (2010). A molecular survey of the diversity of microbial communities in different Amazonian agricultural model systems. Div. 2, 787-809.

Navarrete, A.A., Barreto, C.C., Arnaldo, M., Tsai, S.M. (2013a). Molecular detection on culture medium of Acidobacteria from Amazon soils. Microbiol. Discov. 1, 787–809.

Navarrete, A.A., Kuramae, E.E., de Hollander, M. et al. (2013b). Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. FEMS Microbiol. Eco. 83, 607–621.

Navarrete, A.A., Diniz, T.R., Braga, L.P.P. et al. (2015a). Multi-analytical approach reveals potential microbial indicators in soil for sugarcane model systems. PLoS ONE. 6, e0129765.

Navarrete, A.A., Venturini, A.M., Kyle, M. et al. (2015b). Differential response of Acidobacteria subgroups to forest-to-pasture conversion and their biogeographic patterns in the Western Brazilian Amazon. Front. Microbiol. 6, 1443.

Nakamura, F.M., Germano, M.G., Tsai, S.M. (2014). Capacity of aromatic compound degradation by bacteria from Amazon dark earth. Div. 6, 339-353.

Naether, A., Foesel, B.U., Naegele, V. et al. (2012). Environmental factor saffect acidobacterial communities below the subgroup level in grassland and forest soils. Appl. Environ. Microbiol. 78, 7398–7406.

Olander, L.P., Peter, M.V. (2000). Regulation of soil phosphatase and chitinase activity by N and P availability. Biogeoch. 49, 175–190.

Omori, W.P., Camargo, A.F., Goulart, K.C.S. et al. (2016). Influence of vinasse application in the structure and composition of the bacterial community of the soil under sugarcane cultivation. Int. J. Microbiol. 2016, 11.

Rachid, C.T.C.C., Santos, A.L., Piccolo, M.C. et al. (2013). Effect of sugarcane burning or green harvest methods on the Brazilian cerrado soil bacterial community structure. Plos One. 8, e59342.

Rawat, S.R., Mannisto, M.K., Bromberg, Y. et al. (2012). Comparative genomic and physiological analysis provides insights into the role of Acidobacteria in organic carbon utilization in Arctic tundra soils. FEMS Microbiol. Ecol. 82, 341–355.

Ribeiro, B.T., Lima, J.M., Curi, N., Oliveira, G.C. (2012). Electrochemical attributes of soils influenced by sugarcane vinasse. Biosc. J. 28,1.

Resende, A.S. (2006). Long-term effects of pre-harvest burning and nitrogen and vinasse applications on yield of sugar cane and soil carbon and nitrogen stocks on a plantation in Pernambuco, N.E. Brazil. Plant. Soil. 281, 339–351.

Silva Filho, E.P., Cottas, L.P., Marini, G.B.S. (2010). Avaliação da compactação dos solos em áreas de pastagens e florestas em Porto Velho-Rondônia. Bol. Geo. Mar. 28, 145-155.

Silva, G.G.Z., Cuevas, D.A., Dutilh, B.E., Edwards, R.A. (2014). FOCUS: an alignment-free model to identify organisms in metagenomes using non-negative least squares. Peer J. 2, e425.

Spring, S., Schulze, R., Overmann, J. et al. (2000). Identification and characterization of ecologically significant prokaryotes in the sediment of freshwater lakes: molecular and cultivation studies. FEMS Microbiol. 24, 573–590.

Suleiman, A.K.A., Silva Lourenço, K., Pitombo, L.M. et al. (2018). Recycling organic residues in agriculture impacts soil-borne microbial community structure, function and N_2O emissions. Sci. T. Env. 631–632, 1089–1099.

Tank, M., Bryant, D. (2015). *Chloracidobacterium thermophilum* gen. nov., sp. nov.: an anoxygenic microaerophilic chlorophotoheterotrophic Acidobacterium. Int. J. Syst. Evol. Microbiol. 65, 1426–1430.

Ter Braak, C.J.F., Smilauer, P. (1998). CANOCO Release 4. Reference Manual and Users Guide to CANOCO for Windows: Software for Canonical Community Ordination. Microcomputer Power, Ithaca, USA.

Val-Moraes, S.P., Macedo, H.S., Kishi, L.T. et al. (2016). Liming in the sugarcane burnt system and the green harvest practice affect soil bacterial community in northeastern São Paulo, Brazil. Ant. van Leeuw.109, 1643-1654.

Ward, N.L., Challacombe, J.F., Janssen, P.H. et al. (2009). Three genomes from the phylum Acidobacteria provide insight into the lifestyles of the microorganisms in soils. Appl. Env. Microbiol. 75, 2046–2056.

Wang C., Liu D., Bai E. (2018). Decreasing soil microbial diversity is associated with decreasing microbial biomass under nitrogen addition. Soil Biol. Biochem. 120, 126–133.

Wu L, Liu X, Schadt CW, Zhou J. Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. Appl Environ Microbiol. 2006, 4931–4941

Wust, P.K., Foesel, B.U., Geppert, A. (2016). *Brevitalea aridisoli*, *B. deliciosa* and *Arenimicrobium luteum*, three novel species of Acidobacteria subdivision 4 class Blastocatellia isolated from savanna soil and description of the novel family Pyrinomonadaceae. Int. J. Syst. Evol. Microbiol. 66, 3355-3366.

Yao F., Yang S., Wang Z. et al. 2017. Microbial taxa distribution is associated with ecological trophic cascades along an elevation gradient. Front. Microbiol. 8:2071Zhou, J-Z, Kang, S., Schadt, C. (2008). Spatial scaling of functional gene diversity across various microbial taxa. Proc. Nat. Acad. Sci., 105: 7768-7773

Zimmermann, J., Gonzalez, J.M., Saiz-Jimenez, C. et al. (2005). Detection and phylogenetic relationships of highly diverse uncultured acidobacterial communities in Altamira cave using 23S rRNA sequence analyses. Geomicrobio. J. 22, 379–388.

4. Chapter 3: Ecological co-occurrence and soil physicochemical factors drive the archaeal community in Amazonian soils

Abstract

We evaluated the co-occurrence of archaeal taxonomic groups and soil physicochemical characteristics about the structuring of the archaeal community in Amazonian soil under different land use systems. Soil samples were collected in primary forest (PF), secondary forest (SF), agricultural systems (AG) and cattle pastures (PA). Archaeal community composition was revealed based on high-throughput amplicon sequencing of the 16S rRNA gene. Co-occurrence (group to group and group to physicochemical factor) and statistical analysis were conducted on both data sets. The results revealed strong co-occurrence of archaeal classes, with two groups formed: a) Thaumarchaeota classes, including South Africa Gold Mine – Group 1 (SAGMG-1), the soil Crenarchaeotic group (SCG) and the Crenarchaeota candidate division YNPFFA (positive Spearman correlation, $p < 0.05$), with predominance in PF and SF; and b) Bathyarchaeota_unclassified, Methanomicrobia and Methanobacteria (Euryarchaeota) with the FHMa11 terrestrial group (positive Spearman correlation, $p < 0.05$), with predominance in PA. The number of co-occurrences between groups was lower in SF, AG and PA (~ 30%) than in PF. qPCR analysis revealed that PF also had the largest number of archaeal representatives. Soil texture may be a limiting factor of interactions between groups since the most representative groups, SAGMG-1 and the SCG (over 20% in all sites), were positively associated with coarse sand, the soil factor most correlated with the groups (33% of the total). These results suggest that interactions between archaeal classes belonging to different phyla supply the energy necessary to survive and may be dependent on the number of individuals in the soil environment. In this context, changes in soil physical structure can reduce the representatives of key groups and consequently the co-occurrences of Archaea, which could compromise the natural dynamics of this complex environment.

Keywords: Archaea, tropical rainforest, 16S rRNA gene amplicon sequencing, soil agriculture, syntrophy

4.1. Introduction

In recent decades, Archaea have been the focus of many studies on the microbial ecology of tropical soils due to their ubiquity, abundance and importance in biogeochemical cycles (NAVARRETE et al., 2011; OFFRE; SPANG; SCHLEPER, 2013; ALVES et al., 2018). Because members of the domain Archaea present close relationships with themselves and with soil biotic and abiotic factors, they can be characterized as key biotic factors in tropical ecosystems (MOISSL-EICHINGER et al., 2018). Although large-scale sequencing technology

provides some insight into the diversity and ubiquity of Archaea in natural systems (MOISSL-EICHINGER; HUBER, 2011; DELONG; PACE, 2001), there are still many difficulties in obtaining isolates of Archaea from the soil, which limits our understanding of these interactions in this environment (MOISSL-EICHINGER et al., 2018). These microorganisms can interact with themselves through syntrophy (TAJIMA et al., 2001). Symbiosis or syntrophy (i.e., the physical association of two or more types of microorganisms) (MOISSL-EICHINGER et al., 2018) allows a group of organisms to meet the energy requirements for growth. This feature may be the key to archaeal survival and success in extreme and complex environments, such as soil.

In general, most microorganisms live in communities, and this lifestyle is particularly useful for the maintenance of energy resources and protection against environmental stress (MOISSL-EICHINGER; HUBER, 2011). A well-described example occurs between the archaea *Nanoarchaeum equitans* and *Ignicoccus hospitalis* (HUBER et al., 2002). *I. hospitalis* can be grown in pure culture, while *N. equitans* cannot. *I. hospitalis* has one of the smallest genomes among those of free-living microorganisms (PODAR; ANDERSON; MAKAROVA, 2008), from which all the genes for heterotrophy are missing; therefore, the species is autotrophic. Energy transfer is believed to occur between *I. hospitalis* and its partner *N. equitans* (MOISSL-EICHINGER; HUBER, 2011). This "intimate relationship" (JAHN et al., 2008), in which there is no harm to either party but benefits to *N. equitans*, can be termed syntrophy. To study co-occurrence between groups of microorganisms, the construction of integrated networks based on taxonomic and group abundance data generated by 16S gene sequencing, which serve as computational resources, has been used (DE CHAVES et al., 2019; KHAN et al., 2019). These interactions between Archaea and their partners function in biogeochemical cycles, driving, for example, carbon (C), nitrogen (N), sulfur (S) and ammonia (NH_3) oxidation (MOISSL-EICHINGER; HUBER, 2011). Among the relationships between Archaea, which once constituted the largest known group of extremophilic microorganisms, and abiotic factors, those between Archaea and methane (CH_4), S, or iron (Fe) are the most important in extreme environments. In these environments, the presence of enzymes and specific cellular biochemical resources is required for metabolic processes to occur (RAMPELOTTO, 2013). Considering the subsistence potential of Archaea, how could the availability of resources limit the functionality of these microorganisms in the soil environment? Delgado-Baquerizo et al. (2016) conducted a study with microcosms and evaluated the enzymatic activity involved in fundamental processes, such as climate regulation, food production and waste decomposition. The authors concluded that the

composition of the soil microbial community is relevant about to specific functions, such as N_2O gas production. In contrast, for general functions, such as respiration, soil properties are more important.

In the soil environment, anthropogenic actions, such as the conversion of forest to land for agriculture, have promoted changes in the natural dynamics of microorganisms (KURAMAE et al., 2011). Changes in plant cover and land use management alter pH, moisture and nutrient availability, which are important factors influencing the structure and functionality of microbial communities (NAVARRETE et al., 2013; TAKETANI; TSAI, 2010). The Amazon rainforest is an important reservoir of biomass and biodiversity, where a complex variety of physical, chemical and biotic processes of different directions and intensities occur, contributing to global dynamics (ODUM, 1988; BERNOUX et al., 2002; PINTO-JÚNIOR et al., 2009). In this ecosystem, deforestation for conversion to pasture and cultivation areas has altered the structure of the archaeal community and the structure and diversity of the ammonia-oxidizing Archaea (AOA) (TAKETANI; TSAI, 2010; NAVARRETE et al., 2011; PAULA et al., 2014).

