TEOTONIO SOARES DE CARVALHO

LAND-USE CHANGE AND SOIL BACTERIAL COMMUNITIES IN THE EASTERN AMAZON: IMPLICATIONS FOR BIODIVERSITY CONSERVATION AND ECOSYSTEM FUNCTION

LAVRAS – MG

2015

TEOTONIO SOARES DE CARVALHO

LAND-USE CHANGE AND SOIL BACTERIAL COMMUNITIES IN THE EASTERN AMAZON: IMPLICATIONS FOR BIODIVERSITY CONSERVATION AND ECOSYSTEM FUNCTION

Thesis submitted for the degree of Doctor of Philosophy as a Dual Ph.D. with the Departamento de Ciência do Solo, Universidade Federal de Lavras, Brazil, and the Lancaster Environment Centre, Lancaster University, United Kingdom.

Supervisors

Dr. Fatima Maria de Souza Moreira (Universidade Federal de Lavras)

Dr. Jos Barlow (Lancaster University)

Co-supervisors

Dr. James M Tiedje (Michigan State University)

Dr. Ederson da Conceição Jesus (Embrapa Agrobiologia)

LAVRAS – MG

2015

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

de Carvalho, Teotonio Soares.

Land-use change and soil bacterial communities in the Eastern Amazon: implications for biodiversity conservation and ecosystem function / Teotonio Soares de Carvalho. – Lavras : UFLA, 2015.

130 p.

Tese(doutorado)–Universidade Federal de Lavras, 2015.

Orientador(a): Fatima Maria de Souza Moreira.

Bibliografia.

1. Soil biodiversity. 2. Bacterial communities. 3. Land-use intensification. 4. Amazon. 5. Rainforest. I. Universidade Federal de Lavras. II. Título.

TEOTONIO SOARES DE CARVALHO

LAND-USE CHANGE AND SOIL BACTERIAL COMMUNITIES IN THE EASTERN AMAZON: IMPLICATIONS FOR BIODIVERSITY CONSERVATION AND ECOSYSTEM FUNCTION

Thesis submitted for the degree of Doctor of Philosophy as a Dual Ph.D. with the Departamento de Ciência do Solo, Universidade Federal de Lavras, Brazil, and the Lancaster Environment Centre, Lancaster University, United Kingdom.

APPROVED in April 20, 2015

Dr. James M Tiedje	DPSMS/MSU
Dr. Richard D Bardgett	FLC/MU
Dr. Nicholas J Ostle	LEC/LU
Dr. Siu Mui Tsai	CENA/USP
Dr. Patricia Gomes Cardoso	DBI/UFLA

Dr. Fatima Maria de Souza Moreira (Universidade Federal de Lavras)

Supervisor

LAVRAS – MG

2015

ACKNOWLEDGMENTS

At first, I avoided to write this section because I think it is almost impossible to find the right words to honor those that are part of this work. But I (foolishly) changed my mind, so let me try it.
I thank:

The God I am learning to recognize in *every*one and within myself.

My wife, Caroline, for doing this PhD with me. The many challenges that this PhD brought with it were overcome by both of us, together.

My lovely daughter, Lana, for always reminding of smiling.

My parents, Milton and Josecília, who taught me responsibility and respect.

Professor Fatima MS Moreira, who believed in me since the beginning and always supported me to dedicate myself to one of the things that I most love to do, learning.

I am glad to have had the opportunity of meeting Professor James M Tiedje and knowing his inspiring way of doing science and dealing with those starting in the sciences, like me.

Professor Richard Bardgett, who made me even more fascinated with soil ecology through his books and through the interesting discussions we had while I joined his group.

Professor Jos Barlow for his support in Lancaster.

All the Amazonia Sustentável Network team, especially Toby Gardner and Joice Ferreira.

Isaac Carvalho Soares, my brother, who helped in most of the lab activities.

I thank the society from where the resources for my research came from. I thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais for providing the scholarship and all necessary resources while I was in Lancaster, UK. I thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the scholarship and the Programa de Doutorado Sanduíche no Exterior – PDSE/Capes for providing the scholarship and all necessary resources while I conducted my research in Michigan, USA. I also thank the Conselho Nacional de Desenvolvimento Científico (CNPq, Processo 477786 2011-1) for providing the resources for most of the analysis performed in this work.

The Amazonian bacteria that kindly donated their DNA for these analyses.

All the colleagues in Lancaster, especially my two friends Hannah Griffiths and Tom Walker who received me so well and made my time in Lancaster memorable for the rest of my life.

All colleagues in the Michigan State University and to the Ribosomal Database Project (RDP) team for all the interesting conversations.

All my colleagues in Universidade Federal de Lavras for the very nice environment where I have been throughout these four years.

The lab managers, who were always ready to help me. Helen Quirk, Maryam Thompson, Marlene Aparecida, and Manoel Aparecido.

All my teachers, especially Tia Arlete, Tia Rivanda, Mário Sobral, and Carlos André.

The Escolinha da Tia Arlete, Colégio São Francisco, Escola Estadual Dr. Antônio Garcia Filho, Colégio Magistral, and Universidade Federal de Sergipe.

The Universidade Federal de Lavras and the Departamento de Ciência do Solo.

Lancaster University and the Lancaster Environment Centre.

Michigan State University and the Department of Plant, Soil, and Microbial Sciences.

Finally I am thankful to all my friends and my family.

GENERAL ABSTRACT

Tropical rainforests are the largest global reservoir of biodiversity and regulate global climate and biogeochemical cycles. Soil microorganisms are a key component in the functioning of these environments, but the knowledge of how they are affected by land use intensification and forest degradation is still very limited. In the Brazilian Amazon, land use intensification considerably changes chemical properties, especially acidity, which are known to affect soil bacteria. In Paper 01 I studied how soil bacteria community structure changes along a comprehensive gradient of land use intensity (mechanized agriculture, pasture, secondary forest, disturbed primary forest, and undisturbed primary forest) in the eastern Amazon using next-generation sequencing. I predicted that alpha diversity would be higher in more intensive land uses, but that beta diversity would be reduced, indicating a homogenization of the soil bacterial community. However, that was only supported for alpha diversity because, in general, all components of diversity were higher in more intensive land uses. Soil pH was a strong predictor of alpha diversity, with a linear increase in operational taxonomic units number with increases in soil pH, which was also related to land use intensification. Beta diversity was also positively associated with heterogeneity of the soil acidity, higher in more intensive land uses. Finally, land use intensification imposes considerable shifts in bacterial community structure, which is also related to changes in soil fertility. Land use also affects the predicted functional traits of the community, which is expected to affect ecosystem processes. In Paper 02, I further evaluated whether changes in bacterial community structure associated with land use intensification can alter their stability to drought. For that, I applied a dry/rewet treatment on microcosms taken from four land uses (mechanized agriculture, pasture, secondary forest, and undisturbed primary forest) in the eastern Amazon. I observed that soil bacteria were completely insensitive to a two-month simulated drought, regardless of the land use. Similarly, the process of ammonia oxidation did not change with the dry/rewet treatment. These results suggest that bacteria from these soils are equally adapted to these events in all land uses. Papers 03 and 04 were dedicated to advance some of the methodological limitations that I faced during this work. The first presents a tool to design primers (for polymerase-chain reaction – PCR) for genes of functional relevance, especially developed for gene-targeted metagenomics. In this approach, primers are not necessarily designed from the same position in the alignment and can differ in size, contrary to the frequently used degenerate primers. Furthermore, this tool

takes into account thermodynamic properties of the primers to ensure compatibility among them. I tested this tool on the archaeal ammonia monooxygenase subunit A (*amoA*). Results from *in silico* PCR indicate good coverage (~80%) and high specificity of the designed primers. When tested on DNA extracted from Amazonian soil, these primers allowed the amplification of fragments with the expected length, indicating their suitability for complex samples like soil. In Paper 04, I explore some properties of an index commonly used for studies of stability in soil (as in Paper 02). I describe novel properties of the resilience and resistance indices that were not mentioned in the original or any subsequent work. These properties may cause spurious/erroneous conclusions. I provide recommendations that avoid such problems for both indices.

Keywords: Amazon forest. Below-ground biodiversity. Next-generation sequencing. Ammonia oxidizing organisms. Drivers of bacterial community structure. *16S rRNA gene*.

RESUMO GERAL

As florestas tropicais são as principais reservas mundiais de biodiversidade e regulam o clima e os ciclos biogeoquímicos no planeta. Os microrganismos do solo são componentes-chave para o funcionamento desses ecossistemas, mas o conhecimento de como eles são afetados pela intensificação do uso da terra e pela degradação de florestas ainda é muito limitado. Na Amazônia brasileira, a intensificação do uso da terra altera consideravelmente as propriedades químicas do solo, especialmente a acidez, fatores que conhecidamente afetam as bactérias do solo. No Capítulo 02 eu avaliei como a estrutura das comunidades de bactérias do solo muda ao longo de um gradiente abrangente de intensidade de uso da terra (agricultura mecanizada, pastagem, floresta secundária, floresta primária perturbada, floresta primária não perturbada) na Amazônia oriental usando sequenciamento de nova geração. Eu esperava que a diversidade alfa de bactérias fosse maior em usos da terra mais intensivos, mas que a diversidade beta fosse reduzida, indicando uma homogeneização da comunidade bacteriana do solo. Entretanto, isso foi apenas confirmado para a diversidade alfa porque, em geral, todos os componentes de diversidade foram maiores em sistemas de uso mais intensivos. O pH do solo foi um forte preditor da diversidade alfa, com um aumento linear no número de unidades taxonômicas operacionais com aumentos no pH do solo, que também foi relacionado à intensificação do uso da terra. A diversidade beta também foi positivamente associada à heterogeneidade da acidez do solo, maior em sistemas de uso mais intensivos. Finalmente, a intensificação do uso da terra também impõe mudanças consideráveis na estrutura das comunidades bacterianas, que também se relaciona com mudanças na fertilidade do solo. O uso da terra também afeta os atributos funcionais preditos para a comunidade bacteriana, com possíveis implicações para os processos de ecossistema. No Capítulo 03, eu avaliei se as mudanças na estrutura da comunidade bacteriana associada à intensificação do uso da terra pode alterar a estabilidade delas frente a secas. Para isso, eu apliquei um tratamento de secagem/reidratação em microcosmos originários de quatro usos da terra (agricultura mecanizada, pastagem, floresta secundária e floresta primária) na Amazônia oriental. Eu observei que as bactérias do solo foram completamente insensíveis a dois meses de secagem, independentemente do uso da terra. Da mesma forma, o processo de oxidação de amônia não mudou após o tratamento de secagem/reidratação. Esses resultados sugerem que as bactérias desses solos são igualmente adaptadas a esses eventos em todos os sistemas de uso. Os Capítulos 04 e 05 foram dedicados a avançar algumas das limitações metodológicas que enfrentei durante este trabalho. O

primeiro apresenta uma ferramenta para o desenho de iniciadores (para reação em cadeia da polimerase) para genes de relevância funcional, especialmente desenvolvida para metagenômica direcionada a genes específicos. Nesta abordagem, os iniciadores não são necessariamente desenhados a partir de uma mesma posição no alinhamento e podem diferir em tamanho, ao contrário dos iniciadores degenerados frequentemente utilizados. Além disso, esta ferramenta leva em consideração propriedades termodinâmicas dos iniciadores para garantir que eles são compatíveis entre si. Eu testei esta ferramenta na subunidade A do gene da amônia monooxigenase (*amoA*) de arqueobactérias. Os resultados da PCR *in silico* indicam uma boa cobertura (~80%) e alta especificidade dos iniciadores desenhados. Quando testados em DNA extraído de solos da Amazônia, esses iniciadores permitiram a amplificação de fragmentos com o comprimento esperado, indicando a capacidade de eles funcionarem em amostras complexas como o solo. No Capítulo 05, eu explorei algumas propriedades de um índice comumente utilizado para estudos de estabilidade no solo (como no Capítulo 03). Eu observei que tanto o índice de resiliência quanto o índice de resistência apresentam propriedades que não foram mencionadas no trabalho original que o descreveu, ou em qualquer dos artigos que os utilizaram. Essas propriedades podem causar conclusões espúrias/errôneas. Eu forneço recomendações que evitam esses problemas para ambos os índices.

Palavras chave: Floresta amazônica. Biodiversidade do solo. Sequenciamento de nova geração. Organismos oxidantes de amônia. Estruturadores das comunidades bacterianas. Gene *16S rRNA*.

SUMMARY

FIRST PART 13

1 GENERAL INTRODUCTION 13

1.1 Biodiversity and ecosystem functioning 13

1.2 Soil biodiversity 14

1.3 Land use intensification and soil biodiversity in the Amazon region 15

1.4 Ammonia-oxidizing microorganisms (AOM) 16

1.5 Next-generation sequencing technologies applied to microbial ecology 18

1.6 Objectives 20

2 CONCLUDING REMARKS 20

2.1 Impact of land use intensification on bacterial communities in Amazon 21

2.2 Ability of bacteria from Amazonian soil to cope with a dry/rewet event 22

2.3 Methodological advances 22

2.4 Future research 24

2.5 Conclusions 25

REFERENCES 26

SECOND PART - PAPERS 33

PAPER 1 Soil acidity controls both alpha and beta diversity of soil bacteria in human-modified tropical forest landscapes 33

PAPER 2 Soil bacterial communities are highly stable to a dry/rewet event in the eastern Amazon 65

PAPER 3 A new approach for designing primers for functional genes with emphasis on gene-targeted metagenomics 85

PAPER 4 Pitfalls in a commonly used index for measuring resistance and resilience of ecosystem processes in soil 107

APPENDIX 127

FIRST PART

1 GENERAL INTRODUCTION

1.1 Biodiversity and ecosystem functioning

Biodiversity is defined as the variability among living organisms, including the diversity within and among species, as well as the diversity of ecosystems (United Nations, 1992). Its conservation is considered fundamental for the maintenance of the ecosystem functions and services, which are estimated, in a global scale, at about $145 trillion/yr (Costanza et al., 2014). Besides its functional role, the importance of biodiversity also results from aesthetics and cultural aspects, as well as its intrinsic value (Moonen e Barberi, 2008).

In spite of its importance, biodiversity has declined across the whole world as a result of human activity, with many changes being driven by the conversion of natural environments into agricultural lands, and subsequent intensification of land use (Millennium Ecosystem Assessment, 2005). This loss of biodiversity may have implications on ecosystem functioning and stability as indicated by theoretical models (Naeem, 1998; Yachi and Loreau, 1999), and experimental evidence (Naeem and Li, 1997; Griffiths et al., 2000; Loreau et al., 2001; Maherali e Klironomos, 2007).

However, given the conditions in which those studies were conducted, their results can not be directly extrapolated to real ecosystems because the manipulation of diversity in artificial conditions does not reflect the way communities are naturally assembled or disassembled, especially for soil (Bardgett, 2005). Another limitation of most experiments studying the relationship between biological diversity and soil ecosystem functioning is that

the number of species manipulated are usually orders of magnitude lower than the species richness usually found in soil (Nielsen et al., 2011).

1.2 Soil biodiversity

Despite soil's huge biodiversity and its importance for terrestrial ecosystems, as soil organisms are key drivers of biogeochemical cycles (Falkowski et al., 2008), the soil ecology and the functional importance of below-ground organisms remained poorly considered until recently. For that reason, soil ecology lags behind the ecology of above ground organisms (Wall et al., 2005; Maron et al., 2011; Nielsen et al., 2011). This is especially true for tropical regions because most studies of the functional importance of biodiversity take place in temperate climates and focus on intensive agriculture (Giller et al., 2005).

Soil ecology has some peculiarities that distinguish it from the ecology of above-ground organisms. For instance, the enormous diversity of organisms that coexist in the soil at very small spatial scales seems to violate the principle of competitive exclusion, what is known as "the enigma of soil biodiversity" (Anderson, 1975; Bardgett et al., 2005). It is believed that this phenomenon is a consequence of some characteristics of the soil habitat, especially the large spatial heterogeneity even at very small scales, associated with the high degree of specialization of the soil biota (Nielsen et al., 2010).

Because of its high biodiversity, it is believed that soil communities have a high degree of functional redundancy, i.e., many species per functional group (Setälä et al., 2005). Ecosystems with high functional redundancy are hypothetically more reliable, in the sense of being more stable to perturbation and to species losses (Walker, 1992; Naeem, 1998). However, this view of high

functional diversity in soil has to be taken carefully because some processes in soil are possibly mediated by restricted groups of organisms, such as ammonia oxidizers and organisms that decompose recalcitrant molecules in soil. These processes may be more susceptible to instability due to species losses (Schimel et al., 2005).

Even under high functional redundancy, soil processes are not only affected by species richness, but also by community structure of soil organisms (Griffths et al., 2004; Heemsbergen et al., 2004; Maherali & Klironomos, 2007), in such a way that the loss of key species or functional traits in the community may cause instability in the system even when the species richness remains high in that functional group. That is one believed reason for the idiosyncratic relationship between species richness and soil processes usually observed in soil (Nielsen et al., 2011).