To evaluate the theoretical relationships between archaeal groups in the soil community and soil physicochemical factors, archaeal community composition was revealed using high-throughput DNA sequencing of the amplicon 16S rRNA gene in soil samples from primary forest, secondary forest, agricultural systems of indigenous people, and cattle pasture in the western Brazilian Amazon. The soil physicochemical factors were determined based on the same soil samples used for the molecular analysis of the archaeal community. We hypothesized that the archaeal community structure in soils under different land use systems in the Amazon would be defined based on the relationships between taxonomic groups of Archaea and between archaeal groups and soil physicochemical factors, with differences in these relationships for soils under primary forest, secondary forest, agriculture and pasture.

4.2. Material and Methods

4.2.1. Field sites and soil sampling

The soil samples used in this study were collected in March 2008 and January 2009 during the regional wet season. The studied sites are located in Benjamin Constant, a municipality in the Amazonas state, Brazil (between 4°35´ and 4°42´S and 69°60´ and 70°01´W), along the Solimões River (COELHO et al., 2005; FIDALGO et al., 2005).

The climate in the region is equatorial humid, classified as Af (Koppen´s classification), with an annual average temperature of 25.7 °C and an average precipitation of 2810 mm (COELHO et al., 2005, SALDANHA et al., 2018). At all the sampling points considered in this study, the soil type was Inceptisol (COELHO et al., 2005). The deforestation rate in this region is low, mainly due to limited access and low population density (INPE, 2009). The dominant land use and cover types in the four different landscapes studied were primary forest (PF), secondary forest (SF), agricultural systems of indigenous people (AG) and pasture (PA). In this region, the agricultural systems of indigenous people are primarly based on slash-and-burn practices, the shifting cultivation of annual crops and long fallow periods, which involves the abandonment of areas to allow natural regeneration and a fallow period of approximately three years (FIDALGO et al., 2005). Some pasture areas are present because of governmental policies implemented in the 1970s. Amendments, fertilizers and pesticides were not applied in any of the land use systems before soil sampling. The samples were collected from the 0-20 cm topsoil layer using a cylindrical sampler after removing the litter layer. Three soil samples were collected per sampling period from each sampling site: PF, SF (5-20 years old), AG and PA. The samples were kept refrigerated for transportation and frozen at -20 °C until processing for molecular analyses. The soil samples were kept at room temperature for physicochemical analyses.

4.2.2. Analyses of soil physicochemical factors

The soil samples were air dried at room temperature and analyzed according to (COELHO et al., 2005). The organic matter (OM) was obtained with $Na_2Cr_2O_3$ 4N + H_2SO_4 10N oxidation, and pH was measured with a glass electrode in a soil-liquid suspension with 0.01 M $CaCl_2$. The soil fertility factors (phosphorus (P), potassium (K), iron (Fe), manganese (Mn) and zinc (Zn)) were extracted with a Mehlich I extractor (0.05 N HCl and 0.025 N H_2SO_4 solution). Calcium (Ca) and magnesium (Mg) were analyzed using 1N KCl extraction at a 1:20 ratio, and potential acidity (hydrogen (H) + aluminum (Al)) was analyzed with an SMP extractor. For particle size analysis, 4% NaOH was employed as a dispersant and stirred in fast rotation for 15 minutes. Coarse sand (0.2-2 mm) and fine sand (0.05-0.2 mm) were separated by sieving, clay (<0.002 mm) was separated by pipetting and sedimentation and silt (0.002-0.05 mm) was separated from the sand and clay fractions.

4.2.3. DNA extraction and amplicon sequencing of the 16S rRNA gene

Total DNA was extracted from 250 mg of soil using a Power Soil DNA Isolation Kit (MoBio Laboratories Inc., Carlsbad, CA, USA) according to the manufacturer's instructions. The concentration and purity of the soil DNA were determined spectrophotometrically (NanoDrop Technologies Inc., Wilmington, DE, United States). The DNA samples from each of the four land use systems corresponding to both sampling periods (2008 and 2009) were considered biological replicates and represented six DNA replicates for each land use system. The prokaryotic primers 515F and 926R, which amplify the V4-5 variable region of the 16S rRNA gene, were used for PCR amplification, as described in Parada, Needham, Fuhrman et al. (2015). The equipment used for amplicon sequencing was a MiSeq Personal Sequencing System by Illumina (Illumina, Inc., San Diego, CA, United States) operated to generate 2 x 300 bp paired-end reads at the Argonne National Laboratory (Illinois, United States). In summary, we obtained an average of 51177 genomic sequences per sample.

4.2.4. Sequence processing and taxonomic assignment

All the downstream analyses of the raw Illumina sequences were carried out in the QIIME 2 (version 2017.11) environment (BOLYEN et al., 2019). First, the paired-end reads were merged (FLASH R1.fastq and R2.fastq) to produce consensus sequences. The barcode and primer sequences were removed. Quality control of the consensus sequences was performed using a minimum *Phred* quality score of 20 for each base call (EWING; WEND; GREEN, 1998) in DADA2 software, which is used to produce dereplicated and chimera-free sequences (CALLAHAN; McMURDLE; HOLMES, 2017) without operational taxonomic unit (OTU) clustering. Sequence identification was performed using the q2-feature-classifier (BOKULICH et al., 2018a) and the SILVA-ARB database v.132 (QUAST et al., 2013) trained to the primer set 515F/926R (PARADA; NEEDHAM; FUHRMAN, 2015), with 99% stringency to identify the sequences belonging to the Archaea domain.

4.2.5. Construction of integrated networks of archaeal classes and soil physicochemical factors

The DNA sequences corresponding to the twenty-four samples (a total of 6 samples for each of the four land use types) were used to classify the Archaea at the class level. To visualize the connections between the different archaeal classes and between the archaeal classes and soil

physicochemical factors, a co-occurrence network was constructed. Spearman rank correlation coefficients were calculated (significance: $p < 0.05$) for the factors in the R environment with the "multtest" package: *i*. the relative abundances of the archaeal classes and *ii*. the relative abundances of the archaeal classes and soil physicochemical factors (POLLARD; SANDRINE; van der LAAN, 2005). Spearman rank correlation coefficients from all the dataset combinations with $p < 0.05$ (positive and negative) were used for network construction. The constructed network was integrated into different application modules and analyzed using *Gephi* software version 0.9.2 (BASTIAN; HEYMANN; JACOMY, 2009), which has a graphical interface. The nodes in the network represented the taxa at the class level and the physicochemical factors, whereas the edges represented significant positive or negative correlations between the nodes. The network graphs were based on a set of statistical measures and parameters, such as the numbers of nodes and edges, modularity, the clusterization coefficient, the degree and betweenness centrality distributions, training algorithm parameters and activation functions (BRANDES, 2001; BLONDEL, 2008). Betweenness centrality, which reflects the number of times a node serves as a connector along the shortest path between two other nodes, can indicate the relative importance of each node in a network. The clustering coefficient is a measure of interconnectivity in the neighborhood of a node (KHAN et al., 2019).

4.2.6. Quantitative PCR targeting the total archaeal community

Total archaea in each of the 24 soil samples were quantified using quantitative real-time PCR (qPCR) and the domain-specific primer set 519F-915R (KLINDWORTH et al., 2013) for 16S ribosomal RNA gene amplification. Each 10µL reaction mixture contained 5 µL of Power SYBR Green PCR Master Mix (Applied Biosystems, USA), 5 pmol of each primer and 1 ng of each sample. The reactions were performed by an Applied Biosystems 7300 Real-Time Quantitative PCR System, and the conditions were experimentally determined according to Klindworth et al. (2013). Autoclaved milli-Q water was used as a negative control in each reaction. A standard curve (10^2 to 10^7 copies per reaction) was generated containing a PCR-amplified fragment for the Archaea domain (DSMZ 23604/*Methanolinea mesophila*). The reaction efficiency (E) was determined with the equation $E = 10^{(-1/slope)}$.

4.2.7. Statistical analyses

We used analysis of similarities (ANOSIM) to determine whether to join the samples collected in 2008 and 2009 for each of the four land use types. This test was performed with PAST3 software (HAMMER; HAEPER; RYAN, 2001). Additionally, to test for differences in the archaeal community between land use systems using archaeal abundance matrices, analysis of variance (ANOVA) and Tukey's least significant difference (LSD) tests were performed in R software with the "vegan" package (OKSANEN; KINDT; O´HARA, 2013). In addition, the "corrplot" package (TAIYUN; SIMKO, 2017) in R was used to calculate the r value from the Spearman rank correlations between microbial groups and soil physicochemical parameters in each land use system.

4.3. Results

4.3.1. Similarity test between the sampling periods

The ANOSIM revealed no differences in the Archaea groups and soil physicochemical factors between the samples from different sampling periods (March 2008 and January 2009), which justified considering them as replicates of each of the four land use systems (S1 Table).

4.3.2. Analysis of the soil chemical and physical factors

Although the values of the soil chemical and physical factors varied among the sites (S2 Table), most differences were not statistically significant ($p < 0.05$). Only one soil factor, the physical factor "clay", which is associated with soil texture, showed a significant difference between SF and PF (359 and 330 g/kg, respectively). The SF site contained the largest amount of clay among the land use types. In addition to the change in this factor, changes in the values of fine and coarse sand were noteworthy: when converting areas from forest to pasture, the amount of fine sand increased (~ 170 to 357 g/kg), and the amount of coarse sand decreased (~ 12 to 7 g/kg). The K values also increased when we compared forest areas with pasture and agricultural areas (~ 1.35 to 1.85 $mmolc.dm^{-3}$).

4.3.3. Relative abundances of archaeal classes in the treatments

Amplicon 16S sequencing generated 20566 sequences. Identification with SILVA-ARB v.128 revealed twelve archaeal groups, which were organized at the class level (Fig 1). Two main clusters were generated according to group predominance at the sampled sites: Parvarchaeota uncultured archaeon, Methanobacteria and Methanomicrobia in Euryarchaeota, Bathyarchaeota unculture_archaeon and the Thaumarchaeota FHMa11 terrestrial group formed the first cluster with dominance in PA. The last two classes, Bathyarchaeota and the FHMa11 terrestrial group were very abundant in the soils from all sites (10 to 19% of the total archaea identified), and at a specific site (PA), accounted for 35 and 23%, respectively. Increases of 350% in Bathyarchaeota ($p < 0.05$) and >100% in Methanomicrobia were detected at the PA site. The other cluster was composed of Woesearchaeota_DHVEG-6, Lokiarchaeota_uncultured archaeon, Miscellaneous Euryarcheotic group_uncultured archeon (MEG), Euryarchaeota_Thermoplasmata, candidate division YNPFFA_uncultured archaeon, the Thaumarchaeota soil Crenarchaeotic group (SCG) and South Africa Gold Mine-Group 1 (SAGMCG-1). The SCG and SAGMCG-1 were the most abundant Archaea at the sites (over 20%). The second cluster was concentrated in the PF and SF sites, and the SCG significantly increased by 400% ($p < 0.05$) at the SF site. SAGMCG-1 decreased by 120% in PA; however, the change was not statistically significant. SCG was significantly reduced in PA ($p < 0.05$). At the AG site, the Crenarchaeota candidate division YNPFFA and Euryarchaeota_Thermoplasmata were predominant, accounting for 16% and 25%, respectively, of the total archaea. YNPFFA was represented by twice as many individuals at the AG site compared with the PA site, where this class had a lower abundance. However, none of these results were statistically significant. Instead, the SCG was predominant (55%) and significantly more abundant ($p < 0.05$) at the AG site compared with the other sites. Curiously, the FHMa11 terrestrial group was the only Thaumarchaeota class that increased in abundance (by 63%) in the area of forest converted to pasture; however, this increase was not statistically significant.

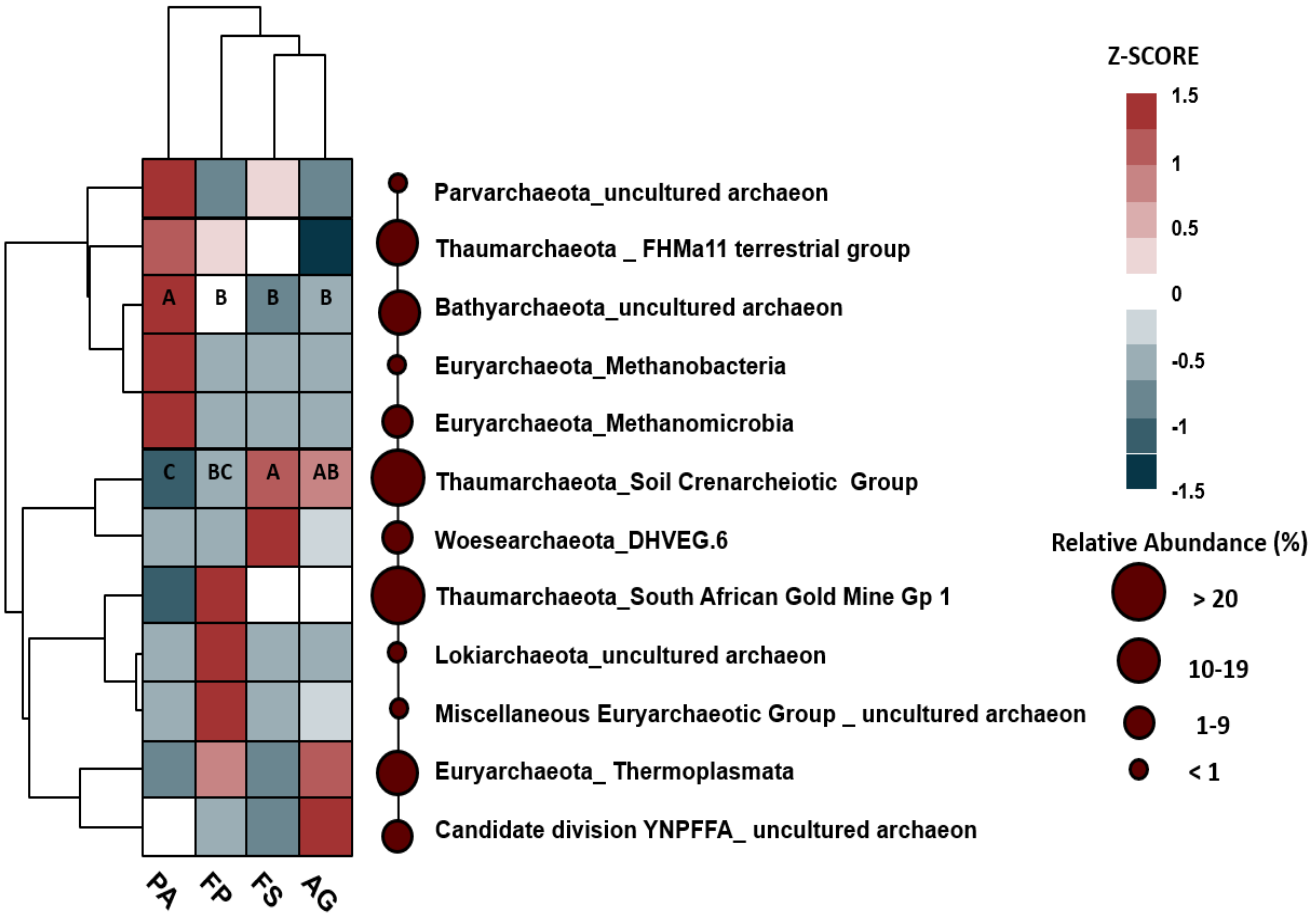


Fig 1. Heatmap created in the R environment with the "pheatmap" package showing the archaeal class abundances in Amazonian soils from primary forest (FP), secondary forest (FS), cropland (AG) and pastures (PA).

4.3.4. Analysis of correlations for land use types

When all the correlations between the Amazonian land use types were observed, a relatively large number of correlations between the Archaea and soil factors was obtained (Fig 2, Table 1). However, when analyzing each land use type separately, the number of correlations (Archaea *vs.* Archaea or Archaea *vs.* soil physicochemical factor) was variable.

Table 1 Total number of correlations per site between the archaeal classes and between the archaeal classes and soil factors

Site	Total correlations (%)*	Archaea vs archaea (%)**	Archaea vs soil factors (%)***
Primary forest	8	50	50
Secondary forest	46	27	73
Agriculture	17	20	80
Pasture	21	42	58

*Proportion of the total number of correlations within each land use type. **Proportion of the total number of correlations between the archaeal classes (i.e., biotic correlations) within each land use type. *** Proportion of the total number of correlations between the archaeal classes and soil factors (i.e., abiotic correlations) within each land use type. Note: the values are derived from the results displayed in the heatmaps.

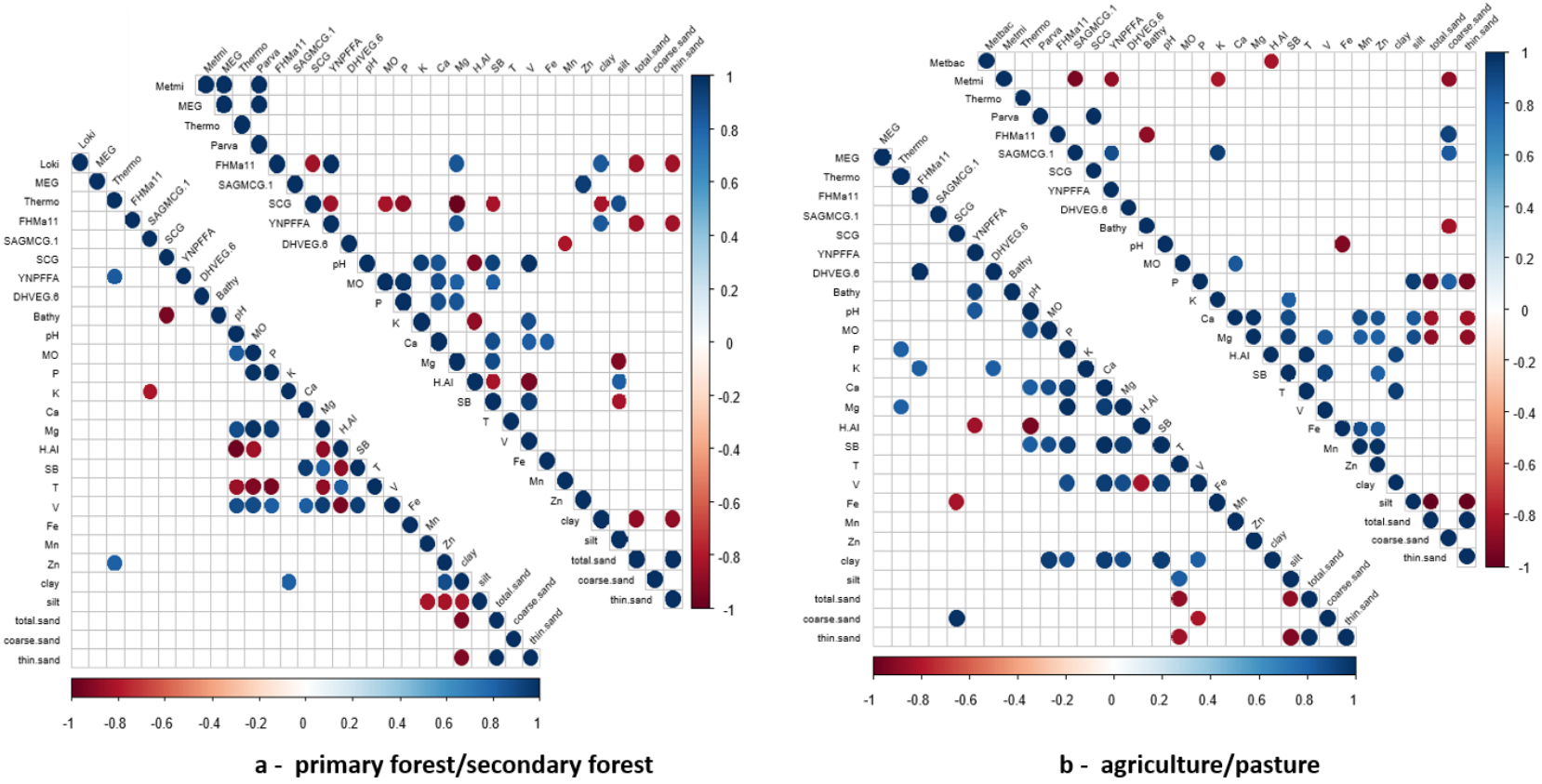


Fig 2. Heatmap of correlations between the archaeal classes and the chemical and physical soil properties, according to significant ($p < 0.05$) Spearman rank correlation coefficient values based on tests performed in the R environment with the "corrplot" package. Positive values (blue circles) represent positive correlations, and negative values (red circles) represent negative correlations. a: primary forest in the lower panel and secondary forest in the upper panel; b: agriculture in the lower panel and pasture in the upper panel.

4.3.5. Correlations between the archaeal classes and soil physicochemical factors in the land use types

The correlation analyses and heatmaps revealed that the AG and SF areas had the largest percentages of Archaea and soil factor correlations, with 80% and 73%, respectively (Fig 2, Table 1). At the PA site, correlations between Archaea and texture were prevalent, accounting for 57% of the total Archaea *vs.* soil factor correlations in this land use type. Unlike at the PA site, correlations between Archaea and soil chemical factors were prevalent at the AG site, corresponding to 88% of the total correlations of this type.

The chemical soil factor most prevalent among the correlations was K, which was present in 25% of the total Archaea *vs*. soil factor correlations, followed by Mg, at 20%. The correlations involving K were present at all sites, except, curiously, the SF site, where most of the correlations between the Archaea and soil factors occurred. Moreover, the correlations involving Mg were most common at the SF site, at 37%. Accounting for only 5% of the correlations between total Archaea and soil chemical factors, correlations with pH, Fe, OM and Mn were less abundant. In relation to correlations between the Archaea and physical soil factors (texture), coarse sand was present in the majority of the correlations (33%), including all texture correlations at the AG and PA sites. At the SF site, clay was the most prevalent (37%), and no correlations were detected between Archaea and coarse sand. Thaumarchaeota SAGMCG-1 was the class with the greatest number of correlations with K (40% of the total correlations) at the PF and PA sites. The SAGMCG-1, FHMa11 and SCG Thaumarchaeota classes were found in 75% of the correlations between archaea and coarse sand, followed by Methanomicrobia in Euryarchaeota, found in 25%. The Thaumarchaeota SCG class was the group with the most correlations between Archaea and soil factors, with 22% of the total correlations. Crenarchaeota candidate division YNPFFA and the Thaumarchaeota FHMa11 terrestrial soil group were the second most common taxa in the correlations between the Archaea and soil factors (19% of the total correlations). Together, these classes composed 60% of the correlations between the Archaea and soil factors.

4.3.6. Correlations between the archaeal classes in each land use type

The correlation analyses ($p < 0.05$) visualized with heatmaps (Fig 2, Table 1) revealed that the land use type with the most correlations between archaeal classes was PF (50% of the total correlations at this site), followed by PA (42%). A minor number of biotic (Archaea *vs*.