1.3 Land use intensification and soil biodiversity in the Amazon region

Tropical rainforests are a great reservoir of biodiversity and participate in the regulation of global climate and biogeochemical processes (Orians et al., 1996). However, these systems are threatened by human activities, and the losses resulting from its economic exploitation in terms of biodiversity and ecosystem functioning stability are poorly understood, especially for soil biota (Moreira et al., 2008).

Land use intensification in the Amazon region homogenizes plant communities (Arroyo-Rodríguez, 2013). Consequently, the diversity of organic substances in the rhizosphere and in the litter, the diversity of soil microhabitats, and the diversity of plant hosts for symbiotic microorganisms may also be reduced with land use intensification (Wardle, 2006). Furthermore, slash-and-

burn practices and the application of fertilizer and lime considerably change soil fertility and acidity (Moreira et al., 2009), which are known drivers of bacterial community structure (Fierer et al., 2006; Jesus et al., 2009). Finally, conversion of tropical forests also has strong impacts on soil physical properties, increasing the susceptibility to compaction (Martins et al., 2012), reducing water infiltration rate in soil (Lal et al. 1979; Lal 1989) and increasing soil erosion (Maeda et al., 2008).

The impact of land use intensification on soil biodiversity is not yet well understood, but it usually reduces the diversity of soil arthropods (Lo-Man-Hung et al., 2011; Braga et al., 2013; Solar 2014). Soil microorganisms seem to respond differently, as arbuscular mycorrhizal fungi (Picone, 2000; Stürmer and Siqueira, 2011) and soil bacteria (Jesus et al., 2009; Rodrigues et al., 2013; Mendes et al., 2015) both tend to have a higher local (alpha) diversity in human-modified environments than in pristine forests.

On the other hand, after conversion of forest to pasture in the western Amazon region, reductions in the bacterial turnover across space (beta-diversity) have been reported (Rodrigues et al., 2013), indicating that total diversity of soil bacteria may be reduced by land use intensification, despite the higher local diversity. This biotic homogenization of soil bacteria may compromise ecosystem stability (Naeem and Li, 1997), especially when larger spatial and temporal scales are considered (Hooper et al., 2005).

1.4 Ammonia-oxidizing microorganisms (AOM)

Ammonia oxidation was proposed as a model process for molecular studies of microbial ecology (Kowalchuk and Stephen, 2001). This process, in which ammonia is converted to nitrite by non-obligate autotrophic

microorganisms (Arp et al., 2007), is the limiting step for nitrification, i.e., conversion of ammonia to nitrate.

Ammonia oxidizers form a very restricted group of organisms, although new insights suggest this may not be as restricted as it was previously thought (Schimel et al., 2005). Until recently, it was believed that ammonia oxidation was only performed by two lineages of aerobic bacteria called ammonia-oxidizer bacteria (AOB). The first lineage belongs to the betaproteobacteria (beta-AOB) and comprises *Nitrosomonas* (including *Nitrosococcus mobilis*) and *Nitrosospira* (including *Nitrosolobus* and *Nitrosovibrio*) species. The second lineage, affiliated with the gammaproteobacteria (gamma-AOB), contains *Nitrosococcus oceani* and *Nitrosococcus halophilus* (Junier et al., 2010).

The recent discovery of ammonia oxidizer Archaea (AOA) in sea water (Venter et al., 2004) and in soil (Treusch et al., 2005) through metagenomic studies, also confirmed by later cultivation of these organisms (Könneke et al., 2005; de la Torre et al., 2008), had a great impact on the study of nitrification. Further studies showed that AOA can not only be found in soil, but can also predominate over AOB in abundance (Leininger et al., 2006) and that they can be more important than AOB to the nitrification process in agricultural soils (Gubry-Rangin et al., 2010), though the opposite may also occur (Jia e Conrad, 2009; Glaser et al., 2010).

These two groups of ammonia oxidizers (AOA and AOB) occupy distinct niches regarding ammonia oxidation due to the differentiated kinetics of the ammonia monooxigenase present in each of them. In AOB, this enzyme is more efficient at high ammonia concentrations, whereas in AOA it is more efficient at low ammonia concentrations (Martens-Habbena et al., 2009). Besides, soil pH strongly influences the dynamic of AOA and AOB communities by favoring AOA at acidic conditions (Nicol et al., 2008). The volume of soil pores occupied by water also affects these two groups, especially

AOB (Gleeson et al., 2010). Other factors such as soil type and land use (Colloff et al., 2008; Hayden et al., 2010; Taketani and Tsai, 2010), and fertilization regime (He et al., 2007; Kong et al., 2010) also influences the AOM dynamics. Their responsiveness to environmental conditions and the importance of ammonia oxidation in the nitrogen cycle make this functional group an interesting model for the impact of anthropogenic pressures on soil processes mediated by soil bacteria.

1.5 Next-generation sequencing technologies applied to microbial ecology

The development of molecular techniques revolutionized microbial ecology by considerably reducing the time and cost of microbial communities assessments when compared to traditional culture-based methods. It also allowed for a more comprehensive analysis of microbial communities by giving access to uncultured organisms, which comprise the majority (99%) of all microorganisms (Xu, 2006).

In the last seven years, microbial ecology has benefited from the development of next-generation sequencing (NGS). This technology allows large-scale sequencing at high speed and can generate billions of reads at costs about 10,000x smaller than that of first-generation sequencing. The application of NGS for studying microbial communities is usually done by whole genome shotgun (WGS) sequencing (Thomas et al., 2012) or by gene-targeted metagenomics (Iwai et al., 2011). The first method is based on sequencing of entire metagenomes from multiples organisms simultaneously. The second approach is based on sequencing amplicons from polymerase chain reaction (PCR).

The advantages of WGS include its suitability for the discovery of novel organisms, genes and pathways (Treusch et al., 2005; Culligan et al., 2013); it allows a detailed functional profile of microbial communities (Langille et al., 2013); it can be used for evaluating differential gene expressions by combining RNA and DNA sequencing (Dupont et al., 2014), among others (Myrold et al., 2014). However, the application of this technique for soil microorganisms is still very limited because the least abundant microorganisms, including many important functional groups (e.g. ammonia oxidizers, nitrogen-fixing bacteria), are not well covered. An alternative to circumvent this limitation is to increase the depth of sequencing, but current limitations in data storage and data processing are still prohibitive for large sample collections (Langille et al., 2013), although novel methods for metagenomics data processing are under development to reduce those limitations (Howe et al., 2014). Another difficulty for applying WGS for soil microorganism is the limited availability of reference genomes or genes for assembling the short reads generated by NGS and to annotate the assembled reads (Myrold et al., 2014).

In gene-targeted metagenomics, the use of PCR coupled with barcoding allows a detailed analysis of specific genes for hundreds of samples simultaneously, considerably reducing the sequencing costs and, consequently, making this technique especially suitable for large sample collections (Iwai et al., 2011). Because of that, it has been widely used on hypervariable regions of ribosomal genes to taxonomically characterize microbial communities (Caporaso et al., 2012; Wang et al., 2007). It has also been used on specific functional genes, allowing the assessment of the diversity within functional groups of the soil biota (Iwai et al., 2009; Wang et al., 2013). However, designing primers with good coverage for functional genes is fundamental to minimize the bias of this approach (Penton et al., 2013), but this remains difficult due to limitations of current available tools for primer design.

1.6 Objectives

The main objective of this thesis was to investigate changes in the diversity and community structure of soil bacteria along a gradient of land use intensity in the eastern Amazon (Paper 01), and the implications of those changes on the ability of these communities to cope with a simulated dry/rewet event (Paper 02). I also aimed to provide two methodological advances for the study of soil ecology. First, I developed a tool for designing primers in gene-targeted metagenomics (Paper 03). Second, I undertook a detailed analysis of the circumstances under which two commonly used indices for resilience/resistance in soil may lead to spurious/erroneous conclusions (Paper 04).

2 CONCLUDING REMARKS

This thesis addressed the impact of land use intensification on bacterial community structure, activity, and stability in the eastern Amazon. It builds on previous work in the Amazon (Jesus et al., 2009; Rodrigues et al., 2013; Mendes et al., 2015) by including a comprehensive gradient of land use intensity in a region that has high anthropogenic pressure on the remaining Amazon forest. It also provides two novel methodological advances for the study of soil ecology, namely, a tool for primer design developed for gene-targeted metagenomics and a detailed description of some pitfalls in two commonly used indexes for measuring stability of soil processes.

2.1 Impact of land use intensification on bacterial communities in Amazon

Increases in local alpha diversity for bacterial communities have been frequently found after conversion of Amazon rainforest to human-modified environments (Jesus et al., 2009; Rodrigues et al., 2013; Mendes et al., 2015). I found similar results for a comprehensive gradient of land use in the eastern Amazon. In addition, I present evidence that reduction in soil acidity in more intensive land uses is the main driver for this frequently observed increase. This is further supported by previous observations of the strong effect of soil pH on bacterial richness in other ecosystems (Fierer and Jackson 2006; Lauber et al., 2009; Rousk et al., 2010).

The sampling scheme used in this study allowed for the assessment of all components of diversity, with at least four replicates for all five land uses. I observed that, at the spatial scale of this study, not only alpha, but also beta and gamma diversities increased in more intensive land uses. This higher heterogeneity (beta diversity) of the bacterial community in more intensive land uses has not yet been reported in the Amazon region and it challenges the hypothesis of homogenization of soil bacterial communities with land use intensification (Rodrigues et al., 2013). I also found that soil pH heterogeneity, higher in more intensive land uses, predicts bacterial beta diversity. Accordingly, changes in soil fertility, especially soil acidity, are closely associated with shifts in bacterial community structure imposed by land use intensification. These changes also altered the predicted functional traits of the bacterial community, with possible implications for ecosystem processes in soil.

2.2 Ability of bacteria from Amazonian soil to cope with a dry/rewet event

Based on known differences among bacteria to cope with drought (Schimel et al., 2007; Placella et al., 2012), I predicted that shifts in bacterial community structure associated with land use intensification alters the stability of soil bacterial communities, and the processes they mediate, to a dry/rewet event. However, these predictions were not supported by the observation that, regardless of the land use, bacterial communities were completely stable to the artificial dry/rewet event used in this study. Similarly, the process of ammonia oxidation was not changed by this event, indicating that this bacteria-mediated process is also highly stable to the experimental treatment.

To the best of my knowledge, this is the first study testing the stability of bacterial community structure to a dry/rewet event in Amazonian soils. It is possible that the observed stability is caused by previous selection of drought tolerant bacteria or the induction of changes in bacterial strategies to cope with dry/rewet events due to previous exposure to drought (Evans and Wallenstein, 2012, 2014; Göransson et al., 2013). However, it is not known to what extent the stability observed here in a two-month simulated drought will continue under the longer dry seasons that occur with increasing intensity of El Niño–Southern Oscillation (ENSO) and Atlantic sea surface temperature oscillations (Chen et al. 2011).

2.3 Methodological advances

Primer Design

In Paper 03, I studied the effect of land use intensity on bacterial community structure using the 16S *rRNA* gene. Another interesting approach is to sequence genes of ecological relevance (e.g. *amoA*, *nosZ*, *nifH*), which can give insights on how land use intensification impacts specific functional groups. But primer design for these genes is still a bottleneck for the application of gene-targeted metagenomics for functional genes.

I developed a thermodynamic-based tool to design primers using sequence data available in online repositories like Fungene (Fish et al., 2013). I used this tool to design primers for the archaeal *amoA* gene, specifically for the *Nitrosophaera* group. These primers were predicted to have high coverage, high specificity, and good thermodynamic properties (e.g. absence of hairpins, primer-primer interactions, differences in melting temperature) based on *in-silico* PCR. When tested in DNA extracted from Amazonian soils, the primers allowed the amplification of fragments with the predicted length.

Stability Indexes

During the analysis of the data collected for Paper 02, I realized that a commonly used index for soil resilience and resistance (Orwin and Wardle, 2004) have mathematical properties that can lead to spurious or erroneous conclusions in naturally stochastic system processes. Based on this observation, I provided computer code to calculate the predicted value of the resistance index under the null model of no disturbance effect. These predictions can be used to assess whether or not values of this index applied on real data can be explained by natural variability.

Because these aspects are not mentioned in the original or any subsequent work, I believe that the Paper 04 makes an important contribution to the understanding of these indexes, showing conditions where it should not be

used, and what exactly they express. This will certainly increase the accuracy of conclusions based on this index.

2.4 Future research

In the studied region, soil bacteria seems to be favored by land use intensification, but assessments of bacterial biomass or quantification of bacterial genes are needed to confirm that. Furthermore, it is possible that the observed increase in bacterial diversity occurs to the detriment of fungal diversity. In addition to analysis of fungal diversity and community structure, analysis of bacteria and fungi biomass, and their ratio, may also give a more comprehensive picture of the impacts of human activities on microbial communities in Amazon.

An important limitation of the approach used here is that biodiversity losses in less abundant functional groups may be masked by the tolerance of the most abundant/generalist groups to land use intensification. Consequently, studies targeting specific groups of soil bacteria may be able to detect hidden biodiversity losses, which are likely to influence ecosystem processes. Finally, analyses focused on microbial genes, rather than on their identity, as in shotgun metagenomics, can also help to clarify the functional implications of the strong shifts in bacterial community structure after rainforest conversion observed in this study.

Experiments manipulating soil pH in the Amazon would help to disentangle the effect of soil acidity on bacterial community structure from other human-induced impacts. Particularly, it would be interesting to investigate whether bacterial communities from agricultural soils would be able to maintain soil functioning after reducing soil pH to levels similar to those under natural

conditions. On the other hand, experiments increasing pH in the forest soil would give interesting insights of whether reduced acidity alleviates the selective pressure on less abundant bacteria and, consequently, increases the (detectable) diversity of soil bacteria.

With regard to the stability of soil ecosystems to environmental changes in Amazon, further studies could test longer and more severe drought treatments, as well as simulated wildfires, because the frequency and intensity of those events are expected to increase in Amazon (Chen et al. 2011). Other components of the soil biota (e.g. fungi, arthropods, nematodes) should also be studied in that respect as they are more likely to be sensitive to those disturbances, especially in human-modified environments.

2.5 Conclusions

In summary, the results presented here indicate that the main impact of land use intensification on soil bacteria in the eastern Amazon is related to shifts in bacterial community structure, and not to losses in bacterial diversity. These shifts are accompanied by changes in the predicted abundance of functional genes, with possible implications for the soil processes mediated by bacteria. I provided evidences that increased soil fertility and reduced soil acidity are the main mediators of the effects of land use intensification on bacteria community structure and diversity in soils from the eastern Amazon. Finally, bacterial communities and the process of ammonia oxidation in the studied soils are also highly stable to dry/rewet events, with no detectable effect of land use on their stability.

REFERENCES

ANDERSON, J. M. The enigma of soil animal species diversity. In: VANĚK, J. (Ed.). Progress in soil zoology. Netherlands: Springer, 1975. p.51-58.

ARP, D. J.; CHAIN, P. S. G.; KLOTZ, M. G. The impact of genome analyses on our understanding of ammonia-oxidizing bacteria. **Annual Review of Microbiology,** Palo Alto, v. 61, p. 503-28, January 2007.

ARROYO-RODRÍGUEZ, V. et al. Plant β-diversity in fragmented rain forests: Testing floristic homogenization and differentiation hypotheses. **Journal of Ecology,** London, v. 101, p. 1449-1458, October 2013.

BARDGETT, R. **The Biology of Soil: A community and ecosystem approach**. Oxford Press. Oxford. 2005. 1-256

BRAGA, R. F. et al. Dung beetle community and functions along a habitat-disturbance gradient in the Amazon: a rapid assessment of ecological functions associated to biodiversity. **PloS one,** San Francisco, v. 8, n. 2, p. e57786-e57786, February 2013.

CAPORASO, J. G. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. **The ISME Journal,** New York, v. 6, n. 8, p. 1621-1624, March 2012.

CHEN, Y. et al. Forecasting Fire Season Severity in South America Using Sea Surface Temperature Anomalies. **Science,** Washington, v. 334, n. November, p. 787-791, November 2011.

COLLOFF, M. J. et al. Detection of nitrogen cycle genes in soils for measuring the effects of changes in land use and management. **Soil Biology and Biochemistry,** New York, v. 40, n. 7, p. 1637-1645, July 2008.

COSTANZA, R. et al. Changes in the global value of ecosystem services. **Global Environmental Change,** New York, v. 26, n. 0, p. 152-158, May 2014.

CULLIGAN, E. P. et al. Metagenomics and novel gene discovery. **Virulence,** Oxon, v. 5, n. 3, p. 399-412, April 2013.

DE LA TORRE, J. R. et al. Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. **Environmental Microbiology,** Hoboken, v. 10, n. 3, p. 810-818, January 2008.

DUPONT, C. L. et al. Genomes and gene expression across light and productivity gradients in eastern subtropical Pacific microbial communities. **ISME Journal**, New York, October 2014.

FALKOWSKI, P. G.; FENCHEL, T.; DELONG, E. F. The microbial engines that drive Earth's biogeochemical cycles. **Science**, New York**,** v. 320, n. 5879, p. 1034-1039, May 2008.