Archaea) correlations was observed at the AG (20%) and SF (27%) sites. Interestingly, at the PF and PA sites, where there was more balance between the correlations (Archaea *vs.* Archaea and Archaea *vs.* soil factors), there was also a greater balance between negative and positive biotic correlations. The PF site presented 50% positive correlations between archaeal classes, while the PA site presented 40%. At each of the SF and AG sites, the number of positive correlations was approximately 80%.

The Crenarchaeota candidate division YNPFFA was present in most of the correlations (40% of the total) and occurred in correlations at all sites. With 27% of the total correlations, Thaumarchaeota FHMa11, the SCG and Methanomicrobia in Euryarchaeota were the second most common groups in the correlations. Curiously, FHMa11 was found in correlations at all the sites except the PF site.

In general, the Thaumarchaeota classes most common at the PF and SF sites were negatively correlated with those most abundant at the PA site. For example, there was a negative correlation between the Bathyarchaeota and Thaumarchaeota SCG (Fig 2) at the PF site. However, at the PA site, where Bathyarchaeota and FHMa11 were the most abundant, they were negatively correlated. At the SF site, FHMa11 showed a negative correlation with the other Thaumarchaeota, the SCG. Nevertheless, at the SF site, YNPFFA showed a negative correlation with the SCG. At the AG site, there was a positive correlation between YNPFFA and Bathyarchaeota. Methanomicrobia (Euryarchaeota) showed a negative correlation at the PA site with Thaumarchaeota SAGMCG-1 and Crenarchaeota uncultured division YNPFFA.

4.3.7. Co-occurrence network between the archaeal classes and between the archaeal classes and soil physicochemical factors

The combination of the taxonomic and soil physicochemical factor data and the network based on all the significant correlations from Spearman's nonparametric test resulted in a network with 21 nodes, 40 edges and modularity of 2.802 (Fig 3). The constructed network improved the visualization and interpretation of the possible relationships between the archaeal classes and the biotic and abiotic factors in the soil. According to the network, the greatest number of observed correlations occurred between the archaeal classes and soil factors (~70% of the total). The greatest proportion occurred between the archaeal classes and chemical factors (60% of the total correlations between the Archaea and soil factors). The nodes with the largest betweenness centrality values represented the candidate division YNPFFA and FHMa11 terrestrial group classes, with values of 26 and 8, respectively. When this analysis was

conducted for each site to verify these connections under specific conditions, only the SF site presented nodes with betweenness centrality values for the classes YNPFFA and SCG (with 3 and 2, respectively) (S1 Fig). The largest clustering coefficient value was 0.5, which was found for the node of Bathyarchaeota. The clay texture factor had the same value (Fig 3). Curiously, all the Thaumarchaeota classes, YNPFFA, Thermoplasmata and, DHVEG-6, presented positive and negative correlations. However, Methanobacteria and Methanomicrobia in Euryarchaeota and Bathyarchaeota presented only negative correlations with biotic or abiotic factors.

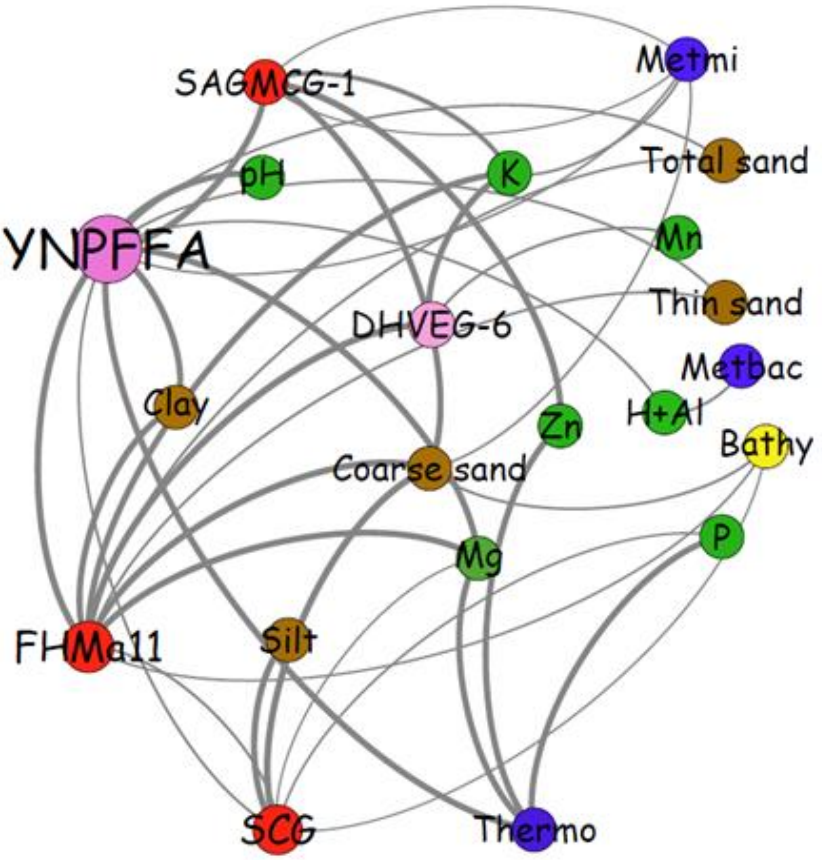


Fig 3. Network correlation analysis of the co-occurrence of archaeal classes and chemical and physical soil properties based on significant ($p < 0.05$) Spearman rank correlation coefficient values. The nodes represent the soil chemical factors (green); soil physical factors (brown); Thaumarchaeota classes (red; FHMa11–FHMa11 terrestrial group, SAGMCG-1–South Africa Gold Mine Group 1, and SCG–soil Crenarchaeota group); Euryarchaeota classes (blue; Thermo–Thermoplasmata, Metbac–Methanobacteria, and Metmi–Methanomicrobia); DHVEG-6–Woesearchaeota (pink); Bathy–Bathyarchaeota (yellow); and YNPFFA–Crenarchaeota candidate division YNPFFA (purple). The thick edges represent positive correlations, and the thin edges represent negative correlations. The sizes of the nodes are scaled according to their betweenness centrality values to indicate their relative importance in the network.

4.3.8. qPCR

The largest number of copies of the 16S ribosomal gene was obtained from the PF site (~ 1125 copies/ng of DNA), followed by the PA site (~ 1000 copies/ng of DNA) (Fig 4). The SF and AG sites presented the smallest numbers, with approximately 300 and 400 copies/ng of DNA, respectively.

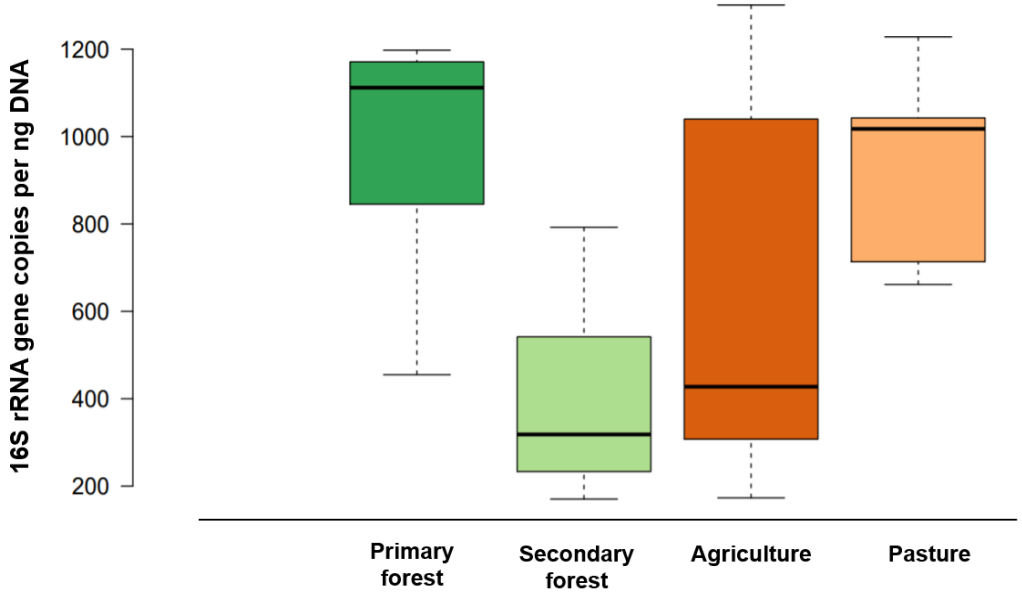


Fig 4. Boxplot of the number of archaeal 16S ribosomal RNA gene copies in each nanogram (ng) of DNA extracted from the six soil samples obtained in the Amazon from sites in primary forest, secondary forest, agricultural areas and pastures. This plot was created in the R environment with the "ggplot2" package.

4.4. Discussion

The 515F and 926R primers (PARADA, NEEDHAM, FUHRMAN, 2015) which amplify the V4-5 variable region of the 16S rRNA gene, were used for PCR amplification and provide safer and more reliable prokaryote identification and abundance results without overestimating certain groups compared with other primers. This lends more validity to the discussion of the results of this study, in which different classes of soil Archaea sampled from four areas in western Amazonia were quantified and the abundances of each class in each area were correlated with biotic (co-occurrence of archaeal groups) and abiotic (chemical factors and soil texture) factors.

Our results showed that at the PF site, the number of correlations between the archaeal classes was proportionally higher than that at the SF, AG and PA sites (Fig 2, S1 Fig). This result coincides with those obtained in the quantification analysis of 16S rRNA gene copies (Fig 4), which indicated that the largest abundance of Archaea representatives occurred at the PF site. A higher abundance of individuals could favor the occurrence of relationships between groups, which would be especially important as a survival strategy in an oligotrophic environment, such as the soil in the Amazonian PF (NAVARRETE et al, 2011; KHAN et al., 2019). Certain microorganisms are functionally complementary, living in syntrophy, in which case it is possible to use a minimal amount of the nutrients available in an environment through

combined metabolic action (MORRIS et al., 2013). The same principle can be used to explain the results obtained at the PA site, which was the area with the second largest number of copies of the 16S rRNA gene (Fig 4) and the largest number of correlations between the Archaea (Table 1, Fig 2). The physicochemical characteristics of the soil environment at the PA site are contradictory to those at the PF site (SUZUKI et al., 2014), which constitutes a more stable environment in this regard. Although the diversity of the Archaea at the PF site relative to that at the PA site (NAVARRETE et al., 2011) may be an important factor for the occurrence of biotic interactions between archaeal classes, our results show that the abundance of individuals is a more relevant factor. In addition to the abundance of individuals, the composition of the archaeal community in each of these areas in the Amazon could also influence the number of interactions between archaeal classes and between archaeal classes and soil factors. According to the physiological needs of a group, syntrophic relationships between specific groups are essential for survival (MORRIS et al., 2013). In this sense, the occurrence of a specific partner group is fundamental for a metabolic process to occur (MORRIS et al., 2013). For example, this is the case with a sugar-fermenting bacterium living in an oligotrophic environment at the bottom of a lake, which is dependent on another (methanogenic) bacterium for hexoses (MULLER, GRIFFIN, STING, 2008).

In the areas with the largest numbers of correlations between the Archaea classes (Table 1; S1 Table) (PF and PA), the Thaumarchaeota SCG and SAGMGC-1 were dominant (in PF). Members of the latter group are responsible for NH_3 oxidation (DUTTA et al., 2019). Methanomicrobia and Bathyarchaeota (Euryarchaeota), associated with methanogenesis (DUTTA et al., 2019), predominated at the PA site. The completion of both physiological processes requires interactions between different groups of microorganisms (MORRIS et al., 2013; YIN; BI; XU, 2018). The reduction in the abundance of the Thaumarchaeota SCG and SAGMGC-1 at the PA site may result in a decrease in the NH_3 oxidation process in this area (NAVARRETE et al., 2011). Therefore, the co-occurrence of these groups would not only coordinate community structuring but also indicate the functionality of these agroecosystems, including the importance of archaeal co-occurrence.