FIERER, N.; JACKSON, R. B. The diversity and biogeography of soil bacterial communities. **Proceedings of the National Academy of Sciences of the United States of America,** Washington, v. 103, n. 3, p. 626-31, January 2006.

GILLER, K. E. et al. Soil biodiversity in rapidly changing tropical landscapes: scaling down and scaling up. In: BARDGETT, R.;USHER, M.*, et al* (Ed.). Biological diversity and function in soils. Cambridge: Cambridge University Press, 2005. p.295-318. ISBN 9780511541926.

GLASER, K. et al. Dynamics of ammonia-oxidizing communities in barley-planted bulk soil and rhizosphere following nitrate and ammonium fertilizer amendment. **FEMS Microbiology Ecology,** Chichester, v. 74, n. 3, p. 575-91, October 2010.

GLEESON, D. B. et al. Response of ammonia oxidizing archaea and bacteria to changing water filled pore space. **Soil Biology and Biochemistry,** New York, v. 42, n. 10, p. 1888-1891, October 2010.

GRIFFITHS, B. S. et al. The relationship between microbial community structure and functional stability, tested experimentally in an upland pasture soil. **Microbial Ecology,** Washignton, v. 47, n. 1, p. 104-13, January 2004.

GRIFFITHS, B. S. et al. Ecosystem response of pasture soil communities to fumigation-induced microbial diversity reductions: an examination of the biodiversity–ecosystem function relationship. **Oikos,** Lund, v. 90, n. 2, p. 279-294, April 2000.

GUBRY-RANGIN, C.; NICOL, G. W.; PROSSER, J. I. Archaea rather than bacteria control nitrification in two agricultural acidic soils. **FEMS Microbiology Ecology,** Chichester, v. 74, n. 3, p. 566-74, October 2010.

HAYDEN, H. L. et al. The abundance of nitrogen cycle genes amoA and nifH depends

on land-uses and soil types in South-Eastern Australia. **Soil Biology and Biochemistry,** New York, v. 42, n. 10, p. 1774-1783, October 2010.

HE, J.-Z. et al. Quantitative analyses of the abundance and composition of ammonia-oxidizing bacteria and ammonia-oxidizing archaea of a Chinese upland red soil under long-term fertilization practices. **Environmental Microbiology,** Hoboken, v. 9, n. 9, p. 2364-74, September 2007.

HEEMSBERGEN, D. A. et al. Biodiversity effects on soil processes explained by interspecific functional dissimilarity. **Science,** New York, v. 306, n. 5698, p. 1019-20, November 2004.

HOOPER, D. U. et al. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. **Ecological Monographs,** Ithaca, v. 75, n. 1, p. 3-35, February 2005.

HOWE, A. C. et al. Tackling soil diversity with the assembly of large, complex metagenomes. **Proceedings of the National Academy of Sciences of the United States of America,** Washington, v. 111, n. 13, p. 4904-4909, April 2014.

IWAI, S. et al. Gene-Targeted Metagenomics (GT Metagenomics) to Explore the Extensive Diversity of Genes of Interest in Microbial Communities. In: de Brujin, F. J. (Ed.). Handbook of molecular microbial ecology: Metagenomics and complementary approaches. Hoboken: John Wiley & Sons, Inc., 2011. p.235-243. ISBN 9781118010518.

IWAI, S. et al. Gene-targeted-metagenomics reveals extensive diversity of aromatic dioxygenase genes in the environment. **ISME Journal,** New York, v. 4, n. 2, p. 279-285, February 2009.

JESUS, E. C. et al. Changes in land use alter the structure of bacterial communities in Western Amazon soils. **The ISME Journal,** New York, v. 3, n. 9, p. 1004-1011, September 2009.

JIA, Z.; CONRAD, R. Bacteria rather than Archaea dominate microbial ammonia oxidation in an agricultural soil. **Environmental Microbiology,** Hoboken, v. 11, n. 7, p. 1658-71, July 2009.

JUNIER, P. et al. Phylogenetic and functional marker genes to study ammonia-oxidizing microorganisms (AOM) in the environment. **Applied Microbiology and Biotechnology,** Washington, v. 85, n. 3, p. 425-40, January 2010.

KONG, A. Y. Y. et al. Impacts of different N management regimes on nitrifier and

denitrifier communities and N cycling in soil microenvironments. **Soil Biology and Biochemistry,** New York, v. 42, n. 9, p. 1523-1533, September 2010.

KONNEKE, M. et al. Isolation of an autotrophic ammonia-oxidizing marine archaeon. **Nature,** London, v. 437, n. 7058, p. 543-546, September 2005.

KOWALCHUK, G. A.; STEPHEN, J. R. Ammonia-oxidizing bacteria: a model for molecular microbial ecology. **Annual Review of Microbiology,** Palo Alto, v. 55, p. 485-529, October 2001.

LAL, R. Agroforestry systems and soil surface management of a tropical alfisoh. **Agroforestry Systems,** Washington, v. 8, p. 197-215, March 1989.

LAL, R.; WILSON, G. F.; OKIGBO, B. N. Changes in properties of an Alfisol produced by various crop covers. **Soil Science,** Philadelphia, v. 127, n. 6, p. 377-382, May 1979.

LANGILLE, M. G. I. et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. **Nature Biotechnology,** London, v. 31, n. 9, p. 814-21, August 2013.

LEININGER, S. et al. Archaea predominate among ammonia-oxidizing prokaryotes in soils. **Nature,** London, v. 442, n. 7104, p. 806-9, August 2006.

LO-MAN-HUNG, N. F. et al. Impact of different land management on soil spiders (Arachnida: Araneae) in two Amazonian areas of Brazil and Colombia. **Journal of Arachnology,** Miami, v. 39, n. 2, p. 296-302, August 2011.

LOREAU, M. et al. Biodiversity and ecosystem functioning: current knowledge and future challenges. **Science,** New York, v. 294, n. 5543, p. 804-8, 2001.

MAEDA, E. E.; FORMAGGIO, A. R.; SHIMABUKURO, Y. E. Impacts of Land Use and Land Cover Changes on Sediment Yield in a Brazilian Amazon Drainage Basin. **Giscience & Remote Sensing**, Columbia**,** v. 45, n. 4, p. 443-453, 2008.

MAHERALI, H.; KLIRONOMOS, J. N. Influence of phylogeny on fungal community assembly and ecosystem functioning. **Science,** New York, v. 316, n. 5832, p. 1746-8, October 2007.

MARON, P.-A.; MOUGEL, C.; RANJARD, L. Soil microbial diversity: Methodological strategy, spatial overview and functional interest. **Comptes Rendus Biologies**, [C.l.], February 2011.

MARTENS-HABBENA, W. et al. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. **Nature,** London, v. 461, n. 7266, p. 976-9, September 2009.

MARTINS, P. C. C. et al. Structural sustainability of cambisol under different land use system. **Revista Brasileira de Ciência do Solo,** Viçosa, v. 36, n. 1, p. 1724-1732, November 2012.

MENDES, L. W. et al. Soil-Borne Microbiome: Linking Diversity to Function. **Microbial Ecology**, Washington, July 2015.

MILLENIUM ECOSYSTEM ASSESSMENT. **Ecosystems and Human Well-Being: Synthesis**. Island Press. Washington D.C. 2005. Available at: <http://www.millenniumassessment.org/documents/document.356.aspx.pdf >. Accessed February 20, 2015.

MOONEN, A.; BARBERI, P. Functional biodiversity: An agroecosystem approach. **Agriculture, Ecosystems & Environment,** [C.l.], v. 127, n. 1-2, p. 7-21, August 2008.

MOREIRA, F. M. D. S. et al. Differentiation in the fertility of Inceptisols as related to land use in the upper Solimões river region, western Amazon. **Science of the Total Environment,** [C.l.], v. 408, n. 2, p. 349-355, December 2009.

MOREIRA, F. M. S. et al. Organismos do Solo em Ecossistemas Brasileiros. In: MOREIRA, F. M. S.;SIQUEIRA, J. O.*, et al* (Ed.). Biodiversidade do solo em ecossistemas brasileiros. Lavras: UFLA, 2008, p. 50-65.

MYROLD, D. D.; ZEGLIN, L. H.; JANSSON, J. K. The Potential of Metagenomic Approaches for Understanding Soil Microbial Processes. **Soil Science Society of America Journal,** v. 78, n. 1, p. 3-10, November 2014.

NAEEM, S. Species redundancy and ecosystem reliability. **Conservation Biology,** Malden, v. 12, n. 1, p. 39-45, July 1998.

NAEEM, S.; LI, S. Biodiversity enhances ecosystem reliability. **Nature,** London, v. 390, n. 6659, p. 507-509, December 1997.

NICOL, G. W. et al. The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. **Environmental Microbiology,** Malden, v. 10, n. 11, p. 2966-2978, August 2008.

NIELSEN, U. N. et al. Soil biodiversity and carbon cycling: a review and synthesis of

studies examining diversity-function relationships. **European Journal of Soil Science,** Malden, v. 62, n. 1, p. 105-116, November 2011.

NIELSEN, U. N. et al. The enigma of soil animal species diversity revisited: the role of small-scale heterogeneity. **PloS one,** San Francisco, v. 5, n. 7, p. e11567-e11567, July 2010.

ORIANS, G. H.; DIRZO, R.; CUSHMAN, J. H. **Biodiversity and Ecosystem Processes in Tropical Forests**. Springer Berlin Heidelberg, 1996. 1-9 ISBN 978-3-642-79757-6.

PENTON, C. R. et al. Functional genes to assess nitrogen cycling and aromatic hydrocarbon degradation: primers and processing matter. **Frontiers in Microbiology,** Lausanne, v. 4, n. September, p. 279-279, September 2013.

PICONE, C. Diversity and Abundance of Arbuscular–Mycorrhizal Fungus Spores in Tropical Forest and Pasture. **Biotropica,** Malden, v. 32, n. 4a, p. 734-750, March 2000.

RODRIGUES, J. L. M. 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, n. 3, p. 988-93, January 2013.

SCHIMEL, J. P.; BENNETT, J.; FIERER, N. Microbial community composition and soil nitrogen cycling: is there really a connection? In: BARDGETT, R.;USHER, M.*, et al* (Ed.). Biological diversity and function in soils. Cambridge: Cambridge University Press, 2005. p.171-188. ISBN 9780511541926.

SETÄLÄ, H.; BERG, M. P.; JONES, T. H. Trophic structure and functional redundancy in soil communities. In: BARDGETT, R.;USHER, M.*, et al* (Ed.): Cambridge University Press, 2005. p.236-249. ISBN 9780511541926.

SOLAR, R. R. D. C. **Effects of land-use change on tropical forest biodiversity: a multi-scale assessment in the Brazilian Amazon**. 2014. 182-182

STURMER, S. L.; SIQUEIRA, J. O. Species richness and spore abundance of arbuscular mycorrhizal fungi across distinct land uses in western Brazilian Amazon. **Mycorrhiza,** Heidelberg, v. 21, n. 4, p. 255-267, May 2011.

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,** v. 59, n. 4, p. 734-43, May 2010.

THOMAS, T.; GILBERT, J.; MEYER, F. Metagenomics - a guide from sampling to data analysis. **Microbial Informatics and Experimentation,** v. 2, n. 1, p. 3-3, February 2012.

TREUSCH, A. H. et al. Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. **Environmental Microbiology,** Washington, v. 7, n. 12, p. 1985-95, December 2005.

UNITED NATIONS. **Convention on biological diversity**. 1992. Available at: <https://www.cbd.int/doc/legal/cbd-en.pdf>. Accessed February 20, 2015.

VENTER, J. C. et al. Environmental genome shotgun sequencing of the Sargasso Sea. **Science,** New York, v. 304, n. 5667, p. 66-74, 2004.

WALKER, B. H. Biodiversity and ecological redundancy. **Conservation Biology,** Malden, v. 6, n. 1, p. 18-23, April 1992.

WALL, D. H.; FITTER, A. H.; PAUL, E. A. Developing new perspectives from advances in soil biodiversity research. In: BARDGETT, R.;USHER, M., *et al* (Ed.). Biological diversity and function in soils. Cambridge: Cambridge University Press, 2005. p.3-28. ISBN 9780511541926.

WANG, Q. et al. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. **Applied and Environmental Microbiology,** Washington, v. 73, n. 16, p. 5261-5267, June 2007.

WANG, Q. et al. Ecological patterns of nifH genes in four terrestrial climatic zones explored with targeted metagenomics using FrameBot, a new informatics tool. **mBio,** Washington, v. 4, n. 5, p. e00592-13, March 2013.

WARDLE, D. A. et al. Ecological linkages between aboveground and belowground biota. **Science,** New York, v. 304, n. 5677, p. 1629-1629, June 2004.

XU, J. Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. **Molecular Ecology,** Malden, v. 15, n. 7, p. 1713-1731, June 2006.

YACHI, S.; LOREAU, M. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. **Proceedings of the National Academy of Sciences of the United States of America,** Washington, v. 96, n. 4, p. 1463-1463, February 1999.

SECOND PART – PAPERS

PAPER 1

Standards of the journal – Ecology

Running title: Soil bacterial diversity in Amazon

Land use intensification in the humid tropics increased both alpha and beta diversity of soil bacteria

Teotonio Soares de Carvalho[1,2]

Ederson da Conceição Jesus[3]

Jos Barlow[2,4]

Toby A Gardner[5]

Isaac Carvalho Soares[3]

James M Tiedje[6]Fatima Maria de Souza Moreira[1]

Authors' affiliations:

[1]Departamento de Ciência do Solo, Universidade Federal de Lavras, Lavras, Minas Gerais, 37200-000, Brazil

[2]Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA14YQ United Kingdom

[3]Embrapa Agrobiologia, CNPAB, Rodovia BR 465, Km 7, CEP 23891-000, Seropédica, Rio de Janeiro, Brazil

[4]Museu Paraense Emilio Goeldi, Av. Magalhães Barata, 376, Belém-Pará-Brazil

Abstract.

[5]Stockholm Environment Institute, Linnégatan 87D, Box 24218, Stockholm, 10451, Sweden

[6]Center for Microbial Ecology, Michigan State University, East Lansing, Michigan, 48824, USA

ABSTRACT

Anthropogenic pressures on tropical forests are rapidly intensifying, but our understanding of their implications on biological diversity is still very limited, especially with regard to soil biota, and in particular soil bacteria communities. Here we evaluated bacterial community structure and diversity across a gradient of land use intensity in the eastern Amazon from undisturbed primary forest, through primary forests varyingly disturbed by fire, regenerating secondary forest, pasture, and mechanized agriculture. Soil bacteria were assessed by paired-end Illumina sequencing of 16S *rRNA* gene fragments (V4 region). The resulting sequences were clustered into operational taxonomic units (OTU) at a 97% similarity threshold. Land use intensification increased the observed bacterial diversity (both OTU richness and community heterogeneity across space) and this effect was strongly associated with changes in soil pH. Moreover, land use intensification and subsequent changes in soil fertility, especially pH, alter bacterial community structure, with pastures and areas of mechanized agriculture displaying the most contrasting communities in relation to undisturbed primary forest. Together, these results indicate that tropical forest conversion impacts soil bacteria not through loss of richness, as previously thought, but mainly by imposing marked shifts on bacterial community structure, with unknown yet potentially important implications for ecological functions and services performed by these communities.

Keywords: *Amazon forest, below-ground biodiversity, next-generation sequencing, drivers of bacterial community structure, 16S rDNA*.

INTRODUCTION

Tropical rainforests harbor some of the most biodiverse fauna and flora on the planet, and have major participation in global climate regulation and biogeochemical cycles (Malhi et al. 2014). These ecosystems are threatened by both widespread forest degradation (e.g. fire, logging, and overhunting), and clearance for agriculture and cattle production, as well as changes in temperature, precipitation and atmospheric chemistry (Malhi et al. 2014). Marked reductions in animal and plant diversity after forest conversion has been widely reported (Turner, 1996; Gibson et al., 2011), but its impact on soil fauna, and in particular on soil bacterial communities, is still poorly understood, despite the key role of these microorganisms in the functioning of terrestrial ecosystems (Bardgett and van der Putten, 2014) and their relevance for agricultural production systems.

In the Brazilian Amazon, rainforest conversion into agricultural systems can increase the availability of mineral nutrients and reduce soil acidity (Moreira et al., 2009; Braz et al., 2013), which is expected to stimulate bacterial growth (Bardgett and Cook, 1998; Rousk et al., 2009). Furthermore, land use intensification can drive the homogenization of plant communities and vegetation structure (Arroyo-Rodríguez, 2013), possibly reducing the diversity of food resources in the rhizosphere and in the litter, the diversity of soil microhabitats, and the diversity of hosts for symbiotic bacteria (Wardle, 2006).

Intriguingly, available evidence suggests that bacteria may respond differently from other taxa, with a number of studies reporting higher alpha (local) diversity following rainforest conversion (Jesus et al., 2009; Tripathi et al., 2012; Rodrigues et al., 2013; Mendes et al., 2015). However, Rodrigues et al (2013) found that, although conversion of primary forest to pasture increases alpha diversity, it can also decrease beta diversity, possibly resulting in lower

gamma (regional) diversity. Yet their suggestion that land use intensification homogenizes bacterial communities in Amazon has to be tested along a comprehensive land use gradient.

This study aims to address this knowledge gap by evaluating changes in bacterial community structure and diversity in the eastern Amazon across a broad land use intensification gradient from undisturbed primary forest, through primary forests varyingly disturbed by fire, regenerating secondary forest, pasture, and mechanized agriculture. Specifically, we test the following hypotheses: (i) Land use intensification homogenizes soil bacterial communities, with increased alpha, but reduced beta and gamma diversities in more intensive land uses; (ii) soil pH and its heterogeneity are positively associated with soil bacterial alpha and beta diversities, respectively, with the former increasing and the latter decreasing in more intensive land uses; (iii) land use intensification in the eastern Amazon changes soil bacterial community structure, including the dominance of two important phyla (Proteobacteria and Acidobacteria); and (iv) this effect is partly associated with increased nutrient availability and reduced acidity in more intensive land uses.

MATERIAL AND METHODS

Sampling

This study was conducted in the eastern Amazon, in the Santarém and Belterra municipalities in the state of Pará, Brazil (see **Figure 1** for geographical coordinates). The mean annual temperature for this area is 26 °C and the mean annual rainfall is 2,150 mm. This region was chosen because it encompasses a land use intensification gradient that is typical of many areas in the eastern Amazon, and across the human-modified tropics more generally, including the expansion of mechanized agriculture, extensive cattle production, and high rates

of forest degradation from unsustainable logging and wildfires (Gardner et al., 2013). Five land uses were chosen as follows:

- Undisturbed primary forest (UPF). Well-preserved primary forests with no signal of logging or previous fire regimes.
- Disturbed primary forest (DPF). Primary forests recently subjected to fire as evidenced by fire scars on trees and/or by remote sensing (Gardner et al. 2013). We did not find any evidence of logging within this disturbance class.
- Secondary forest (SF). Areas previously deforested for agriculture, 13-20 years after abandonment.
- Pastures (Pa). Areas covered with *Brachiaria* sp. and used for cattle production.
- Mechanized agriculture (MA). Areas under intensive agriculture for production of maize, soybean, and upland rice (often in the same year).

All areas of secondary forest, pasture, or mechanized agriculture were deforested more than 20 years ago. Although land use classification is straightforward for UPF, Pa and MA, in some cases it can be much harder to distinguish between DPF and SF: forests that have suffered multiple burns become increasingly similar to secondary forests (e.g. Barlow and Peres, 2008), and it can be difficult to detect fine-scale spatial heterogeneity in historical patterns of forest clearance or burn intensity. As such, these two categories should be considered as broad indicators of forest disturbance and regeneration status. Details about the land use classification used in this work are provided in Gardner et al. (2013). Additional information about changes in forest structure can be found in Berenguer et al. (2014).

The sampling was conducted in April 2013, during the rainy season, in four micro-catchments selected from 18 catchments sampled as part of the Sustainable Amazon Network (Gardner et al. 2013). Four 250 m transects per

land use were distributed using a stratified-random sampling design among these catchments (Figure 1; see Gardner et al., 2013, for details). The exception was MA, with five transects. At five sampling points equally spaced (50 m) along each transect, two composite soil samples (one with 50 g for molecular analyses and another with 500 g for physicochemical analysis) were taken separately by pooling three subsamples collected from the 0-10 cm depth, after removing the soil litter layer. All samples for molecular analysis were immediately placed on ice for transport and permanently stored at -80 ºC until being lyophilized. Soil chemical and physical attributes were analyzed at the Department of Soil Science of the Federal University of Lavras, Minas Gerais, Brazil.

DNA Extraction

DNA was extracted from 0.5 g of lyophilized soil using the PowerSoil DNA Isolation Kit (MO BIO Laboratories Inc., Carlsbad, CA, USA) following the protocol suggested by Marty et al. (2012). Briefly, after adding the soil to the tubes containing the beads, a solution of sterilized skim milk 80 mg g^{-1} (Thermo Scientific™ Oxoid™ Skim Milk) was added to the soil suspension, which was heated at 70 °C for 10 min and subjected to bead-beating at maximum speed for 2 min. All subsequent steps were performed according to the manufacturer's instructions.

PCR Amplification and Sequencing

For PCR amplification, the primers 515f/806r targeting the V4 hypervariable region of the 16S rRNA gene were used as described in Caporaso et al. (2012). The amplicons were pooled in equimolar concentration using the SequalPrep plate normalization kit (Invitrogen), and the final concentration was quantified with a Qubit 2.0 Fluorometer (Invitrogen), and the Qubit dsDNA HS Assay Kit (Invitrogen). Illumina MiSeq paired-end (2 x 150 bp) sequencing was carried

out in the Research Technology Support Facility of the Michigan State University, U.S.A.

Sequence Analysis

The UPARSE pipeline (Edgar, 2013) was used for reads merging and quality filtering by truncating the sequence length at 250 bp, with a conservative maxee value of 0.25 (equivalent to one incorrect nucleotide for each of four sequences). Sequences were dereplicated and, after singletons were removed, they were clustered into OTU's (operational taxonomic units) at 97% similarity threshold using UPARSE. After that, the filtered sequences were mapped to the representative OTU's. To assign taxonomy for each representative OTU, the RDP classifier was used (Wang et al., 2007) with a threshold of 0.5.

Diversity Analysis

Hill's number, also known as true diversity index, was used to measure diversity because it integrates the most commonly used diversity indices in a single framework and it has desirable mathematical properties that make it intuitively interpretable with regards to the diversity concept (Hill, 1973; Jost, 2006). This diversity measure is calculated based on the formula:

$$^{q}D = \left(\sum_{i=1}^{S} p_i^q\right)^{1/(1-q)}$$

Where p_i is the abundance of the i-th species in the community, S is the total number of species (OTU's in this study), and q, referred as the order of the index, determines the measure's sensitivity to the presence of rare species. Following the recommendation of Chao et al. (2014), we measured diversity with three values of q, as follows: diversity of all OTU's (q = 0), "typical" OTU's (q = 1), and dominant OTU's (q = 2). To standardize the sampling effort, the OTU table was rarefied to the minimum sequencing depth (19,203 reads).

The variation in the identities/abundances of species among sampling units (β diversity in the sense of Whitttaker (1972)) was calculated separately for each transect by dividing the total diversity observed in the transect (ɣ diversity) by the mean diversity across sampling points in each transect (α diversity) according to the multiplicative partitioning framework presented by Jost (2007).

The effect of land use on the calculated diversity measures was evaluated using linear models estimated by generalized least square, due to heteroscedasticity, with the undisturbed primary forest as the baseline group.

To assess the relationship between OTU richness (response variable) and soil pH (explanatory variable), a linear mixed effects model was used with a random intercept for transects to account for the dependence among sampling points within transect. The significance of the effect of pH was assessed by the likelihood ratio test (Zuur et al., 2009). A regression analysis was used to assess the relationship between true beta diversity (using Hill's number of order 1) and the standard deviation of soil pH for each transect.

Community Analysis

For studying the community structure, the relative abundance of each OTU was estimated by repeatedly sampling from a Dirichlet distribution as outlined by Fernandes et al. (2014). The estimated relative abundances were square-root transformed to obtain the equivalent to the Hellinger transformation. Redundancy analysis (RDA) was used to assess the effect of each explanatory matrix (land use, soil fertility, and spatial variables) on the community structure. To verify the robustness of the ordination to the relative abundance estimation, the analysis was repeated 1000 times, resulting in practically no difference in the final ordination. The soil explanatory matrix used was based on the first two principal components calculated from the standardized soil physicochemical dataset containing the variables indicated in Appendix A. The spatial variables

were obtained through principal coordinates of neighbor matrices as described in Declerck et al (2011). Partial RDA was used to partition the community variation into the three explanatory matrices. To test the effect of land use on the individual abundances of the most generalist OTU's (those occurring in more than 20% of all sites), the ALDEx2 package was used in R according to Fernandes et al. (2014).

For comparing the (log-transformed) relative abundances of Acidobacteria, Proteobacteria, and the Proteobacteria:Acidobacteria ratio among land uses, linear mixed effects models were used with random intercept for transects. After verifying the significance of the model using likelihood ratio tests, the significance of treatment contrasts was assessed with the arm package in R, using a Markov chain Monte-Carlo method (MCMC; 10,000 simulations) according to Gelman and Hill (2007).

RESULTS

Diversity and land use intensification. About 6 million good quality sequences were obtained from all samples, from which 323,431 were discarded for occurring only once (singletons) as they may result from sequencing errors and inflate diversity measures (Huse et al., 2010). Clustering them at 97% similarity resulted in 12,928 operational taxonomic units (OTU's).

Bacterial diversity was generally higher in the disturbed systems than in the undisturbed primary forest (UPF) (**Figure 2**). Based on the Hill's number (^{q}D), for all OTU's (q = 0), "typical" OTU's (q = 1), and dominant OTU's (q = 2), we observed that all components of diversity (alpha, beta, and gamma) were consistently higher in mechanized agriculture (MA) than in UPF, except for the beta diversity of dominant OTU's. The same was observed for pasture, except for alpha diversity of typical OTU's and for all components of diversity for dominant OTU's. Compared to UPF, disturbed primary forest (DPF) and

secondary forest (SF) showed higher gamma diversity for all OTU's, but lower alpha and gamma diversity for dominant OTU's. Beta diversity was higher in SF than in UPF for all OTU's and typical OTU's.

Combining the results of all land uses, observed OTU richness was positively associated with soil pH ($P < 0.001$, $\chi^2(1) = 62$) (**Figure 3-A**). In the same way, beta diversity (Hill's number of order 1) for each transect was positively related to the standard deviation of soil pH in the transect, regardless the land use ($P < 0.001$, $DF_{residuals} = 19$, $F = 46$) (**Figure 3-B**).

Community structure and land use intensification. The bacterial community structure varied significantly along the gradient of land use intensity as revealed by redundancy analysis (RDA, P < 0.001; $F = 16.527$, $DF_{residuals} = 99$, 1000 permutations). Variance partitioning using soil variables and land use as constraining factors showed that land use explained 38% (adjusted R^2) of the total variation in community structure, while the soil variables explained 34% (adjusted R^2). The covariation between soil and land use uniquely explained 25% (adjusted R^2) of the total community variation. The projection of soil variables, especially pH, on the ordination axis resulting from RDA – using land use as constraining factor – showed that the soil variables were strongly associated with the first axis, which represented about 80% of the variance explained by land use (**Figure 4**). Spatial variables explained 45% (adjusted R^2) of the total community variation (RDA, $P < 0.001$, $F = 2.72$, $DF_{residuals} = 88$, 1000 permutations). However, 75% of the spatial structure of the community (variation explained by spatial variables) was explained by land use and soil variables.

For the 4,893 most generalists OTU's (those occurring in more than 20% of all sites), a significant effect of land use on abundance was observed in 3,776 (Kruskal-Wallis test; $P < 0.05$ after false discovery rate correction) (**Figure 5**).

For the phyla Proteobacteria and Acidobacteria, we observed reduced dominance of the former in MA ($P \leq 0.001$) and PA ($P \leq 0.05$) in relation to UPF; and increased dominance of the latter in MA ($P \leq 0.05$) in relation to UPF. The Proteobacteria:Acidobacteria ratio was significantly lower in MA ($P \leq 0.001$) and PA ($P \leq 0.05$) than in UPF (**Figure 6-A**). The dominant taxonomic classes were also contrasting among land uses (**Figure 6-B**).

DISCUSSION

Diversity and land use intensification. Microbial diversity is considered an indicator of soil health (Nielsen and Winding, 2002) because more diverse ecosystems are predicted to be more reliable (Naeem and Li, 1997; Naeem, 1998) and productive (Yachi and Loreau, 1999). However, our results suggest that bacterial diversity may not reflect soil biological quality in Amazon, if one accepts overall diversity as quality. In this study, all components of diversity (alpha, beta, and gamma) were generally higher in the more intensive land uses, especially mechanized agriculture (MA) and pasture (PA), than in the undisturbed primary forest (UPF). This was mainly due to the less abundant operational taxonomic units (OTU's) because as the sensitiveness of the index to rare species was reduced (increasing the order q) the difference in diversity either disappeared or was inverted, except for MA (Figure 2).

Increased bacterial alpha diversity after conversion of rainforests to anthropogenic systems has been extensively reported (Jesus et al., 2009; Tripathi et al., 2012; Rodrigues et al., 2013; Mendes et al., 2015). The clear relationship that we observed between OTU richness and soil pH indicates a likely mechanism by which land use intensification increases bacterial alpha diversity in Amazon (Figure 3a), supporting previous studies that report a linear increase in OTU richness with increasing soil pH (within the range of pH of our study,

3.7 - 7.4) in other ecosystems (Fierer and Jackson, 2006; Lauber et al., 2009; Rousk et al., 2010; Tripathi et al., 2012).

The natural acidic and nutrient-poor conditions of the soils studied here are unsuitable for agriculture and cattle production, usually requiring large quantities of lime to neutralize the pH and fertilizers to elevate the nutrient content (Sanchez et al., 1983). To a smaller extent, slash-and-burn practices or occasional forest fires may elevate soil pH and increase nutrient contents in secondary and disturbed primary forests through the deposition of ashes from combusted forest biomass (Ohno and Erich, 1990; Giardina et al., 2000; Moreira et al., 2009).

Liming of acidic soils increases bacterial growth (Lupawayi et al., 2009; Rousk et al., 2009) and soil fertilization favors bacteria to the detriment of fungi, leading to a bacteria-dominated food web (Bardgett and Cook, 1998; de Vries et al., 2006). This dominance is related to increased mineralization rates (Wardle et al., 2004) and N-losses by leaching (de Vries et al., 2006, 2012). In this context, the observed increases in bacterial diversity after forest conversion should probably be interpreted as decreased, rather than increased, soil biological quality for farming systems and forest regeneration.

While our results regarding the alpha diversity of soil bacteria are closely aligned with those of other researchers, the fact that we found higher beta diversity in more intensive land uses (Figure 2) challenges the hypothesis of biotic homogenization of bacterial communities after rainforest conversion (Rodrigues et al., 2013). However, these authors studied only two contrasting land uses (primary forest and pasture) whereas we studied a more comprehensive gradient with at least four replicates for each of the five studied land uses. In fact, this gradient is clearly reflected by the structure of bacterial communities (Figure 4). The differences between our results and those of Rodrigues et al. (2013) may also be explained by differences in forest

structure and dynamics between the two studied Amazon regions. In the western Amazon studied by Rodrigues et al. (2013), the rate of stem turnover (rate in which trees die and are replaced) is higher than in the eastern Amazon (region used in our study) partly because of differences in soil age and type of parent material (Quesada et al., 2011) and their implications on soil fertility and physical conditions (Quesada et al., 2012). It could be that the higher intrinsic rate of natural disturbances in western Amazonian forests (Quesada et al., 2012) results in higher heterogeneity of bacterial communities in these soils when compared to those in eastern Amazonia.

The high levels of beta diversity we observed in soil bacteria may be related to concomitantly high levels of heterogeneity in soil pH in more intensive land uses diversity (Figure 3b). This is not the first report of increased bacterial beta diversity after rainforest conversion (Lee-Cruz et al., 2014), but the first to relate it to soil pH heterogeneity. A recent study of bacterial diversity in grasslands reported pH and plant richness as predictors of bacterial beta diversity (mean community dissimilarity) (Prober et al., 2015). But in our study, pH heterogeneity had a remarkably dominant effect on bacterial beta diversity, in spite of the much higher plant diversity in UPF as compared to MA and PA. Therefore, at the scale of our study, all components of diversity of soil bacteria appear to be uncoupled from plant diversity.

Community structure and land use intensification. Supporting our last two hypothesis, we found evidence that changes in soil parameters, especially pH, mediate the effect of land use intensification on bacterial community structure, as indicated by the results of the redundancy analysis and variance partitioning. Mechanized agriculture results in a completely distinct bacterial community in relation to that of UPF as indicated by the analysis of the most generalist OTU's (Figure 5), whereas PA and SF have an intermediary impact, sharing both some of the OTU's most abundant in MA and some of those most abundant in UPF.

Members of Acidobacteria subdivisions were among the dominant classes in all land uses (Figure 6) and they also were strongly related to the first ordination axis (Figure 4). In agreement with previous studies, Acidobacteria subdivisions 1, 2, 3, and 13 were more dominant in acidic, nutrient-poor conditions; and subgroups 4, 6, 7, and 16 were more dominant in neutral, nutrient-rich conditions (Jones et al., 2009; Lauber et al., 2009; Navarrete et al., 2013).

At the phylum level, Acidobacteria dominance was higher in MA than in UPF, while the dominance of Proteobacteria was considerably reduced in MA and PA in relation to UPF (Figure 6). These two phyla, and their abundance ratio, have been extensively used as indicators of the soil nutritional status because Proteobacteria usually prefers labile organic C pools, whereas Acidobacteria is adapted to low organic C quality and/or quantity (Fierer et al., 2007). This ratio is usually higher in the rhizosphere or when sucrose is artificially added to soil (Fierer et al., 2007), and reduced when the vegetation is removed (Thomson et al., 2010, 2013). The observed reduction of that ratio in PA and MA suggests that the removal of the forest cover reduces the availability of labile substrates in the soil or that other unrecognized factors are shifting this ratio in these tropical soils.