At the AG and SF sites, where the number of correlations between the archaeal classes and soil factors was larger than the number of correlations between archaeal classes (Table 1, Fig 2, Fig S1), the number of 16S rRNA ribosomal gene copies was also lower than that at the PF and PA sites (Fig 4). At the SF site, a large amount of clay, which promotes a greater cation exchange capacity and consequent availability of nutrients, OM and moisture retention (DON; SCHULZE, 2008; KAISER; ZECH, 2000; SHI; MARSCHNER, 2013), may favor

groups capable of using these resources directly through independent soil and microbial physiologies (MORRIS et al., 2013). Interestingly, the NH_3-oxidizing Thaumarchaeota FHMa11 terrestrial group and SCG (DUTTA et al., 2019) were positively correlated with the large quantities of clay found in this area (S2 Table). The FHMa11 terrestrial group and SCG predominate in forest regeneration areas, even at different soil depths (O´SULLIVAN et al., 2012), where they have also been associated with texture and pH (O´SULLIVAN et al., 2012). The FHMa11 terrestrial group was negatively related to pH and positively related to the SCG. In our study, the SCG was the most abundant class at the SF site, and although changes in the values of soil chemical factors were not statistically significant, larger pH values were obtained at the AG and SF sites than at the PF site (S2 Table). At the PA site, the site with the lowest pH value (S2 Table), the FHMa11 terrestrial group increased in abundance (Fig 1), corroborating the results of the previous study. Similarly, in a study conducted in Australian pasture systems (O´SULLIVAN et al., 2012), the AOA community was assessed, and pH was inversely correlated with AOA *amo*A gene abundance in the soil. In these Amazonian soils, however, the lowest pH value was found at the PA site, which leads us to conclude, after observing the Australian results, that this area has the largest number of archaeal 16S rRNA gene copies. However, there was a reduction in the SCG and SAGMCG-1 classes at this site and the total amount of Archaea. It is therefore assumed that the terrestrial soil FHMa11 group composed the AOA community present in the Australian pasture.

In addition to the terrestrial soil FHMa11 group, the Crenarchaeota candidate division YNPFFA is also likely to be associated with pH; however, the highest abundance of both classes was found at the AG site, along with positive correlations (Fig 1, S1 Fig). Information about the physiology of the Crenarchaeota candidate division YNPFFA, which could provide us with insights into the functional role of this class in these Amazonian soil ecosystems, is lacking.

Other soil factors related to texture, such as the availability of coarse and fine sand, presented high potential relevance to the distribution of archaeal classes. Coarse sand showed reduced values at the PA site compared with the other sites, while fine sand increased at the PA site (S2 Table). This occurs in soils as a result of the translocation of sand in soil layers due to changes in soil use (PETR; JOSEF, 2018). This process alters aggregate structure and porosity (REINERT; REICHERT, 2006), with a consequent reduction in O_2 availability (STEFANOSKI et al., 2013). This condition may promote the redistribution of microbial groups in the soil. Recent molecular analysis using the 16S rRNA gene revealed that there is a greater diversity of bacteria where there are more clay particles because it retains more moisture (KIM; CROWLEY, 2013; TISDALL; SMITH; RENGASAMY, 1997). Sand particles

(coarse or fine), which retain less water than clay, attract more Archaea (KIM; CROWLEY, 2013). This explains the increase in the abundance of the Bathyarchaeota and Methanomicrobia classes at the PA site. However, in pasture soil, there is a reduction in total porosity, predominantly for particles with a smaller diameter, due to animal trampling; for example, higher water contents accumulate in the 0.20-0.225 m layer (SUZUKI et al., 2014). These conditions in the PA soil could have direct implications for archaeal groups, especially those related to the methanogenesis process (which is anaerobic) (OFFRE; SPANG; SCHLEPER, 2013). This explains the increase in the abundance of the Bathyarchaeota and Methanomicrobia classes at the PA site. It has been suggested that Bathyarchaeota carry out acetogenesis (acetate production) (McINERNEY; SIEBER; GUNSALUS, 2009; HE et al, 2016; LAZAR et al., 2016; YU et al., 2018), generating a substrate for methanogenesis, which could be performed by Methanomicrobia, indicating a syntrophic relationship.

However, if the arrangement and oscillation of the amounts of macro- and micropores in the soil aggregates resulting from particle movement can contribute to the success of certain archaeal groups, then they may also promote the failure of others, such as Thaumarchaeota, which is involved in the aerobic process of NH_3 oxidation. In this case, ammonia-oxidizing bacteria (AOB), acting under anaerobic conditions, would have an advantage over AOA. In addition, methanogenic Archaea may directly influence this process. Studies have reported diazotrophic methanogens (LEIGH, 2000; BAE et al., 2018), which may be responsible for the availability of large amounts of organic N in the system (BAE et al., 2018). In oxide soils, methanogenic N_2 fixation may be the most important route of N_2 availability, consequently impacting NH_3 demand by AOA and AOB (BAE et al., 2018). In particular, among the four land use types, the NH_3 concentrations were highest at the PA site (NAVARRETE et al., 2011), with an approximately 40% increase over the NH_3 values found in the forest areas. AOA are known to have impaired NH_3 oxidation abilities when subjected to high NH_4 concentrations (LI et al, 2010). This could explain the reduction in SAGMCG-1 and SCG AOA in this area.

Candidate division YNPFFA and terrestrial group FHMa11 showed the highest values of betweenness centrality in the correlation networks (Fig 3, S1 Fig), suggesting that the presence of these groups in the Amazonian soils, especially at the SF and AG sites, is highly relevant to the dynamics of these soil systems (KHAN et al., 2019). Interestingly, terrestrial group FHMa11, a group of AOA, was the only Thaumarchaeota class that did not decrease in abundance at the PA site, increasing instead (Fig 1). Unfortunately, although studies using independent methods and cultivation techniques have provided much information about uncultured microorganisms in recent years, there is not enough physiological information about

this class to allow us to understand the success of Thaumarchaeota under these conditions. Recently, based on genomic evidence, some authors suggested that certain Thaumarchaeota might participate in methanogenesis. HUA et al. (2019) studied methyl coenzyme M-reductase (Mcr) genes in assembled archaeal genomes from metagenomic analysis of samples collected from recent sources in China. The study revealed the expansion of the archaeal groups that participate in CH_4 metabolism. Among them, Thaumarchaeota contains *mtr*AH genes, suggesting the metabolism of CH_4 precursor methylated compounds. This information indicates their close association with methanogenic Euryarchaeota and Bathyarchaeota. In this context, the co-occurrence of terrestrial group FHMa11 with Bathyarchaeota and Methanomicrobia at the PA site in this study may be related to the ecological role of these Archaea in the sampled Amazonian soils. Another important piece of information supporting this possibility is the fact that these Archaea seem to be adapted to higher concentrations of NO_3 and NH_4, which were observed in this area. In general, due to their oligotrophic lifestyle, microorganisms such as AOA are unable to live in environments with excess nutrients (MULLER; GRIFFIN; STING, 2008). The co-occurrence of terrestrial group FHMa11 with Bathyarchaeota and Methanomicrobia was reported in a study evaluating archaeal communities in igneous rocks in different horizons in India (DUTTA et al., 2019). In the granitic horizon, where large proportions of Thaumarchaeota and an uneven distribution of Euryarchaeota and Bathyarchaeota are affiliated with Methanomicrobia and terrestrial group FHMa11 were found, the sources of organic C were very low (4-48 $mg.kg^{-1}$), and the sources of NO_3 were variable.

Previous studies conducted on indigenous peoples' agricultural systems and cattle pastures in the same areas of western Amazonia indicated that the bacterial community was affected by changes in chemical factors (JESUS et al, 2009), and archaeal communities can be changed under these conditions, but it is unknown which factor is the most important (NAVARRETE et al., 2011). In assessing how the process of forest conversion to pasture in western Amazonia can affect co-occurrence patterns in prokaryotic communities, KHAN et al. (2019) revealed new information on changes in metabolic pathways related to the N cycle. Soil physicochemical properties, such as temperature, the C/N ratio and $H^+ + Al^3$, had a significant impact on bacterial communities in the two contrasting ecosystems. However, only the Archaea Crenarchaeota, which is abundant in the forest (3.6%) (where the study was conducted), responded to the changes, presenting a strong positive correlation with $H^+ + Al^3$. As evidenced in a previous study by Delgado-Baquerizo et al. (2016), changes in the structure of the microbial community in the PF area could compromise specific archaeal functions, such as N_2O greenhouse gas production and denitrification. At the GA and SF sites, where the community

is dependent on soil properties, basic functions such as respiration by microorganisms could be altered according to the nutrient availability in the soil. This would explain the smaller number of Archaea in these areas compared to those at the PF and PA sites.

Soil is a complex environment for the survival of Archaea. Given the inherent characteristics of these organisms, they evolved in extreme environments, which required evolutionary adaptations essential to survival. Due to the need to obtain energy, the Archaea found in extreme environments need to perform more functions. By observing the genomes of these microorganisms, we can better understand how they live. Bathyarchaeota is known to have larger genomes than methanogenic Euryarchaeota (KELLNER et al., 2018). Having small genomes can mean major limitations in terms of functionality in an environment, a problem that can be managed by performing syntrophy with organisms that have more genes (MOISSL-EICHINGER et al., 2018; MORRIS et al., 2013). The Thaumarchaeota terrestrial group FHMa11, which co-occurs with the Bathyarchaeota and methanogenic Euryarchaeota in western Amazon soils, has the potential to be syntrophic. In the near future, the answer to this question could be obtained through microbial ecology.

4.5. Conclusions

In natural environments exposed to less anthropogenic impact, such as the primary forest in the western Amazon, the occurrence of several highly abundant groups was more important than soil properties in structuring the archaeal community. Interaction between archaeal groups is essential for their acquisition of energy and the performance of their functions. When soil nutrient availability increases and soil physical structure changes, which can happen at a secondary forest, agricultural and pasture sites, the number of interactions between Archaea and abiotic factors can be sustained, favoring the selection of groups that are able to meet the demands of the new conditions.

References

ALVES, R. J. E. et al. Unifying the global phylogeny and environmental distribution of ammonia-oxidising archaea based on amoA genes. **Nature Communications**, London, v. 9, n. 1, p. 1–17, 2018.

BAE, HEE-SUNG et al. Methanogens are major contributors to nitrogen fixation in soils of the Florida Everglades. **Applied Environmental Microbiology**, Washington, v. 84, p. 7 e02222-17. 2018.

BASTIAN, M.; HEYMANN, S.; JACOMY, M. Gephi: an open source software for exploring and manipulating networks. **International Association for the Advancement of Artificial Intelligence Conference Weblogs Social Media**, San Jose, CA, v. 8, p. 361-362. 2009.

BERNOUX, M. Brazil's soil carbon stocks. **Soil Science Society America Journal**, Madison, v. 66, p. 888-896, 2002.

BLONDEL, V.D. et al. Fast unfolding of communities in large networks. **Journal of Statystical Mechanics**, P10008. 2008. DOI: 10.1088/1742-5468/2008/10/P10008.

BOLYEN, E. et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. **Nature Biotechnology**, London, v. 37, p. 848–857, 2019.

BOKULICH, N.A. et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. **Microbiome**, London, v. 6, p. 90. 2018a. doi: 10.1186/s40168-018-0470-z.

BRANDES, U. A faster algorithm for betweenness centrality. **Journal of Mathematical Sociology**, v. 25, p. 163-177. 2001.

CALLAHAN, B. J.; MCMURDIE, P. J.; HOLMES, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. **The ISME Journal**, London, v. 11, n. 12, p. 2639–2643, 2017.