These results have implications for ecosystem conservation and the long-term sustainability of agricultural land-uses because, although forest conversion does not seem to reduce bacterial diversity, it considerably changes bacterial community structure, even after establishment of secondary vegetation. These changes are expected to affect ecosystem processes mediated by bacteria (Waldrop et al., 2006; McGuire et al., 2010; Kaiser et al., 2010). In fact, the metagenome predictions (Appendix B) suggest that these changes in bacterial community structure are altering its functional traits, especially in MA, when compared to UPF, although further studies targeting functional genes or using shotgun metagenomics are necessary to confirm those predictions.

Consequently, assessments of the impacts of land use intensification on tropical biota should include soil bacteria, but focus on changes in bacterial community structure and their effects on ecosystem functioning, rather than on diversity losses. Future research is also needed to verify whether the apparent favoring of soil bacteria after land use intensification in tropical forest landscapes is occurring at the detriment of soil fungi.

ACKNOWLEDGEMENTS

We are grateful to the following for financial support: Instituto Nacional de Ciência e Tecnologia – Biodiversidade e Uso da Terra na Amazônia (CNPq 574008/2008-0, and 400640/2012-0), Empresa Brasileira de Pesquisa Agropecuária – Embrapa (SEG: 02.08.06.005.00), the UK government Darwin Initiative (17-023), The Nature Conservancy, the UK Natural Environment Research Council (NERC) (NE/F01614X/1 and NE/G000816/1, NE/F015356/2; NE/l018123/1; NE/K016431/1), the Swedish Formas 2013-1571, Conselho Nacional de Desenvolvimento Científico e Tecnológico for the research productivity scholarship and grant to F.M.S. Moreira (processes 304574/2010-4 and 477786/2011-1), the support at MSU from RDP under US Department of Energy, Grant # DE-FG02-99ER62848, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the scholarship provided to the first author. We also thank the farmers and workers unions of Santarém, Belterra and Paragominas and all collaborating private landowners for their support

LITERATURE CITED

Arroyo-Rodríguez, V., M. Rös, F. Escobar, F. P. L. Melo, B. A. Santos, M. Tabarelli, and R. Chazdon. 2013. Plant β-diversity in fragmented rain forests: Testing floristic homogenization and differentiation hypotheses. Journal of Ecology **101**:1449-1458.

Bardgett, R. D., and R. Cook. 1998. Functional aspects of soil animal diversity in agricultural grasslands. Applied Soil Ecology **10**:263-276.

Bardgett, R. D., and W. H. van der Putten. 2014. Belowground biodiversity and ecosystem functioning. Nature **515**:505-511.

Barlow, J., and C. A. Peres. 2008. Fire-mediated dieback and compositional cascade in an Amazonian forest. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **363**:1787-1794.

Berenguer, E., J. Ferreira, T. A. Gardner, L. E. O. C. Aragão, P. B. De Camargo, C. E. Cerri, M. Durigan, R. C. D. Oliveira, I. C. G. Vieira, and J. Barlow. 2014. A large-scale field assessment of carbon stocks in human-modified tropical forests. Global Change Biology **2005**:3713-3726.

Caporaso, J. G., et al. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME Journal **6**:1621-1624.

Chao, A., N. J. Gotelli, and T. C. Hsieh. 2014. Rarefaction and extrapolation with Hill numbers : a framework for sampling and estimation in species diversity studies. Ecological Monographs **84**:45-67.

De Souza Braz, A. M., A. R. Fernandes, and L. R. F. Alleoni. 2013. Soil attributes after the conversion from forest to pasture in Amazon. Land Degradation and Development **24**:33-38.

de Vries, F. T., J. Bloem, H. Quirk, C. J. Stevens, R. Bol, and R. D. Bardgett. 2012. Extensive Management Promotes Plant and Microbial Nitrogen Retention in Temperate Grassland. PLoS ONE **7**:1-12.

de Vries, F. T., E. Hoffland, N. van Eekeren, L. Brussaard, and J. Bloem. 2006. Fungal/bacterial ratios in grasslands with contrasting nitrogen management. Soil Biology and Biochemistry **38**:2092-2103.

Declerck, S. A. J., J. S. Coronel, P. Legendre, and L. Brendonck. 2011. Scale dependency of processes structuring metacommunities of cladocerans in temporary pools of High-Andes wetlands. Ecography **34**:296-305.

Edgar, R. C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature methods **10**:996-998.

Fernandes, A. D., J. N. Reid, J. M. Macklaim, T. A. McMurrough, D. R. Edgell, and G. B. Gloor. 2014. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. Microbiome **2**:15-15.

Fierer, N., M. A. Bradford, and R. B. Jackson. 2007. Toward an ecological classification of soil bacteria. Ecology **88**:1354-1364.

Fierer, N., and R. Jackson. 2006. The diversity and biogeography of soil bacterial communities. Proceedings of the National Academy of Sciences of the United States of America **103**:626-631.

Gardner, T. A., et al. 2013. A social and ecological assessment of tropical land uses at multiple scales : the Sustainable Amazon Network. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **368**:20120166-20120166.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York.

Giardina, C., R. Sanford, I. Døckersmith, and V. Jaramillo. 2000. The effects of slash burning on ecosystem nutrients during the land preparation phase of shifting cultivation. Plant and Soil **220**:247-260.

Gibson, L., et al. 2011. Primary forests are irreplaceable for sustaining tropical biodiversity. Nature **478**:378-381.

Hill, M. O. 1973. Diversity and Evenness: A Unifying Notation and Its Consequences. Ecology **54**:427-432.

Huse, S. M., D. M. Welch, H. G. Morrison, and M. L. Sogin. 2010. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology **12**:1889-1898.

Jesus, E. C., T. L. Marsh, J. M. Tiedje, and F. M. S. Moreira. 2009. Changes in land use alter the structure of bacterial communities in Western Amazon soils. The ISME Journal **3**:1004-1011.

Jones, R. T., M. S. Robeson, C. L. Lauber, M. Hamady, R. Knight, and N. Fierer. 2009. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. The ISME Journal **3**:442-453.

Jost, L. 2006. Entropy and diversity. Oikos **113**:363-375.

Jost, L. 2007. Partitioning diversity into independent alpha and beta components. Ecology **88**:2427-2439.

Kaiser, C., M. Koranda, B. Kitzler, L. Fuchslueger, J. Schnecker, P. Schweiger, F. Rasche, S. Zechmeister-Boltenstern, A. Sessitsch, and A. Richter. 2010. Belowground carbon allocation by trees drives seasonal patterns of extracellular enzyme activities by altering microbial community composition in a beech forest soil. The New Phytologist **187**:843-858.

Lauber, C. L., M. Hamady, R. Knight, and N. Fierer. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Applied and Environmental Microbiology **75**:5111-5120.

Lee-Cruz, L., D. P. Edwards, B. M. Tripathi, and J. M. Adams. 2013. Impact of logging and forest conversion to oil palm plantations on soil bacterial communities in Borneo. Applied and Environmental Microbiology **79**:7290-7297.

Lupwayi, N. Z., M. A. Arshad, R. H. Azooz, and Y. K. Soon. 2009. Soil microbial response to wood ash or lime applied to annual crops and perennial grass in an acid soil of northwestern Alberta. Canadian Journal of Soil Science **89**:169-177.

Malhi, Y., T. a. Gardner, G. R. Goldsmith, M. R. Silman, and P. Zelazowski. 2014. Tropical Forests in the Anthropocene. Annual Review of Environment and Resources **39**:125-159.

Marty, F., J.-f. Ghiglione, S. Païssé, H. Gueuné, M. C. M. V. Loosdrecht, and G. Muyzer. 2012. Evaluation and optimization of nucleic acid extraction methods for the molecular analysis of bacterial communities associated with corroded carbon steel. Biofouling **28**:363-380.

McGuire, K. L., and K. K. Treseder. 2010. Microbial communities and their relevance for ecosystem models: Decomposition as a case study. Soil Biology and Biochemistry **42**:529-535.

Mendes, L. W., S. M. Tsai, A. A. Navarrete, M. de Hollander, J. A. van Veen, and E. E. Kuramae. 2015. Soil-Borne Microbiome: Linking Diversity to Function. Microbial Ecology.

Moreira, F. M. S., R. S. A. Nóbrega, E. D. C. Jesus, D. F. Ferreira, and D. V. Pérez. 2009. Differentiation in the fertility of Inceptisols as related to land use in the upper Solimões river region, western Amazon. Science of the Total Environment **408**:349-355.

Naeem, S. 1998. Species redundancy and ecosystem reliability. Conservation Biology **12**:39-45.

Naeem, S., and S. Li. 1997. Biodiversity enhances ecosystem reliability. Nature **390**:507-509.

Navarrete, A. A., E. E. Kuramae, M. de Hollander, A. S. Pijl, J. A. van Veen, and S. M. Tsai. 2013. Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. FEMS Microbiology Ecology **83**:607-621.

Nielsen, M. N., and A. Winding. 2002. Microorganisms as Indicators of Soil Health. National Environmental Research Institute, Aarhus, Denmark.

Ohno, T., and M. Susan Erich. 1990. Effect of wood ash application on soil pH and soil test nutrient levels. Agriculture, Ecosystems & Environment **32**:223-239.

Prober, S. M., et al. 2015. Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. Ecology Letters **18**:85-95.

Quesada, C. A., J. Lloyd, L. O. Anderson, N. M. Fyllas, M. Schwarz, and C. I. Czimczik. 2011. Soils of Amazonia with particular reference to the RAINFOR sites. Biogeosciences **8**:1415-1440.

Quesada, C. A., et al. 2012. Basin-wide variations in Amazon forest structure and function are mediated by both soils and climate. Biogeosciences **9**:2203-2246.

Rodrigues, J. L. M., et al. 2013. 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 **110**:988-993.

Rousk, J., E. Bååth, P. C. Brookes, C. L. Lauber, C. Lozupone, J. G. Caporaso, R. Knight, and N. Fierer. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. The ISME Journal **4**:1340-1351.

Rousk, J., P. C. Brookes, and E. Bååth. 2009. Contrasting soil pH effects on fungal and bacterial growth suggest functional redundancy in carbon mineralization. Applied and Environmental Microbiology **75**:1589-1596.

Sanchez, P. A., J. H. Villachica, and D. E. Bandy. 1983. Soil Fertility Dynamics after Clearing a Tropical Rainforest in Peru1. Soil Science Society of America Journal **47**:1171-1171.

Thomson, B. C., N. Ostle, N. McNamara, M. J. Bailey, A. S. Whiteley, and R. I. Griffiths. 2010. Vegetation affects the relative abundances of dominant soil bacterial taxa and soil respiration rates in an upland grassland soil. Microbial Ecology **59**:335-343.

Thomson, B. C., N. J. Ostle, N. P. McNamara, S. Oakley, A. S. Whiteley, M. J. Bailey, and R. I. Griffiths. 2013. Plant soil interactions alter carbon cycling in an upland grassland soil. Frontiers in Microbiology **4**:253-253.

Tripathi, B. M., M. Kim, D. Singh, L. Lee-Cruz, A. Lai-Hoe, A. N. Ainuddin, R. Go, R. A. Rahim, M. H. A. Husni, J. Chun, and J. M. Adams. 2012. Tropical soil

bacterial communities in Malaysia: pH dominates in the equatorial tropics too. Microbial Ecology **64**:474-484.

Turner, I. M. 1996. Species loss in fragments of tropical rain forest: a review of the evidence. Journal of Applied Ecology **33**:200-209.

Waldrop, M. P., and M. K. Firestone. 2006. Response of microbial community composition and function to soil climate change. Microbial Ecology **52**:716-724.

Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology **73**:5261-5267.

Wardle, D. A. 2006. The influence of biotic interactions on soil biodiversity. Ecology Letters **9**:870-886.

Wardle, D. A., L. R. Walker, and R. D. Bardgett. 2004. Ecosystem properties and forest decline in contrasting long-term chronosequences. Science **305**:509-513.

Whittaker, R. H. 1972. Evolution and Measurement of Species Diversity. Taxon **21**:213-251.

Yachi, S., and M. Loreau. 1999. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. Proceedings of the National Academy of Sciences of the United States of America **96**:1463-1463.

Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Springer, New York, NY.

LIST OF FIGURES

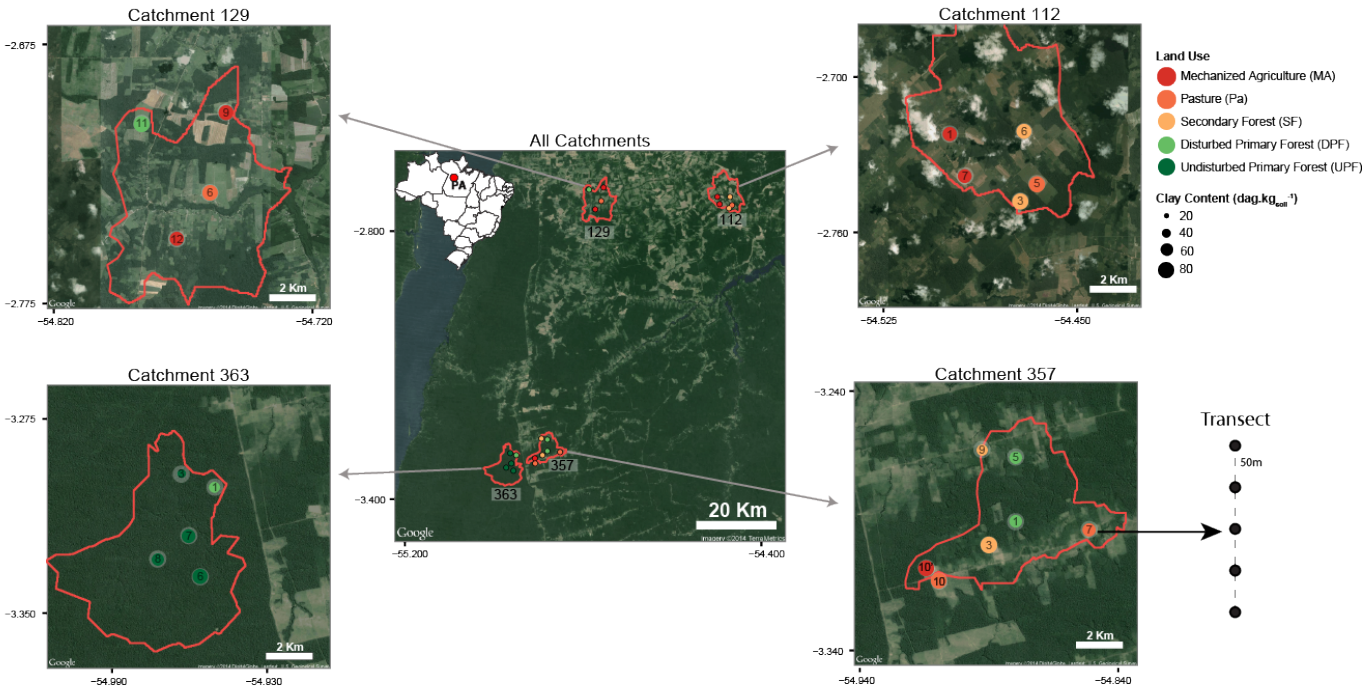

Figure 1. Satellite images of the four catchments used in this study. Each point represents one transect with five sampling units spaced by 50 meters. Point colors and sizes indicate, respectively, the land use and the mean clay content for each transect as shown in the legend. The x and y axis represent the latitude and longitude, respectively, in decimal degrees.

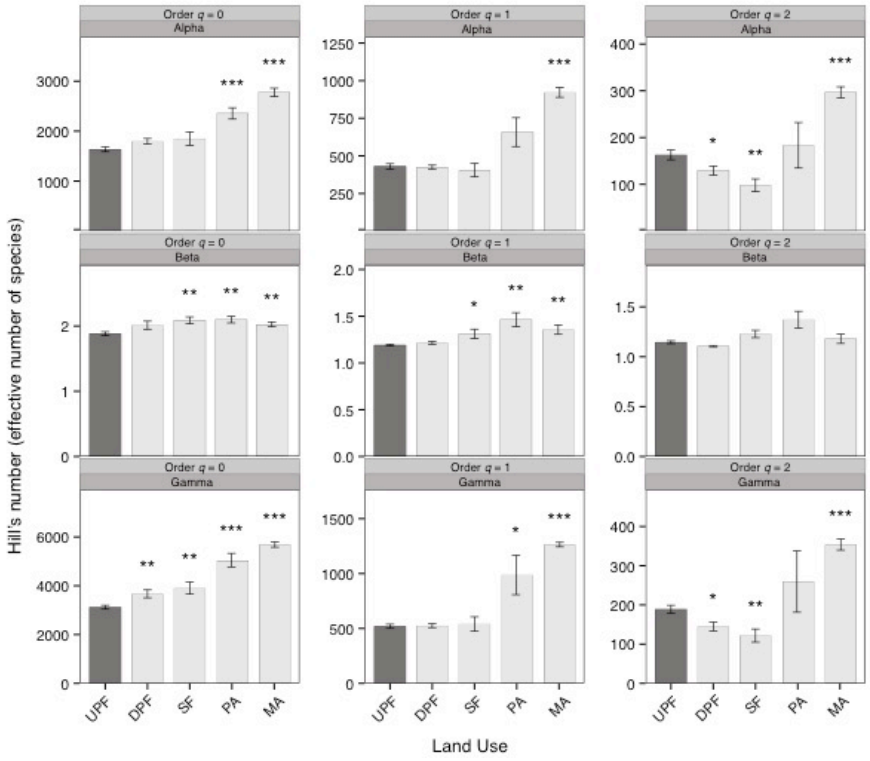

Figure 2. Alpha, beta and gamma bacterial diversity for all OTU's (q = 0), "typical" OTU's (q = 1), and dominant OTU's (q = 2). The bar height indicates the mean for each land use system and the error bar shows the standard error of the mean (UPF, undisturbed primary forest, n = 4; DPF, disturbed primary forest, n = 4; SF, secondary forest, n = 4; PA, pasture, n = 4; MA, mechanized agriculture, n = 5). The means for each land use is compared to the reference (UPF), in dark grey, using generalized least squares. * Significant at $P \leq 0.05$; ** Significant at $P \leq 0.01$; *** Significant at $P \leq 0.001$.

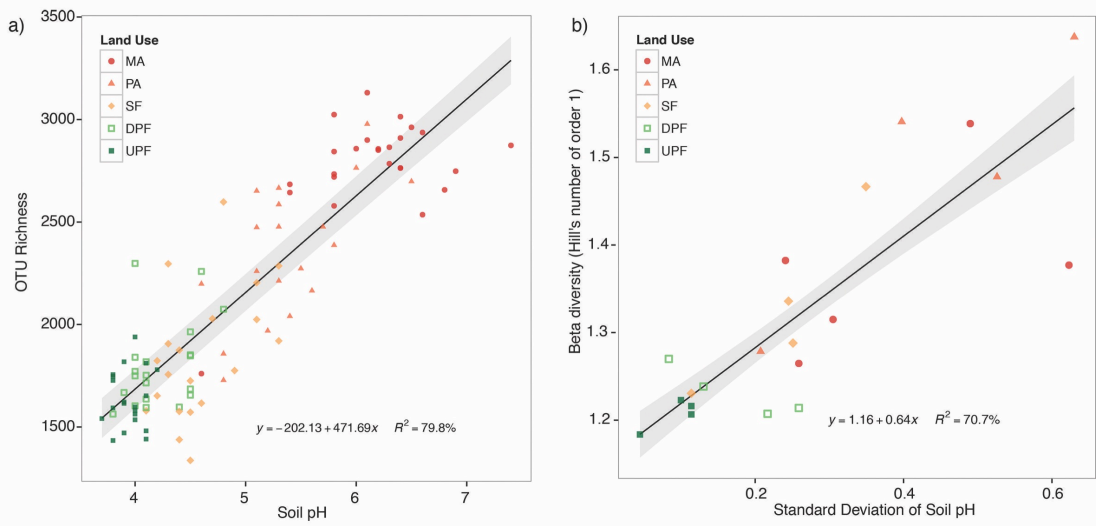

Figure 3. Models relating a) OTU Richness and soil pH; b) beta diversity and standard deviation of soil pH for each transect. The fitted values for each model is represented as the black line and their standard errors are indicated by the shaded area. Both models are significant at $P < 0.001$. UPF, undisturbed primary forest; DPF, disturbed primary forest; SF, secondary forest; PA, pasture; MA, mechanized agriculture.

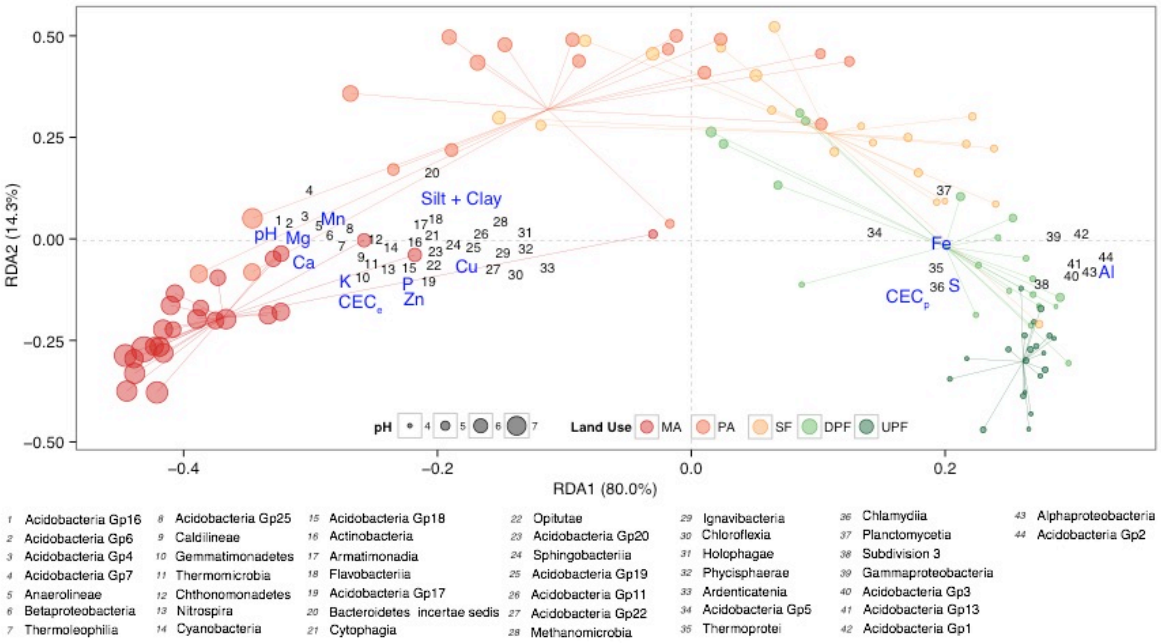

Figure 4. Redundancy analysis of the effect of land use on the abundance of soil bacteria in the eastern Amazon. Taxonomic classes (in black) and soil variables (in blue) were projected on the ordination axis using Kendall rank and Pearson correlation, respectively, and only those with significant correlation ($P \leq 0.05$; 1000 permutations) are shown. Taxonomic classes are indicated by numbers according to the legend at bottom. Point sizes indicates soil pH and their colors represent the land use system according to the legend (UPF, undisturbed primary forest; DPF, disturbed primary forest; SF, secondary forest; PA, pasture; MA, mechanized agriculture). The percentages indicated in both axes indicate the fraction of the community variation explained by land use (38% of the total variation) that is represented by the respective axis. CEC_p, potential cation exchange capacity; CEC_e, effective cation exchange capacity.

Figure 5. Heatmap of (center log-ratio) scaled relative abundances of the most generalist OTU's (occurrence > 20% for all sites). Among the 4,893 OTU's that fell in this category, only those significantly affected by land use (Kruskal-Wallis test; $P < 0.05$ after false discovery correction) are shown (3,776 in total). The OTU's are arranged in rows; the samples, in columns; and the cell colors indicate the scaled relative abundance according to the legend at bottom. Rows are grouped into taxonomic classes as indicated in the right axis, and columns are grouped into land uses, as indicated at top. UPF, undisturbed primary forest; DPF, disturbed primary forest; SF, secondary forest; PA, pasture; MA, mechanized agriculture

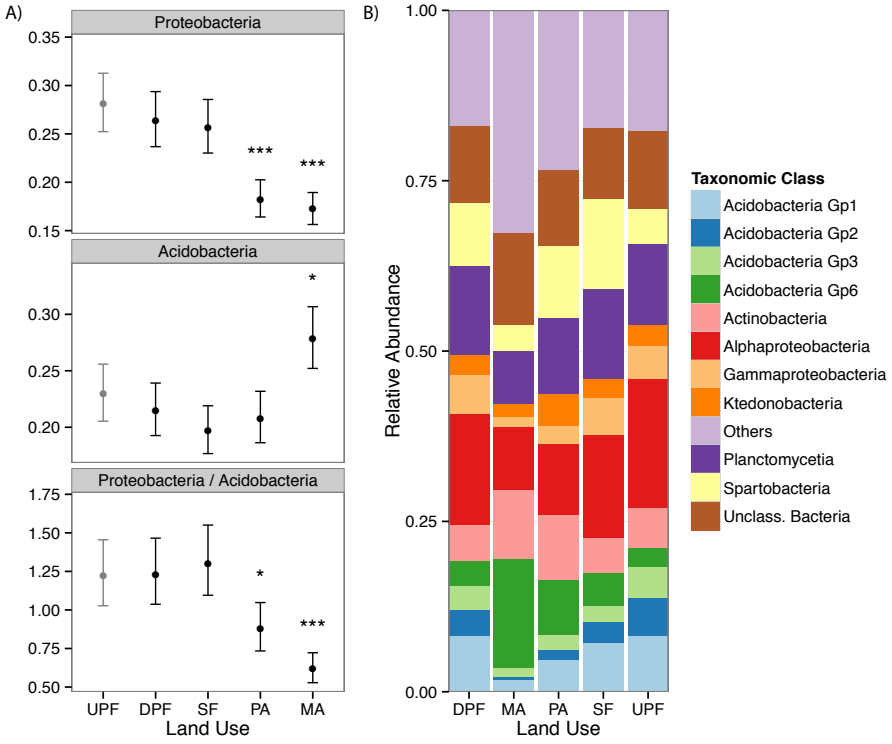

Figure 6. Dominant bacterial taxa in eastern Amazon soils. A) Mean relative abundances of the phylum Proteobacteria, Acidobacteria; and Protebacteria:Acidobacteria ratio. Vertical bars indicate the 95% confidence interval calculated using MCMC (10,000 simulations). Means are compared to UPF (in grey). * Significant at $P \leq 0.05$; *** Significant at $P \leq 0.001$. B) Relative abundances of the dominant taxonomic classes. UPF, undisturbed primary forest; DPF, disturbed primary forest; SF, secondary forest; PA, pasture; MA, mechanized agriculture.

Teotonio S. de Carvalho, Ederson C. Jesus, Jos Barlow, Toby A. Gardner, Isaac C. Soares, James M. Tiedje, Fatima M.S. Moreira. 2015. Land use intensification in the humid tropics increased both alpha and beta diversity of soil bacteria. *Ecology*.

APPENDIX A. Soil properties in the sampled sites.

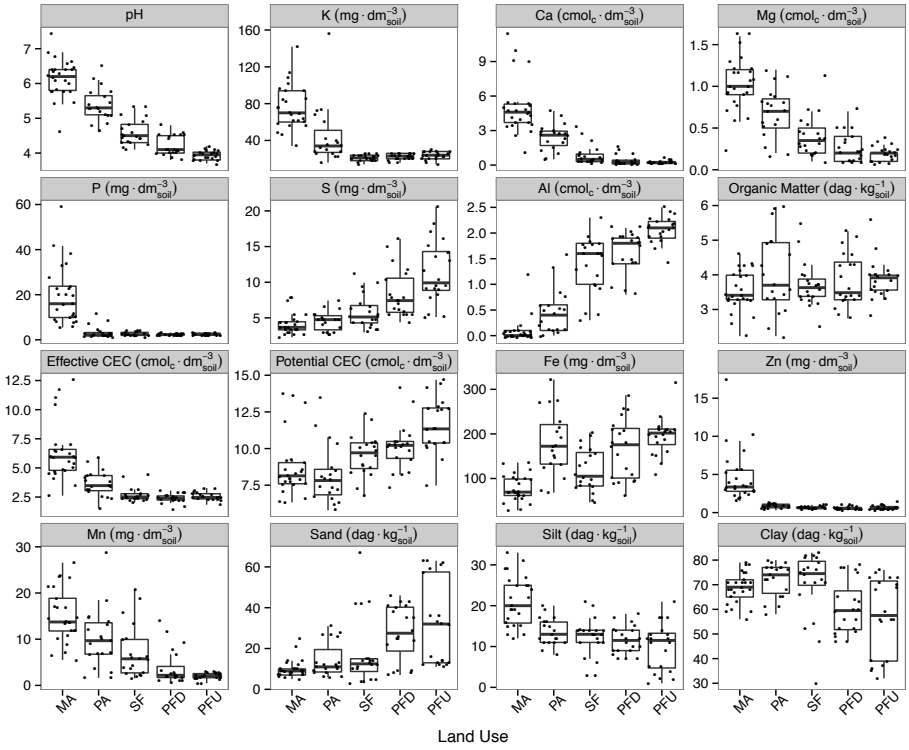

Figure A1. Boxplots of physicochemical parameters of soils from eastern Amazon. Points were jittered to facilitate visualization. UPF, undisturbed primary forest; DPF, disturbed primary forest; SF, secondary forest; PA, pasture; MA, mechanized agriculture.

Teotonio S. de Carvalho, Ederson C. Jesus, Jos Barlow, Toby A. Gardner, Isaac C. Soares, James M. Tiedje, Fatima M.S. Moreira. 2015. Land use intensification in the humid tropics increased both alpha and beta diversity of soil bacteria. *Ecology*.

APPENDIX B. Metagenome predictions using PICRUSt

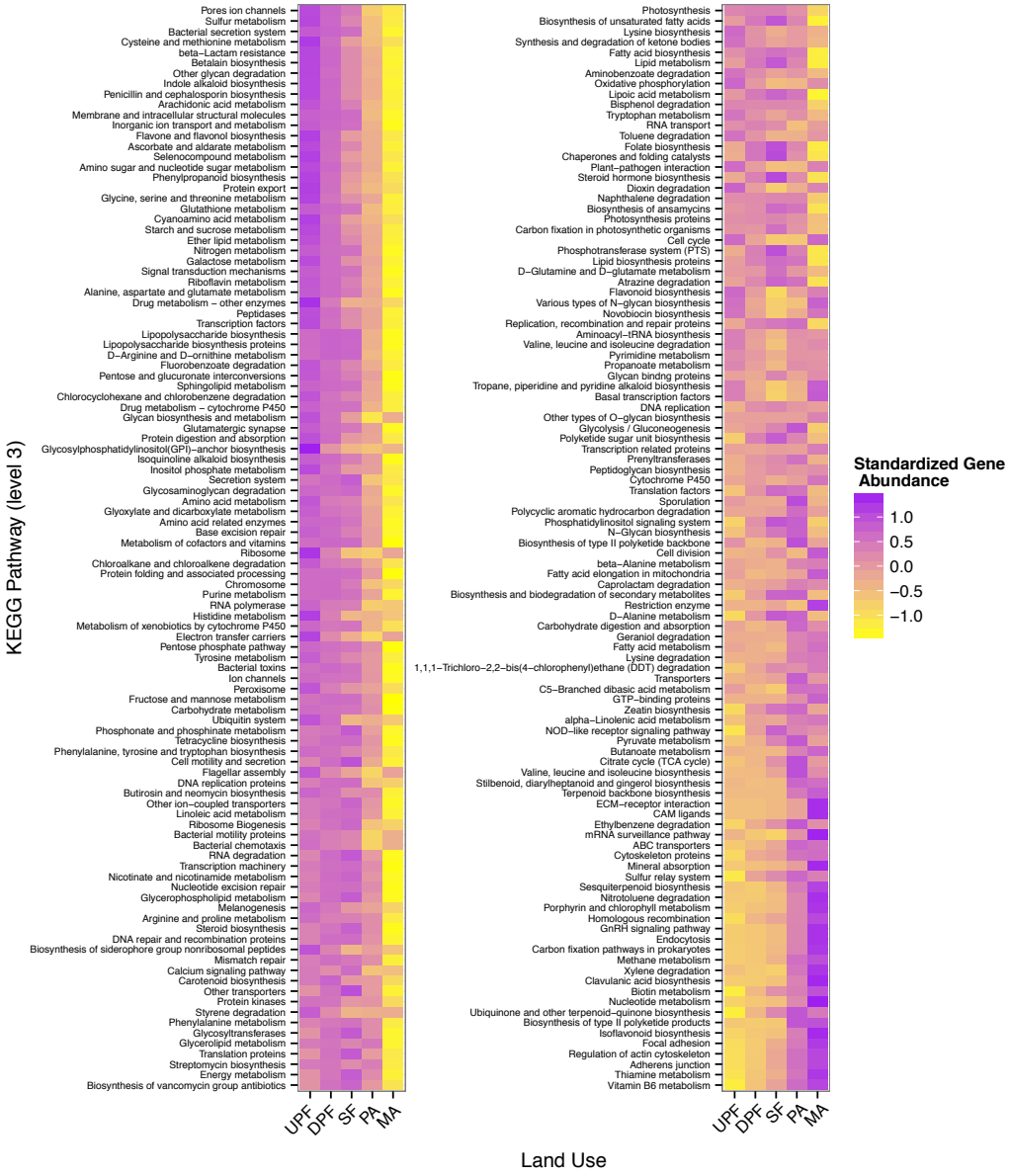

Figure B1. Heatmap of predicted gene abundances aggregated into KEGG pathways (level 3) for bacterial communities in eastern Amazon. Colors indicate the means of standardized (centered and scaled) gene abundance for each pathway in each land use according to the legend.