COELHO, M.R. et al. **Solos das áreas-piloto do Projeto BiosBrasil**. (Conservation and Sustainable Management of Below-Ground Biodiversity: Phase I) Município de Benjamin Constant, Estado do Amazonas. Rio de Janeiro: Embrapa Solos, 2005. 159 p. (Boletim de Pesquisa e Desenvolvimento, 67). Disponível em: www.biosbrasil.ufla.br or http://www.cnps.embrapa.br/solosbr/publicaoes/levantamentodesolos. Acesso em: 21 fev. 2020.

DE CHAVES, M.G. et al. Acidobacteria subgroups and their metabolic potential for carbon degradation in sugarcane soil amended with vinasse and nitrogen fertilizers. **Frontiers in Microbiology,** Lausanne, v. 10, p. 1680, 2019. doi: 10.3389/fmicb.2019.01680.

DELGADO-BAQUERIZO, M. Relative importance of soil properties and microbial community for soil functionality: insights from microbial swap experiment. **Functional Ecology,** Toronto, v. 30, p. 1862-73, 2016. doi: 10.1111/1365-2435.12674.

DELONG, E.F.; PACE, N.R. Environmental diversity of Archaea and Bacteria. **Systematic Biology**, Oxford, v. 50, p. 470-478. 2001.

DON, A.; SCHULZE, E.D. Controls on fluxes and export of dissolved organic carbon in grasslands with contrasting soil types. **Biogeochemistry**, London, v. 91, p. 117-131, 2008.

DUTTA, A. et al. Archaeal communities in deep terrestrial subsurface underneath the Deccan Traps, India. **Frontiers in Microbiology**, Lausanne, v. 10, p. 1362, 2019.

EWING, B.H.L., WEND, M.C., GREEN. P. Base-calling of automated sequencer traces using phred: I. Accuracy assessment. **Genome Research**, New York, v. 8, p. 175-85, 1998. doi:10.1101/gr.8.3.175.

FIDALGO, E.C.C. et al. **Levantamento do uso e cobertura da terra de seis áreas amostrais relacionadas ao projeto BiosBrasil** (Conservation and Sustainable Management of Below-Ground Biodiversity: Phase I), Município de Benjamin Constant (AM). Rio de Janeiro: Embrapa Solos, 2005. 2005. 15 p. (Boletim Pesquisa e Desenvolvimento, 71).

HAMMER, Ø.; HARPER, D.A.T.; RYAN, P.D. Past: Paleontological statistics software package for education and data analysis. **Palaeontologia Electronica**, v. 4, n. 1, p. 4-9, 2001. Available from: https://palaeo-electronica.org/2001_1/past/issue1_01.htm.

HUBER, H. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. **Nature**, London, v. 417, p. 63-67, 2002.

HUA, Z.S. et al. Insights into the ecological roles and evolution of methyl-coenzyme M reductase-containing hot spring Archaea. **Nature Communications**, London, v. 10, n. 4574, 2019. doi:10.1038/s41467-019-125.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS - INPE. Ministério da Ciência e Tecnologia do Brasil. **Relatório de gestão do exercício de 2009.** Prestação de contas ordinária anual. São José dos Campos, 2009. 58 p.

JAHN, U. et al. *Nanoarchaeum equitans* and *Ignicoccus hospitalis*: new insights into a unique, intimate association of two archaea. **Journal in Bacteriology**, Washington, DC, v. 190, p. 1743-1750, 2008.

JESUS, E.C. Changes in land use alter the structure of bacterial communities in western Amazon soils. **The ISME Journal**, London, v.3, p. 1004-1011, 2009.

KAISER, K.; ZECH, W. Dissolved organic matter sorption by mineral constituents of subsoil clay fractions. **Journal of Plant Nutrition Soil Science**, New York, v. 163, p. 531-535, 2000.

KELLNER, S. Genome size evolution in the Archaea. **Emerging Topics in Life Sciences**, London, v. 2, p. 595-605, 2018.

KHAN, M.A.W. Deforestation impacts network co-occurrence patterns of microbial communities in Amazon soils. **FEMS Microbiology Ecology**, Amsterdam, v. 95, p. fiy230. 2019.

KIM, J.S.; CROWLEY, D.E. Size fractionation and microbial community structure of soil aggregates. **Journal of Agricultural Chemistry and Environment**, New York, v.2, p. 75-80, 2013.

KLINDWORTH, A. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. **Nucleic Acid Research**, Oxford, v. 41, p. e1, 2013. doi: 10.1093/nar/gks808

KURAMAE, E.E. Soil characteristics more strongly influence soil bacterial communities than land-use type. **FEMS Microbiology Ecology**, Amsterdam, v. 79, p. 12-24. 2012. doi: 10.1111/j.1574-6941.2011.01192.x.

LAZAR, C.S. et al. Genomic evidence for distinct carbon substrate preferences and ecological niches of Bathyarchaeota in estuarine sediments. **Environmental of Microbiology**, Oxford, v.18, p. 1200–1211, 2016.

LEIGH, J.A. Nitrogen fixation in methanogens: the archaeal perspective. **Current Issues in Molecular Biology**, London, v. 2, p. 125–131, 2000.

LI, M. et al. Spatial distribution and abundances of ammonia-oxidizing archaea (AOA) and ammonia-oxidizing bacteria (AOB) in mangrove sediments. **Applied Microbiology and Biotechnology**, London, v. 89, p. 1243–1254, 2010. doi: 10.1007/s00253-010-2929-0.

MCINERNEY, M.J.; SIEBER, J.R.; GUNSALUS, R.P. Syntrophy in anaerobic global carbon cycles. **Current Opinion in Biotechnology**, Amsterdam, v. 20, p. 623–632, 2009.

MOISSL-EICHINGER, C. et al. Archaea are interactive components of complex microbiomes. **Trends in Microbiology**, Amsterdam, v. 26, n. 1, p. 70-85, 2018. doi:10.1016/j.tim.2017.07.004.

MOISSL-EICHINGER, C.; HUBER, H. Archaeal symbionts and parasites. **Currently Opinion in Microbiology**, Amsterdam, v. 14, n. 3, p. 364-370, 2011.

MORRIS, B.E. et al. Microbial syntrophy: interaction for the common good. **FEMS Microbiology Review**, Amsterdam, v. 37, p. 384–406, 2013.

MUELLER, N.; GRIFFIN, B.M.; STING, S.B. Dominant sugar utilizers in sediment of Lake Constance depend on syntrophic cooperation with methanogenic partner organisms. **Environmental Microbiology**, Baltimore, v.10, p. 1501–1511, 2008.

NAVARRETE, A. A. et al. Land-use systems affect Archaeal community structure and functional diversity in western Amazon soils. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 35, n. 5, p. 1527–1540, 2011.

NAVARRETE, A.A. et al. Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. **FEMS Microbiology Ecology**, Amsterdam, v. 83, p. 607-621, 2013.

ODUM EP. **Ecologia**. Rio de Janeiro: Guanabara Koogan, 1988.

OFFRE, P.; SPANG, A.; SCHLEPER, C. Archaea in biogeochemical cycles. **Annual Review of Microbiology**, Palo Alto, v. 67, p. 437–457, 2013.

OKSANEN, J.; KINDT, R.; O'HARA, B. Package 'vegan'. **Community Ecology Package**, New York, v.2, n.9, p.1-295, 2013.

O'SULLIVAN, C.A. et al. Archaeal ammonia oxidizers are abundant in acidic, coarse-textured Australian soils. **Soil Research**, Clayton South, v. 49, p. 715-724, 2012. doi:10.1071/SR11288

PAULA, F.S. et al. Land use change alters functional gene diversity, composition and abundance in Amazon forest soil microbial communities. **Molecular Ecology**, Amsterdam, v. 23, p. 2988-2999, 2014. doi.org/10.1111/mec.12786.

PARADA, A.E.; NEEDHAM, D.M.; FUHRMAN, J.A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. **Environmental Microbiology**, Oxford, v. 18, p. 1403-14, 2015. doi:10.1111/1462-2920.13023

PETR, L.N.; JOSEF, H. Article translocation of soil particles during secondary soil tillage along contour lines. **Water**, Basel, v. 10, p. 568, 2018. doi:10.3390/w10050568.

PINTO-JUNIOR O.B. et al. Soil efflux CO_2 in mature transitional tropical forest Amazon and pasture area. **Acta Amazônica**, Manaus, v. 39, p. 4, 2009.

PODAR M, ANDERSON I, MAKAROVA KS. A genomic analysis of the archaeal system *Ignicoccus hospitalis –Nanoarchaeum equitans*. **Genome Biology**, London, v. 9, R158, 2008.

POLLARD, K.S.; SANDRINE, D.; VAN DER LAAN, M.J. Multiple testing procedures: the multtest package and applications to genomics. **In: Bioinformatics and computational biology solutions using R and bioconductor**. New York: Springer, 2005. pp. 52.

QUAST, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. **Nucleic Acid Research**, Oxford, v. 41, p. D590-D596, 2012.

RAMPELOTTO, P.H. Extremophiles and extreme environments. **Life**, Basel, v. 3, p. 482-485, 2013. doi:10.3390/life3030482.

REINERT, D.J.; REICHERT, J.M. Coluna de areia para medir a retenção de água no solo - protótipos e teste. **Ciência Rural**, Santa Maria, v.36, p.1931-1935, 2006. doi:10.1590/S0103-84782006000600044.

SALDANHA, L.S. Caracterização morfológica de briófitas no município de Benjamin Constant - AM. **Biota Amazônia. Open Journal System**, Macapá, v. 8, n. 2, p. 48-52, 2018.

SHI, A.D.; MARSCHNER, P. Addition of a clay subsoil to a sandy top soil alters CO_2 release and the interactions in residue mixtures. **Science of the Total Environment**, Amsterdam, v. 465, p. 248-254, 2013.

STEFANOSKI, D.C. et al. Soil use and management and its impact on physical quality. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 17, p. 1301-1309, 2013. doi:10.1590/S1415-43662013001200008.

SUZUKI, L.E.S.A. et al. Estrutura e armazenamento de água em um argissolo sob pastagem cultivada, floresta nativa e povoamento de eucalipto no Rio Grande do Sul. **Revista Brasileira de Ciência do Solo**, Viçosa. v. 38, p. 94-106. 2014.

TAJIMA, K. et al. Phylogenetic analysis of archaeal 16S rRNA libraries from the rumen suggests the existence of a novel group of archaea not associated with known methanogens. **FEMS Microbiology Letters**, Amsterdam, v. 200, p. 67–72, 2001

TAKETANI, R.G.; TSAI, S.M. The influence of different land uses on the structure of archaeal communities in Amazonian anthrosols based on 16S rRNA and amoA genes. **Microbial Ecology**, New York, v. 59, p. 734-743, 2010. doi: 10.1007/s00248-010-9638-1.

TAIYUN, W.; SIMKO, V. **R package "corrplot"**: visualization of a correlation matrix (version 0.84). 2017. Available from https://github.com/taiyun/corrplot

TISDALL, J.M.; SMITH, A.C.; RENGASAMY, P.C. Aggregation of soil by fungal hyphae. **Australian Journal of Soil Research**, Clayton, v. 35, p. 55-60, 1997.

YIN, Z.; BI, X.; XU, C. Ammonia-oxidizing Archaea (AOA) play with ammonia oxidizing Bacteria (AOB) in nitrogen removal from wastewater. **Review Article Hindawi Archaea**, London, v. 2018, p. 9, doi:10.1155/2018/8429145.

YU, T. et al. Growth of sedimentary Bathyarchaeota on lignin as an energy source. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v.115, p. 6022–6027, 2018.