Metagenome prediction

Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt 1.0.0) was used to predict gene abundances from 16S *rRNA* data, according to the recommended protocol (Langille et al. 2013) available at http://picrust.github.com/picrust/. Briefly, paired-end merged and filtered 16S sequences were clustered into OTU's at 94% similarity threshold using closed-reference OTU picking in QIIME 1.8.0 (Corporaso et al., 2010). The OTU's abundances were normalized by 16S rRNA gene copy number before metagenome predictions. The obtained gene abundances were aggregated into Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (level 3).

LITERATURE CITED

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data, pp. 335-336.

Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R., Beiko, R.G., Huttenhower, C., 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature Biotechnology 31, 814-821.

PAPER 2

Standards of the journal – Soil Biology and Biochemistry

Soil bacterial communities are highly stable to a dry/rewet event in the eastern Amazon

ABSTRACT

Soil bacteria may experience increased exposures to drought after rainforest degradation and conversion to agricultural systems, especially under the predicted effects of climate change. Because soil bacteria differ in their ability to cope with dry/rewet events, we hypothesized that changes in bacterial community structure associated with land use intensification alter the stability of bacterial communities, and of the processes they mediate, to those events. To test these hypotheses, we conducted a greenhouse experiment with a total of 200 microcosms (undisturbed soil samples) taken from four land uses (primary forest, secondary forest, mechanized agriculture, and pasture) in the eastern Amazon. Half of these microcosms were left with no water addition for a 60 days period, after which the soil was rewetted to 80% of the field capacity (dry/rewet treatment). The remaining microcosms were used as control and kept at 80% of the field capacity for the whole duration of the experiment. At 1, 5, 10, 20, and 40 days after rewetting, microcosms from both treatments were destructively analyzed for the potential of ammonia oxidation and for the bacterial community structure using next-generation sequencing. Regardless of the land use, bacterial community structure and richness were insensitive to the dry/rewet treatment at both the OTU or phyla level. Similarly, the potential of ammonia oxidation was not affected by the dry/rewet treatment in any land use, but it was higher in primary forest.

INTRODUCTION

Soil bacteria mediate several ecosystems processes and are key components in biogeochemical cycles (Falkowski et al., 2008). In the Amazon region, land use intensification alters bacterial community structure in soil at several taxonomic levels (Jesus et al., 2009; Rodrigues et al., 2013; Paper 01) with possible impacts on ecosystem processes. Rainforest conversion can also degrade the soil structure and increase the susceptibility to compaction (Martins et al., 2012), reducing water infiltration rate in soil (Lal et al. 1979; Lal 1989), and favoring runoff and erosion processes (Maeda et al., 2008) that reduce water availability in soil. Furthermore, changes in the global climate are expected to alter precipitation regimes and increase drought periods (Bernstein et al., 2007), specially in areas where primary forests were degraded or deforested (Laurance and Willianson, 2001).

There are several mechanisms that soil bacteria use to resist/recover from drought, including: accumulation of osmolytes within the cell (Schimel et al., 2007), production of exopolysaccharides (Chang et al., 2007; Ward et al., 2009), differentiation into dormant life forms such as cyst-like cells (Soina et al., 2004) and spores (Laubach and Rice, 1916). The ability of soil bacteria to cope with dry/rewet events is apparently conserved in some phyla like Actinobacteria, Chloroflex, and Verrucomicrobia (Placella et al., 2012). These differences among bacterial taxa may lead to competitive advantages for drought tolerant bacteria after a dry/rewet event that can result in transient changes in bacterial community structure after these events. The magnitude of these changes possibly depends on previous exposures to drought, which may select drought tolerant bacteria in the community, or change bacteria strategy to survive under low moisture conditions (Evans and Wallenstein, 2014).

Because land use intensification alters the soil moisture regime and imposes considerable changes in bacterial community structure in the Amazon region, we hypothesized that the ability of soil bacterial communities to cope with dry/rewet events varies among land uses. We also hypothesized that the same occurs for the processes that soil bacteria mediate. We tested these hypotheses by applying a dry/rewet event on microcosms containing undisturbed soil samples taken from four land uses (mechanized agriculture, pasture, secondary forest, and primary forest) in the eastern Amazon, and by analyzing the subsequent impacts on bacterial community structure and on the process of ammonia oxidation.

MATERIAL AND METHODS

Sampling

The sites used in this study are located in the Eastern Amazon, in the Belterra municipality (Pará state, Brazil). In this region, the mean annual temperature is 26 °C and mean annual pluviosity is 2,150 mm. Geographical coordinates of the area are given in **Figure 1**.

To assess how land use affects the ability of bacterial communities to resist/recover from drought, we selected four land uses as follows: well-preserved primary forest; secondary forest with 20 years of regeneration; pasture covered with *Brachiaria* sp. and used for cattle production; and mechanized agriculture mainly destined to soybean production.

The sampling was conducted in April 2013, during the rainy season. For each land use, one transect with 250 m was used with five sampling points equally spaced (50 m). These transects were sampled as part of the Rede Amazonia Sustentável project (Gardner et al. 2013). In each sampling point, 10 undisturbed soil samples were taken in pairs by manually inserting PVC cylinders with 10 cm of height and 8.5 cm of diameter in the soil (Figure 1). The resulting 200 samples were kept at room temperature and carefully transported to the Departament of Soil Science at the Federal University of Lavras (Lavras, Minas Gerais, Brazil).

Dry/Rewet Experiment

Microcosms were prepared by placing the PVC cylinders with soil (undisturbed samples) in plastic pots with a sterilized cotton fabric in the bottom. To allow the leachate to return to the soil, a cotton cord was used to connect the bottom of the soil cylinder to the bottom of the plastic pot used as support. They were kept in a greenhouse with uncontrolled temperature ranging from 20 to 29 °C.

For each pair of microcosms, one was randomly assigned to the dry/rewet (DR) treatment and did not receive any water addition for 60 days, what reduced their moisture content by about 20% (±2%). After that period, they were rehydrated and kept at 80% of the field capacity (0.50 $g_{water}.g_{soil}^{-1}$ for primary and secondary forest; 0.32 $g_{water}.g_{soil}^{-1}$ for pasture and mechanized agriculture) for the rest of the experiment. The microcosms assigned to the control treatment were kept at constant moisture content (80% field capacity) by weekly adding an amount of sterilized deionized water equivalent to the water lost by evaporation, as measured by the weight loss for each microcosm.

Potential of ammonia oxidation (PAO) and samples for molecular analysis

At 1, 5, 10, 20, and 40 days after the end of the DR treatment, one pair of microcosm (DR and control) was destructively analyzed for each sampling point in each land use, totalizing 40 microcosms at each time. Two soil samples were taken from each microcosm, one for the PAO analysis (15 g) and one for molecular analysis (30 g). The samples for molecular analysis were immediately frozen at -80°C and further lyophilized. The samples for the potential of ammonia oxidation were immediately used in the assay described by Norton & Stark (2011). Briefly, 7 g of fresh soil was suspended in 50 mL of a 1 mM NH_4SO_4 solution buffered with phosphate (1 mM PO_4^{2-} at pH 7.2) in an Erlenmeyer flask. The soil suspension was shaken at 200 rpm during 24 hours at 25 °C. After 2, 4, 22, and 24 hours of agitation, a 2 mL aliquot was taken without interrupting the shaking to keep the proportion soil:solution constant. These aliquots were centrifuged, filtered, and analyzed for nitrate/nitrite concentration by ion chromatography (ICS 1100, Dionex, USA). The POA for the first sampling after the end of the drought period could not be determined due to problems in the equipment.

Bacterial community

To assess the bacterial community structure, the V4 region of the 16S *rRNA* gene was sequenced using the Illumina MiSeq platform. DNA extraction from soil, polymerase chain reaction conditions, sequencing, and sequence analysis were performed as described in Paper 01.

Statistical Analysis

To model the potential of ammonia oxidation, a two-level hierarchical linear model was used Gelman and Hill (2007). In the first level, the rate of increase in nitrate concentration along time (2, 4, 22, and 24h) was estimated for each combination of the three factors studied (land use, time after rewetting, and treatment). In the second level, means for that rate in each combination of the three factors were estimated, as well as the variances for combinations of treatment and land use. The model was estimated using the no-U-turn Bayesian sampler (Hoffman and Gelman, 2014) in Stan 2.5 with uninformative priors. Four Markov chains with 10,000 iterations were used, and convergence for all parameters was checked using the Gelman–Rubin Rhat statistic (Gelman and Rubin, 1992). Treatment contrasts were calculated using samples from the posterior distribution as outlined in Gelman and Hill (2007).

The effect of DR treatment on bacterial community structure at the OTU level was tested separately for each land use with Principal Response Curves (den Brink and Ter Braak, 1999) after removing the effect of sampling site using the residuals of linear models as outlined in Borcard et al. (2011). To test the effect of DR treatment on the (log-transformed) relative abundances of the dominant phyla, a linear mixed effect model was used with a random intercept for sampling point and time after rewetting. The significance of the DR treatment was tested using log-likelihood ratio tests (Zuur et al., 2009). To test for differences in the bacterial community structure among sampling points

within land use, a permutational multivariate analysis of variance (Permanova) (Anderson, 2001) was performed on the distance matrix calculated from the hellinger-transformed OTU table, only for the control treatment.

RESULTS AND DISCUSSION

This is the first report of the tolerance of soil bacterial communities to a dry/rewet event in the Amazon region. In general, regardless of the land use, the soil bacterial communities were highly stable and did not respond to the dry/rewet (DR) treatment at any time for all evaluated parameters.

To asses bacterial community structure, we sequenced the V4 region of the 16S *rRNA* gene using the Illumina MiSeq platform. In total, about 8 millions good quality sequences were obtained for all 200 samples. From these, 650,791 occurred only once (singletons) and were removed because they may be the result of sequencing errors and can inflate diversity measures (Huse et al., 2010). Using a similarity threshold of 97%, these sequences were clustered into 11,891 operational taxonomic units (OTU's). In agreement with the results presented in Paper 01, the observed OTU number was lower in primary forest than in any other land use ($P < 0.001$), but it was not affected by the DR treatment ($P > 0.05$; likelihood ratio tests) (**Figure 2**). Similarly, the dominant phyla were contrasting among land uses, but their relative abundances did not change after DR treatment ($P > 0.05$; likelihood ratio tests) (**Figure 3**).

The principal response curves analysis of the community structure at the OTU level also indicated no effect of DR ($P > 0.05$) for all land uses. The nonmetric multidimensional scaling ordinations for each land use confirm those results by showing that, for each pair of microcosms, the separation between DR and control treatments are not higher than the natural variability in those communities (**Figure 4**). In addition, the bacterial communities are spatially structured and differ significantly among sampling points as indicated by Permanova ($P < 0.001$). Therefore, if there were an undetected effect of DR on the community structure, it would be smaller than the spatial variability.

The ability of bacteria to recover from dry/rewet stresses seems to be phylogenetically conserved (Placella et al., 2012). For instance, Actinobacteria, Chloroflexi, and Verrucomicrobia are expected to recover their activity after rewetting relatively quickly (4-24 hours), whereas Proteobacteria and Acidobacteria may respond slowly and recover their activity only 72 hours after rewetting. These differences can give competitive advantage to fast responders and cause a transient change on the community structure following a dry/rewet event.

On the other hand, a recent study showed that bacterial taxa can shift their strategy to respond to dry/rewet events when they are previously exposed to drought (Evans et al., 2014). Therefore, previous exposure to low soil water availability may not only select drought tolerant bacteria but also induce drought sensitive bacteria to invest C and N resources in exopolysaccharides or protectant osmolytes (Schimel et al., 2007). This effect may reduce differences in competitive ability among bacteria after dry/rewet events and is a possible reason for the stability of bacterial community structure to these events. Similar increases in soil microbial community stability to drought after previous exposures have been recently reported (De Vries et al., 2012; Evans and Wallenstein, 2012; Göransson et al., 2013). That is a likely explanation for the bacterial community stability observed in our study because the region from where samples were taken have a tropical monsoon climate (*Am* – Koppen's classification) with a dry season lasting 4-5 months (Guimberteau et al., 2014).

Another recent work using next-generation sequencing analysis on 16S *rRNA* gene and transcripts revealed that bacterial community structure can be completely insensitive to extreme dry/rewet events even though bacterial activity can be strongly affected (Barnard et al., 2013). These authors detected strong changes in the relative abundances of RNA transcripts for some bacterial phyla during dry-down, which completely disappeared 2 hours after soil

rewetting. Because the transcript relative abundances returned to pre-drought levels in such a short time, this recovery could not be attributed to bacterial growth (increase in the number of cells), but rather indicated that the bacterial community structure was indeed stable as shown by their DNA-based analysis. Other studies using PLFA (Gordon et al., 2008) or DNA-based techniques (Fierer et al., 2003; Griffiths et al., 2003; Kaisermann et al., 2015) also support the stability of bacteria to variations in soil moisture.

With regard to the potential of ammonia oxidation (PAO), although an effect of land use was observed, with higher values of this variable in primary forest than in any other land use, the DR treatment did not affect the PAO in any land use, indicating that this process is highly stable in these soils (**Figure 5**). Because we only started to measure POA 5 days after rewetting, we can not know whether the process was unchanged by the DR treatment or if it recovered in that period. Previous works reported that the impact of dry/rewet events on nitrification potential or on ammonia oxidizers depends on historical soil moisture regimes (Fierer and Schimel, 2002; Placella and Firestone, 2013; Thion and Prosser, 2014)). Therefore, the lack of response of the POA to DR treatment in our study is possibly due to previous drought events in these soils. The higher nitrification potential in primary forest than in other land uses is consistent with previous reports of decreased abundance of *amoA* gene (Paula et al. 2014) and decreased nitrification rates (Neil at al., 1995) after conversion of Amazonian rainforests to pastures.

REFERENCES

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral ecology 26, 32-46.

Arevalo, L.A., Alegre, J.C., Bandy, D.E., Szott, L.T., 1998. The effect of cattle grazing on soil physical and chemical properties in a silvopastoral system in the Peruvian Amazon. Agroforestry Systems 40, 109-124.

Barnard, R.L., Osborne, C.A., Firestone, M.K., 2013. Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. The ISME Journal 7, 2229-2241.

Bernstein, L., Bosch, P., Canziani, O., Chen, Z., Christ, R., Davidson, O., 2007. Climate change 2007: synthesis report. Summary for policymakers. IPCC, Geneva, Switzerland.

Borcard, D., Gillet, F., Legendre, P., 2011. Numerical ecology with R. Springer Science & Business Media, New York.

Chang, W.S., van de Mortel, M., Nielsen, L., Nino de Guzman, G., Li, X., Halverson, L.J., 2007. Alginate Production by Pseudomonas putida Creates a Hydrated Microenvironment and Contributes to Biofilm Architecture and Stress Tolerance under Water-Limiting Conditions. Journal of Bacteriology 189, 8290-8299.

Comte, I., Davidson, R., Lucotte, M., de Carvalho, C.J.R., de Assis Oliveira, F., da Silva, B.P., Rousseau, G.X., 2012. Physicochemical properties of soils in the Brazilian Amazon following fire-free land preparation and slash-and-burn practices. Agriculture, Ecosystems and Environment 156, 108-115.

De Vries, F., Liiri, M., Bjornlund, L., Bowker, M., Christensen, S., Setala, H., Bardgett, R., 2012. Land use alters the resistance and resilience of soil food webs to drought. Nature Climate Change 2, 276-280.

den Brink, P.J., Ter Braak, C.J.F., 1999. Principal response curves: Analysis of time-dependent multivariate responses of biological community to stress. Environmental Toxicology and Chemistry 18, 138-148.

Evans, S.E., Wallenstein, M.D., 2012. Soil microbial community response to

drying and rewetting stress: Does historical precipitation regime matter? Biogeochemistry 109, 101-116.

Evans, S.E., Wallenstein, M.D., 2014. Climate change alters ecological strategies of soil bacteria. Ecology Letters 17, 155-164.

Falkowski, P.G., Fenchel, T., Delong, E.F., 2008. The microbial engines that drive Earth's biogeochemical cycles. Science 320, 1034-1039.

Fierer, N., Schimel, J.P., 2002. Effects of drying-rewetting frequency on soil carbon and nitrogen transformations. Soil Biology and Biochemistry 34, 777-787.

Fierer, N., Schimel, J.P., Holden, P.A., 2003. Influence of drying-rewetting frequency on soil bacterial community structure. Microbial Ecology 45, 63-71.

Fuchslueger, L., Kastl, E.M., Bauer, F., Kienzl, S., Hasibeder, R., Ladreiter-Knauss, T., Schmitt, M., Bahn, M., Schloter, M., Richter, a., Szukics, U., 2014. Effects of drought on nitrogen turnover and abundances of ammonia-oxidizers in mountain grassland. Biogeosciences Discussions 11, 9183-9214.