5. FINAL CONSIDERATIONS

The use of a multi-analytical approach including molecular techniques and bioinformatics allowed us improve information about “key” genes, process, and archaeal and bacterial groups associated with N and C cycles in tropical soils. In this sense, this thesis reached the goal to fill the “gaps” found in previous studies that were performed in the same area and soil conditions, which used a non-integrative analytical approach.

The studies performed in this thesis in tropical soils confirmed that the use of N fertilizers as management in sugarcane crop and the conversion of forest to pasture or agriculture in Amazonian promote physicochemical changes that affect specific groups from Archaea and Bacteria associated to N and C cycles. The application of vinasse and N fertilizer in sugarcane soils may promote an increase of pathogenic denitrification *Mycobacterium* and decrease of Gp13 and Gp18 Acidobacteria associated with hemicellulose degradation, which represents an alert around this practice. Besides, this allowed to access essential groups associated with N cycle as the Archaea AOA South Africa Gold Mine – Group 1 (SAGMG-1) and *Nitrosopumillus* that are highly present in natural conditions (primary forest) and agricultural practices as N fertilizers application in soil, respectively. In primary forest, it is possible that the high abundance of Archaea was capable to keep the natural nutrient dynamic in the soil. Unfortunately, in sugarcane crops, the increase of groups AOA/AOB may represent a potential increase in N_2O production. In addition, the results suggest a potential increase of archaeal methanogenic groups, as *Methanomicrobia* and *Methanobacteria*, in the Amazonian pasture.

However, if in soil conditions with fertilizers, with higher salt contents and low levels of O_2, there are losses of N and C as GHG production from microbiota, other resilient groups, with interesting metabolic capabilities, can grow. In the ammonification process under N fertilizers, increases of *Natronomonas* (Archaea halophile) and *Anaeromyxobacter* (a Bacteria able to reduce N_2O) may represent a good perspective for application in biotechnology and sustainable agriculture practices.

APPENDIX

Appendix A. Supplementary Material of Chapter 2

Supplementary Table S1 *Spearman's* correlation values r_s between the Acidobacteria classes and physicochemical factors quantified in the soil samples

Acidobacteria subgroup	OM	P	S	K	Ca	Mg	H+Al	SB	CEC	V	B	Cu	Fe	Mn	Zn	C	C/N	N	pH
Gp1												0.6 **0.008**	0.479 **0.043**	0.45 **0.059**					
Gp2	0.479 **0.043**		-0.54 **0.022**					-0.48 **0.043**											
Gp3																			
Gp4						0.592 **0.01**	-0.494 **0.036**										-0.59 **0.01**	0.56 **0.015**	0.617 **0.006**
Gp5																			
Gp6																	0.523 **0.025**	-0.518 **0.027**	-0.509 **0.031**
Gp7																			
Gp9			-0.48 **0.042**			-0.47 **0.046**													
Gp10																			
Gp13																			
Gp17	-0.53 **0.022**	0.707 **0.001**						0.505 **0.032**		0.5 **0.03**	0.46 **0.06**				0.482 **0.042**				
Gp18					-0.48 **0.042**				-0.558 **0.016**										
Gp21																			
Gp23					-0.6 **0.009**			-0.608 **0.007**		-0.5 **0.03**									
Gp25			0.488 **0.039**	0.559 **0.016**										0.487 **0.039**					
Gp26																			

Spearman's correlation coefficient; Highlighted: significant at $P \leq 0.05$

Supplementary Table S2. *Spearman's* correlation values r_s between Acidobacteria classes and carbon degradation genes based on GeoChip 5.0M detection in samples

Acidobacteria subgroups	AceA	AceB	acetylglucosaminidase	amyA	ara	cda	cellobiase	endoglucanase	chitinase	glucoamylase	pectinase	pme	pula	RgaE	xyla	xylanase
	-0.465		-0.482	-0.478				-0.501	-0.461	-0.558	-0.501			-0.463		-0.511
Gp1	**0.05**		**0.042**	**0.044**				**0.034**	**0.053**	**0.016**	**0.034**			**0.052**		**0.03**
		0.474														
Gp2		**0.046**														
Gp3																
Gp4																
Gp5																
Gp6																
Gp7																
Gp9																
Gp10																
Gp13																
		0.515	0.519	0.505	0.509		0.504	0.539	0.516		0.523	0.56				0.566
Gp17		**0.028**	**0.027**	**0.032**	**0.031**		**0.033**	**0.02**	**0.028**		**0.025**	**0.015**				**0.015**
	-0.564	-0.51	-0.546	-0.602	-0.48	-0.48	-0.614	-0.552	-0.562	-0.664	-0.527	-0.455	-0.455	-0.583	-0.525	-0.562
Gp18	**0.014**	**0.029**	**0.019**	**0.008**	**0.043**	**0.045**	**0.007**	**0.017**	**0.015**	**0.003**	**0.024**	**0.056**	**0.056**	**0.011**	**0.025**	**0.015**
		0.575	0.486	0.454	0.488	0.501		0.017	0.462							0.473
Gp21		**0.013**	**0.04**	**0.058**	**0.039**	**0.034**		**0.017**	**0.052**							**0.046**
Gp23																
Gp25	-0.496		-0.472	-0.488	-0.484		-0.451			-0.562						
	0.035		**0.047**	**0.04**	**0.041**		**0.06**			**0.015**						
Gp26																

Spearman's correlation coefficient; Highlighted: significant at $P \leq 0.05$

Supplementary Table S3. *Spearman* Correlation values r_s between the classes of Acidobacteria samples in the three treatments with level of significance P ≤ 0.05

	Gp1	Gp2	Gp3	Gp4	Gp5	Gp6	Gp7	Gp9	Gp10	Gp11	Gp13	Gp17	Gp18	Gp21	Gp22	Gp23	Gp25	Gp26	Others
Gp1											0.557		0.654						
											0.0159		**0.0032**						
Gp2					0.597			0.584											
					0.009			**0.011**											
Gp3																			
Gp4								-0.518											
								0.0276											
Gp5						-0.497		0.486											
						0.0355		**0.04**											
Gp6																			
Gp7																			
Gp9																			
Gp10																		-0.469	
																		0.048	
Gp13																			
Gp17							-0.639												
							0.004												
Gp18																			
Gp21																0.625			
																0.005			
Gp22																			
Gp23																			
Gp25																			
Gp26																			
Others				-0.5		0.707						0.562							1
				0.04		**0.0008**						**0.015**							**0.0000002**

Spearman correlation coefficient; Highlighted: value of significance P ≤ 0.05

Appendix B. Supplementary Material of Chapter 3

Table S1 Similarity analysis test (ANOSIM) between soil samples collected in 2008 and 2009

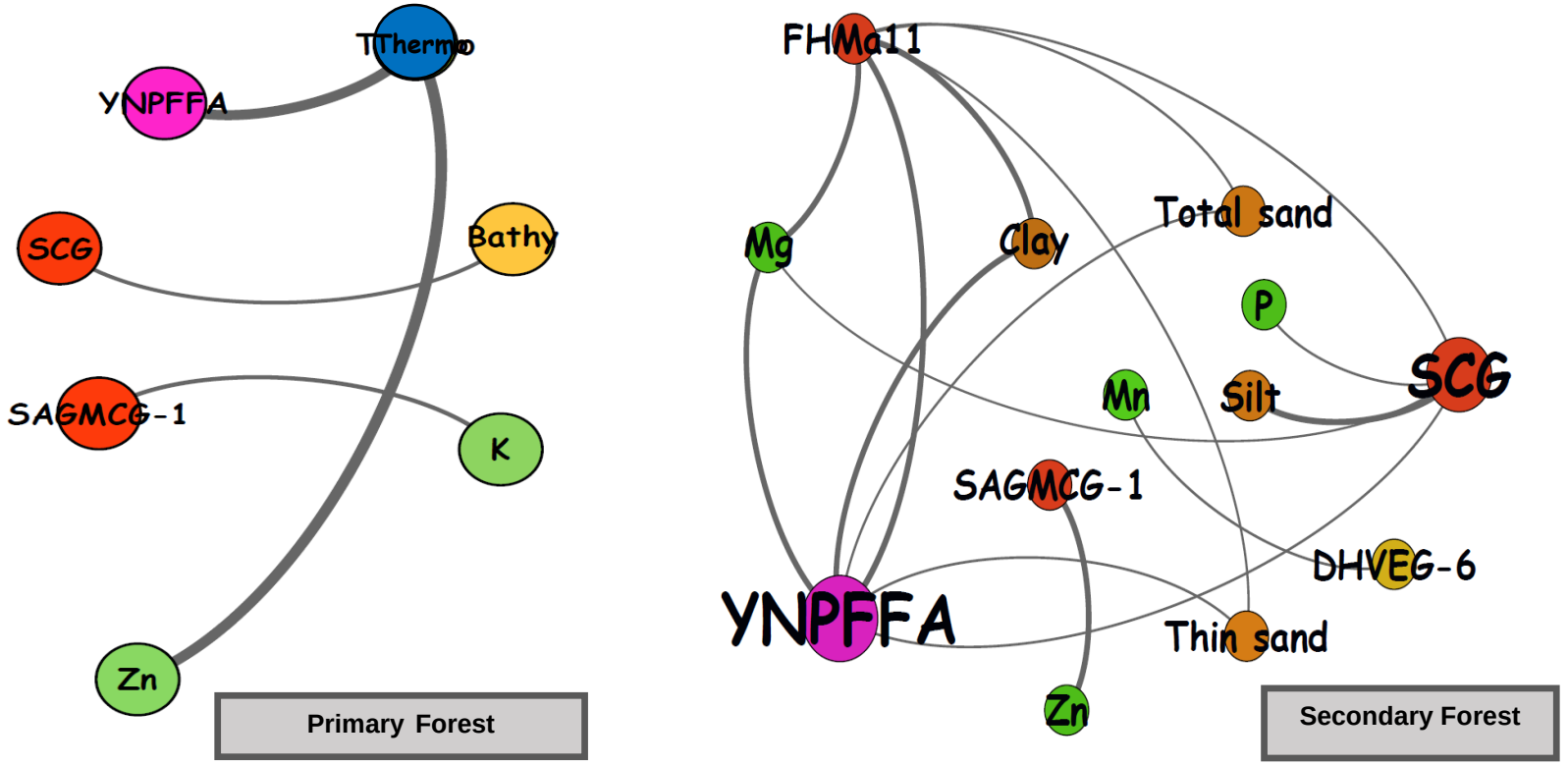
Site	Archaea Groups		Soil Factors	
	R	P	R	P
Primary Forest	0,1111	0,3908	0,2222	0,2008
Secondary Forest	-0,333	0,8107	-0,1852	0,7971
Agriculture	0,4815	0,1052	-0,1852	1
Pasture	-0,1852	1	-0,2222	0,7012

Table S2 Soil physicochemical factors of the 0-20 cm topsoil layer under different land use in the Amazon region - Brazil (PDF)

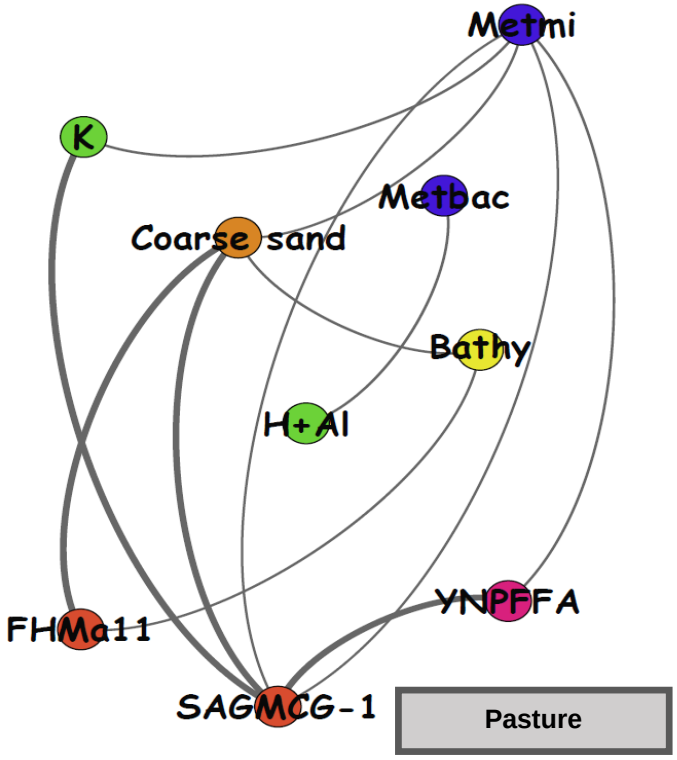
Site	pH	MO	P	Fe	Mn	Zn	K	Mg	H+Al	SB	Clay	Silt	Total sand	Coarse sand	Thin sand
	$CaCl_2$	$g.dm^{-3}$		$mg.dm^{-3}$				$mmolc.dm^{-3}$				g/Kg			
Primary Forest	4,0±0,3	30,6±11	10,3±6	126,3±83	16,4±19	1,1±0,5	1,3±0,3	6,6±3,7	126,5±64	38,6±28	330,5±66	498,0±32	171,5±48	11,6± 10	159,8±47
Secondary Forest	4,2±0,5	30,5±14	10,3±3	89,5±31	21,1±22	2,1±2,1	1,4±0,4	9,0±5,0	108,3±93	56,1±33	**358,8*±114**	446,1±48	195,0±104	11,6±4	183,3±106
Agriculture	4,6±0,6	32,5±15	9,3±4,5	64,8±41	30,7±40	1,3±0,5	1,9±1,2	16,5±15	76,3±77	75,4±60	240,3±54	474,6±178	285,0±216	20,0±14	265,0±216
Pasture	3,9±0,1	29,3±9	6,3±3,2	106,8±59	14,5±19	1,0±0,8	1,8±1,6	5,1±4,2	123,6±28	16,0±8	264,0±70	372,6±115	363,3±176	6,6±8	356,6±183

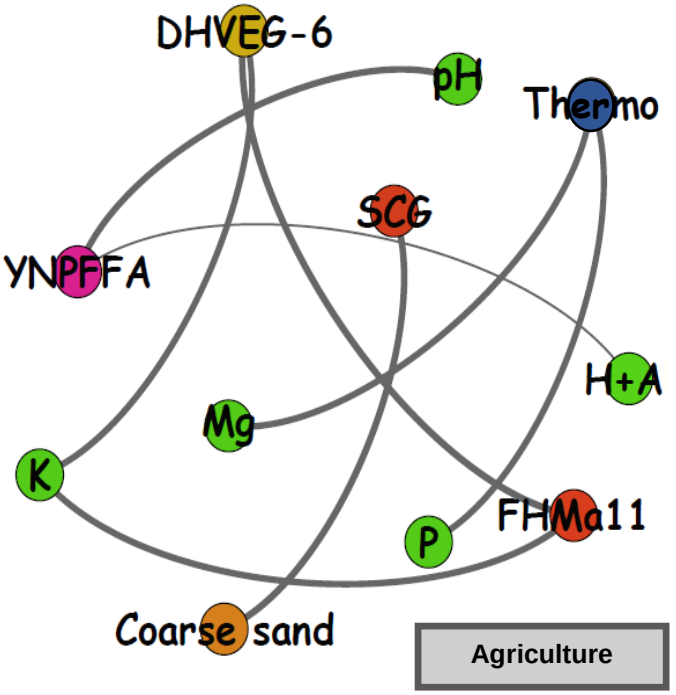
Mean and Standard Deviation based in samples replicates in each site. *** p < 0,05 (Tukey LSD)**

Fig S1. Network correlation analysis between Archaea classes and chemical and physical soil properties, according to significant values ($p < 0.05$) in *Spearman*'s rank correlation coefficient per site. Nodes represent soil chemical factors (green); physical factors (brown); Thaumarchaeota classes (red): FHMa11 – FHMa11 terrestrial group, SAGMCG-1 – South Africa Golden Mine Group 1, SCG – Soil Crenarchaeota group; Euryarchaeota classes: Metbac – Methanobacteria, Metmi – Methanomicrobia and Thermo –Thermoplasmata (blue); Woesearchaeota – DHVEG – 6 (pumpkin); Bathyarchaeota – Bathy (yellow); Crenarchaeota candidate division ~ YNPFFA (purple). Thick edges represent positive correlations, and thin edges represent negative correlations. The large circle represents the largest values of the betweenness centrality statistic factor.



Continuing. Fig S1. Network correlation analysis between Archaea classes and chemical and physical soil properties, according to significant values ($p < 0.05$) in *Spearman*'s rank correlation coefficient per site. Nodes represent soil chemical factors (green); physical factors (brown); Thaumarchaeota classes (red): FHMa11 – FHMa11 terrestrial group, SAGMCG-1 – South Africa Golden Mine Group 1, SCG – Soil Crenarchaeota group; Euryarchaeota classes: Metbac – Methanobacteria, Metmi – Methanomicrobia and Thermo –Thermoplasmata (blue); Woesearchaeota – DHVEG – 6 (pumpkin); Bathyarchaeota – Bathy (yellow); Crenarchaeota candidate division ~ YNPFFA (purple). Thick edges represent positive correlations, and thin edges represent negative correlations. The large circle represents the largest values of the betweenness centrality statistic factor.





Appendix C. Scientific and academic contributions

Obs: direct (marked with an asterisk) or indirect consequences of the work developed during this thesis.

1. Scientific papers

1.1. *__M.G. de Chaves__, G.G. Silva, R. Rossetto, R.A. Edwards, S.M. Tsai, A.A. Navarrete. Acidobacteria subgroups and their metabolic potential for carbon degradation in sugarcane soil amended with vinasse and nitrogen fertilizers. Frontiers in Microbiology**,** Lausanne, v. 10, p. 1680, 2019. https://doi.org/10.3389/fmicb.2019.01680.

1.2. Mendes, L.W.; **de Chaves, M.G.**; Fonseca, M.C.; Mendes, R.; Raaijmakers, J.M.; Tsai, S.M. Resistance Breeding of Common Bean Shapes the Physiology of the Rhizosphere Microbiome. Frontiers in Microbiology, Lausanne, v.10 p. 1-10. 2019 https://doi.org/10.3389/fmicb.2019.02252.

1.3. Eduardo Mariano, Beatriz N. Boschiero, Luis F. Merlotti, Alexandre Pedrinho, **Miriam G. Chaves**, Siu M. Tsai, Paulo C.O. Trivelin. Nitrogen rate and nitrification inhibitor DCD affecting the fate of ^{15}N fertilizer and microbial functional genes in sugarcane soil: a laboratory study. Manuscript Submitted - APSOIL_2020_286.

1.4. *__Miriam Gonçalves de Chaves__; Andressa Monteiro Venturini; Luis Fernando Merloti; Dayane Juliate Barros; Raffaella Rossetto; Siu Mui Tsai; Acacio Aparecido Navarrete. Responses of Archaea and Bacteria N-cycling genes and functional processes in sugarcane soils under vinasse and nitrogen fertilization as indicated by metagenome and GeoChip v.5.0M. Manuscript Submitted – Frontiers in Microbiotecnhology_Agosto de 2020.

2. Book Chapter

Pereira, A. P. A.; de Souza, A.J.; **de Chaves, M.G.**; Fracetto, G. G. M.; Garcia, K. G.V., Filho, P. F.M., Cardoso, E. J. B. N. Mechanisms of the phytomicrobiome for enhancing soil fertility and health In: New and Future Developments in Microbial Biotechnology and Bioengineering - Phytomicrobiome for Sustainable Agriculture. 1 ed. Amsterdam: Elsevier, 2020. v.1, 462 p. https://www.elsevier.com/books/new-and-future-developments-in-microbial-biotechnology-andbioengineering/verma/978-0-444-64325-4

3. Co-orientation

Samela Ketelyn Moreira Oliveira. Impactos da mudança do uso do solo na Amazônia nos gases de efeito estufa CO_2, N_2O E CH_4". Pré - iniciação científica – Laboratório de Biologia Celular e Molecular/ CENA – USP. 2017.

4. Class in short course

de Chaves, M.G.; Paula, F. S. Introdução às Técnicas Moleculares Aplicadas em Biotecnologia e Bioinformática, CENA-USP. 2019.

5. Award

Prêmio Profa. Dra. Leda Mendonça-Hagler - melhor trabalho na área Microbiologia Ambiental, 30º Congresso Brasileiro de Microbiologia - Sociedade Brasileira de Microbiologia – 2019. Souza, L. F.; Obregon, D.; **Chaves, M. G.;** Paula, F. S.; Kroeger, M.; Moraes, M. T.; Moreira, M. Z.; Nüsslein, K.; Tsai, S. M. Soil acidity correction as a modulator of methane cycling microorganisms in the amazon forest-to-pasture conversion.

6. Poster presentation

6.1. *de Chaves, M. G.; Navarrete, A.; Tsai, S. M. Archaeal, bacterial and fungal responses to vinasse and nitrogen application to the sugarcane soil, 2017. VIII Simpósio de Microbiologia Aplicada; Inst.promotora/financiadora: Instituto de Biociências da Universidade Estadual Paulista "Júlio de Mesquita Filho". UNESP – Rio Claro.

6.2. *de Chaves M.G.; Silva G.G.Z.; Rossetto R.; Edwards R.A.; Tsai S.M.; Navarrete A.A. Metabolic potential of Acidobacteria for carbon degradation in soils cultivated with sugarcane fertilized with nitrogen and vinasse. 15° Symposium on Bacterial Genetics and Ecology – BAGECO, Lisboa – Portugal, 2019.

6.3. Souza, L. F.; Obregon, D.; **de Chaves, M. G.;** Paula, F. S.; Kroeger, M.; Moraes, M. T.; Moreira, M. Z.; Nüsslein, K.; Tsai, S. M. Soil acidity correction as a modulator of methane cycling microorganisms in the amazon forest-to-pasture conversion. 30º Congresso Brasileiro de Microbiologia, Maceió – Brasil, 2019.

6.4. *de Chaves, M.G., Merloti, L.F. , de Souza, L. F. , Tsai, S. M., Navarrete, A.A. Ecological co-occurrence and soil physicochemical factors drivers archaeal community in amazon soils under different land use systems. 30º Congresso Brasileiro de Microbiologia, Maceió – Brasil, 2019.

6.5. de Chaves, M.G.; Carrer, H.; Borges, A. Plataforma Moodle como ferramenta interativa para o ensino de Bioquímica no Ensino Superior In: V Workshop sobre Docência no Ensino Superior, Piracicaba – Brasil, 2017.

6.6. Bonassi, R.C.; **de Chaves M.G.**; Vasquez G.H.; Navarrete, A.A. Aplicação de *Nitrospirillum amazonense* em cana-planta via adubação biológica MICROGEO®. III Congresso Cidades Sustentáveis do Noroeste Paulista. Fernandópolis/SP – Brasil, 2019.

7. Monitoring in undergraduate discipline

7.1. "Bioquímica". Obrigatória. Local: Esalq-USP. Professora responsável: Dra. Helaine Carrer. Turmas: Engenharia Agronômica e Engenharia Florestal. 2017. Bolsista PAE.

7.2. "Biodiversidade e Conservação: um enfoque molecular". Optativa. Local: CENA-USP. 2018. Turmas: Ciências Biológicas e Engenharia Agronômica. Professora responsável: Dra. Tsai Siu Mui.