Gelman, A., Hill, J., 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York.

Gelman, A., Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. Statistical Science 7, 457-472.

Göransson, H., Godbold, D.L., Jones, D.L., Rousk, J., 2013. Bacterial growth and respiration responses upon rewetting dry forest soils: Impact of drought-legacy. Soil Biology and Biochemistry 57, 477-486.

Gordon, H., Haygarth, P.M., Bardgett, R.D., 2008. Drying and rewetting effects on soil microbial community composition and nutrient leaching. Soil Biology and Biochemistry 40, 302-311.

Griffiths, R.I., Whiteley, A.S., Anthony, G., Donnell, O., Bailey, M.J., Donnell, A.G.O., 2003. Physiological and Community Responses of Established Grassland Bacterial Populations to Water Stress Physiological and Community Responses of Established Grassland Bacterial Populations to Water Stress. Applied and Environmental Microbiology 69, 6961-6968.

Guimberteau, M., Ducharne, a., Ciais, P., Boisier, J.P., Peng, S., De Weirdt, M., Verbeeck, H., 2014. Testing conceptual and physically based soil hydrology schemes against observations for the Amazon Basin. Geoscientific Model Development 7, 1115-1136.

Hoffman, M.D., Gelman, A., 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research 15, 30-30.

Huete, A.R., Didan, K., Shimabukuro, Y.E., Ratana, P., Saleska, S.R., Hutyra, L.R., Yang, W., Nemani, R.R., Myneni, R., 2006. Amazon rainforests green-up with sunlight in dry season. Geophysical Research Letters 33, 2-5.

Huse, S.M., Welch, D.M., Morrison, H.G., Sogin, M.L., 2010. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology 12, 1889-1898.

Jesus, E.C., Marsh, T.L., Tiedje, J.M., Moreira, F.M.S., 2009. Changes in land use alter the structure of bacterial communities in Western Amazon soils. The ISME Journal 3, 1004-1011.

Kaisermann, A., Maron, P.A., Beaumelle, L., Lata, J.C., 2015. Fungal communities are more sensitive indicators to non-extreme soil moisture variations than bacterial communities. Applied Soil Ecology 86, 158-164.

Lal, R., 1989. Agroforestry systems and soil surface management of a tropical alfisoh. Agroforestry Systems 8, 197-215.

Lal, R., Wilson, G.F., Okigbo, B.N., 1979. Changes in properties of an Alfisol produced by various crop covers. Soil Science 127, 377-382.

Laubach, C.A., Rice, J.L., 1916. Spore-Bearing Bacteria in Soil. Journal of Bacteriology 1, 513-518.

Laurance, W.F., Bruce Williamson, G., 2001. Positive feedbacks among forest fragmentation, drought, and climate change in the Amazon. Conservation Biology 15, 1529-1535.

Maeda, E.E., Formaggio, A.R., Shimabukuro, Y.E., 2008. Impacts of Land Use and Land Cover Changes on Sediment Yield in a Brazilian Amazon Drainage Basin. Giscience & Remote Sensing 45, 443-453.

Martins, P.C.C., Dias Junior, M.D.S., Ajayi, A.E., Moreira, F.M.D.S., 2012. Structural sustainability of cambisol under different land use system. Revista Brasileira de Ciência do Solo 36, 1724-1732.

Neill, C., Piccolo, M.C., Steudler, P.A., Melillo, J.M., Feigl, B.J., Cerri, C.C., 1995. Nitrogen dynamics in soils of forests and active pastures in the western Brazilian Amazon Basin. Soil Biology and Biochemistry 27, 1167-1175.

Norton, J.M., Stark, J.M., 2011. Regulation and Measurement of Nitrification in Terrestrial Systems, In: Abelson, J.N., Simon, M.I. (Eds.). Academic Press, San Diego, CA, pp. 343-368.

Paula, F.S., Rodrigues, J.L.M., Zhou, J., Wu, L., Mueller, R.C., Mirza, B.S., Bohannan, B.J.M., Nüsslein, K., Deng, Y., Tiedje, J.M., Pellizari, V.H., 2014. Land use change alters functional gene diversity, composition and abundance in Amazon forest soil microbial communities. Molecular ecology, 2988-2999.

Placella, S.A., Brodie, E.L., Firestone, M.K., 2012. Rainfall-induced carbon dioxide pulses result from sequential resuscitation of phylogenetically clustered microbial groups. Proceedings of the National Academy of Sciences 109, 10931-10936.

Placella, S.a., Firestone, M.K., 2013. Transcriptional response of nitrifying communities to wetting of dry soil. Applied and Environmental Microbiology 79, 3294-3302.

Schimel, J., Balser, T.C., Wallenstein, M., 2007. Microbial stress-response physiology and its implications for ecosystem function. Ecology 88, 1386-1394.

Soina, V.S., Mulyukin, A.L., Demkina, E.V., Vorobyova, E.A., El-Registan, G.I., 2004. The structure of resting bacterial populations in soil and subsoil permafrost. Astrobiology 4, 345-358.

Stan Development, T., 2014. Stan Modeling Language Users Guide and Reference Manual, Version 2.5.0.

Thion, C., Prosser, J.I., 2014. Differential response of non-adapted ammonia oxidising archaea and bacteria to drying rewetting stress. FEMS Microbiology Ecology, 1-10.

Ward, N.L., Challacombe, J.F., Janssen, P.H., Henrissat, B., Coutinho, P.M., Wu,

M., Xie, G., Haft, D.H., Sait, M., Badger, J., Barabote, R.D., Bradley, B., Brettin, T.S., Brinkac, L.M., Bruce, D., Creasy, T., Daugherty, S.C., Davidsen, T.M., DeBoy, R.T., Detter, J.C., Dodson, R.J., Durkin, A.S., Ganapathy, A., Gwinn-Giglio, M., Han, C.S., Khouri, H., Kiss, H., Kothari, S.P., Madupu, R., Nelson, K.E., Nelson, W.C., Paulsen, I., Penn, K., Ren, Q., Rosovitz, M.J., Selengut, J.D., Shrivastava, S., Sullivan, S.A., Tapia, R., Thompson, L.S., Watkins, K.L., Yang, Q., Yu, C., Zafar, N., Zhou, L., Kuske, C.R., 2009. Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. Applied and Environmental Microbiology 75, 2046-2056.

Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., 2009. Mixed effects models and extensions in ecology with R. Springer, New York.

LIST OF FIGURES

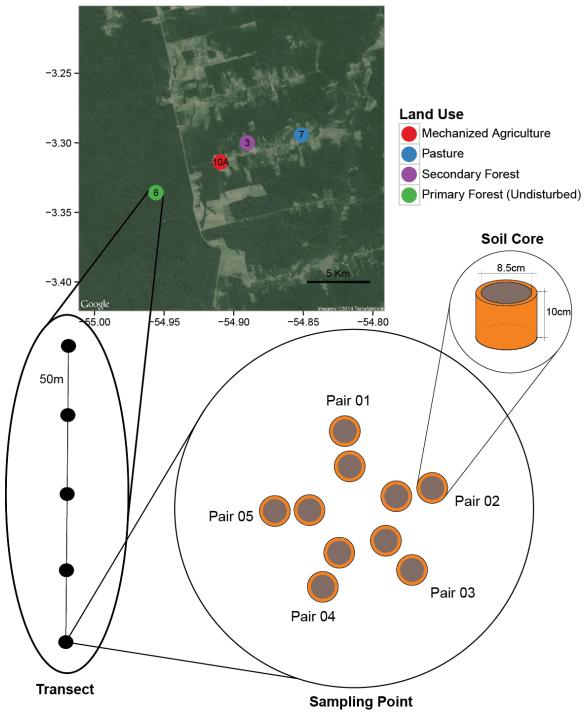

Figure 1. *Sampling scheme*. Each point in the satellite image represents a transect, from which five sampling points equally spaced (50m) were used. In each sampling point, 5 pairs of soil cores (diameter 8cm, height 10cm) were taken. The land use for each transect is indicated in the legend at right.

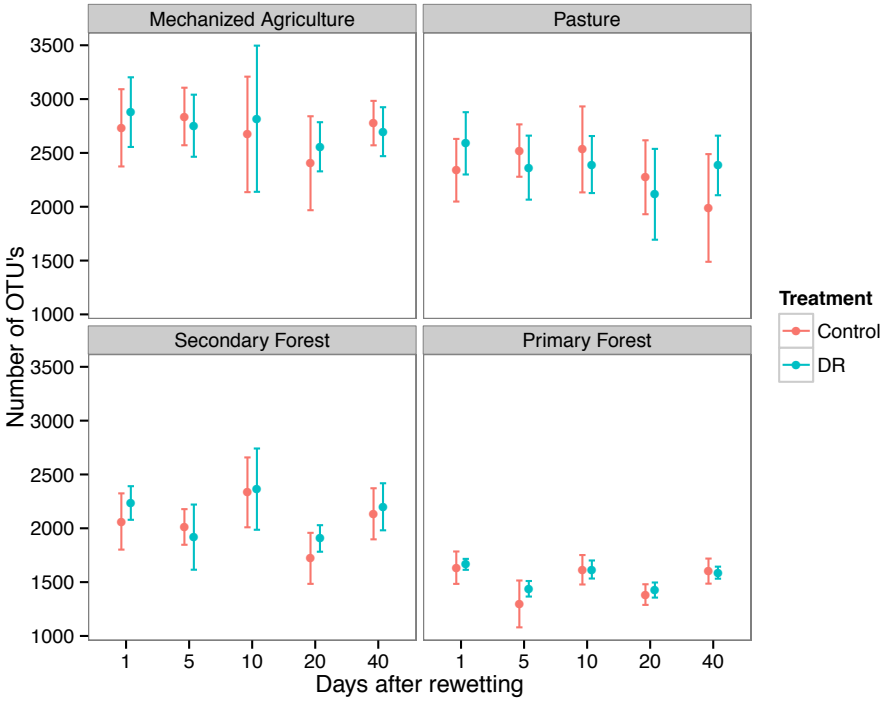

Figure 2. Effects of dry/rewet (DR) on the number of operational taxonomy units (OTU's) in Amazonian soils under four land uses. Points represent the mean (n = 5) and bars indicate the 95% confidence interval. Treatments are indicated by colors according to the legend at right.

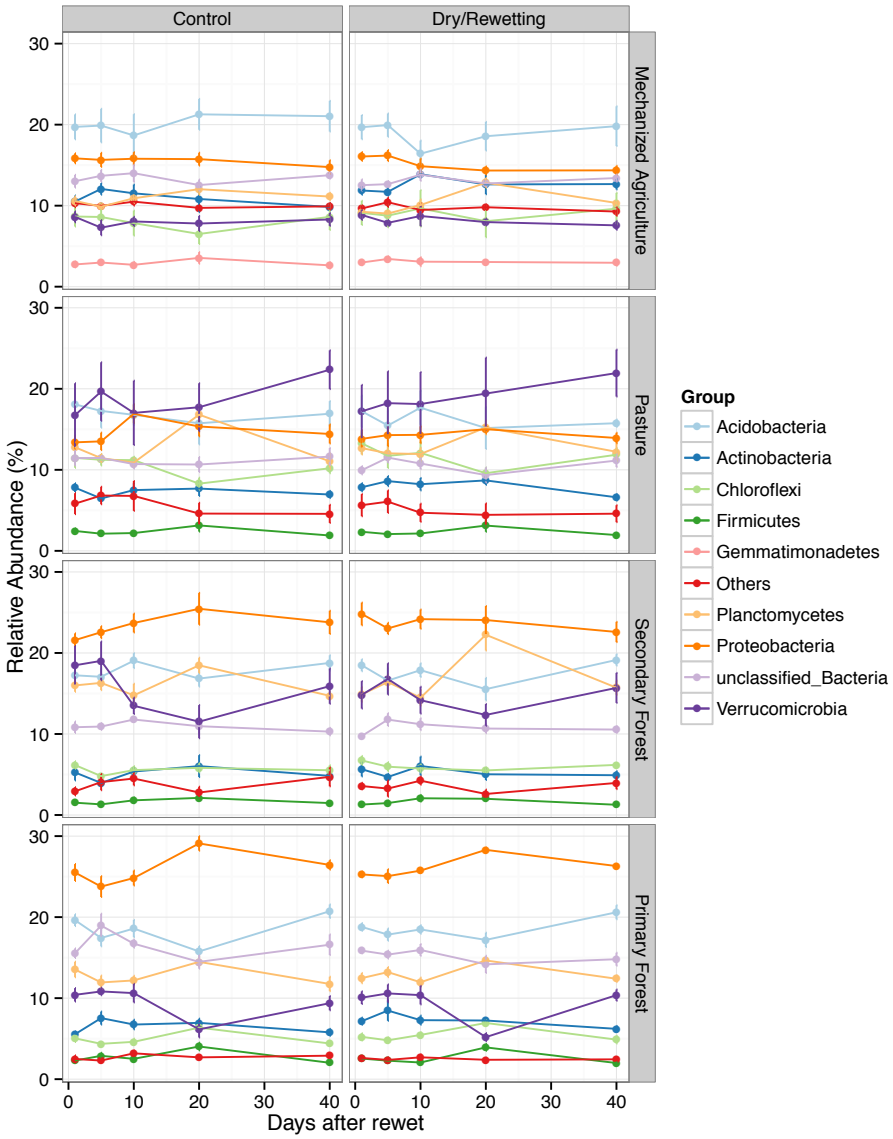

Figure 3. Effects of dry/rewet treatment on the relative abundances of dominant phyla of bacteria in Amazonian soils under four land uses. Points represent the mean (n = 5) and bars indicate the standard error. Treatments are indicated by colors according to the legend at right. Colors represent the phyla according to the legend at right.

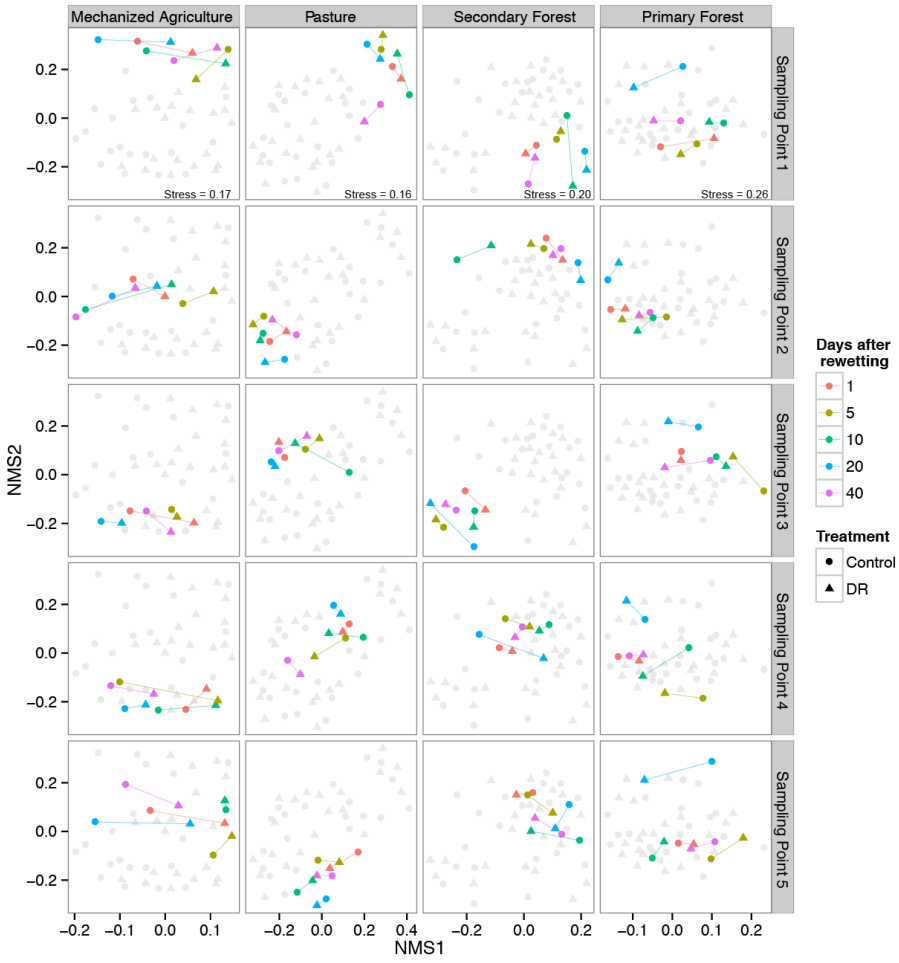

Figure 4. Nonmetric multidimensional scaling ordinations of bacterial communities from Amazonian soils in a gradient of land use intensity treated (DR) or not (control) with a dry/rewet event. The ordination for each land use (columns) is divided into the five sampling points (rows). In each panel, samples belonging to the respective point are colored according to the sampling time (days after rewetting) as indicated in the legend at right, whereas those belonging to other sampling points are colored in grey. The shape of the points indicates the treatment according to the legend at right. Samples corresponding to each pair of control and DR treatments are connected by a segment.

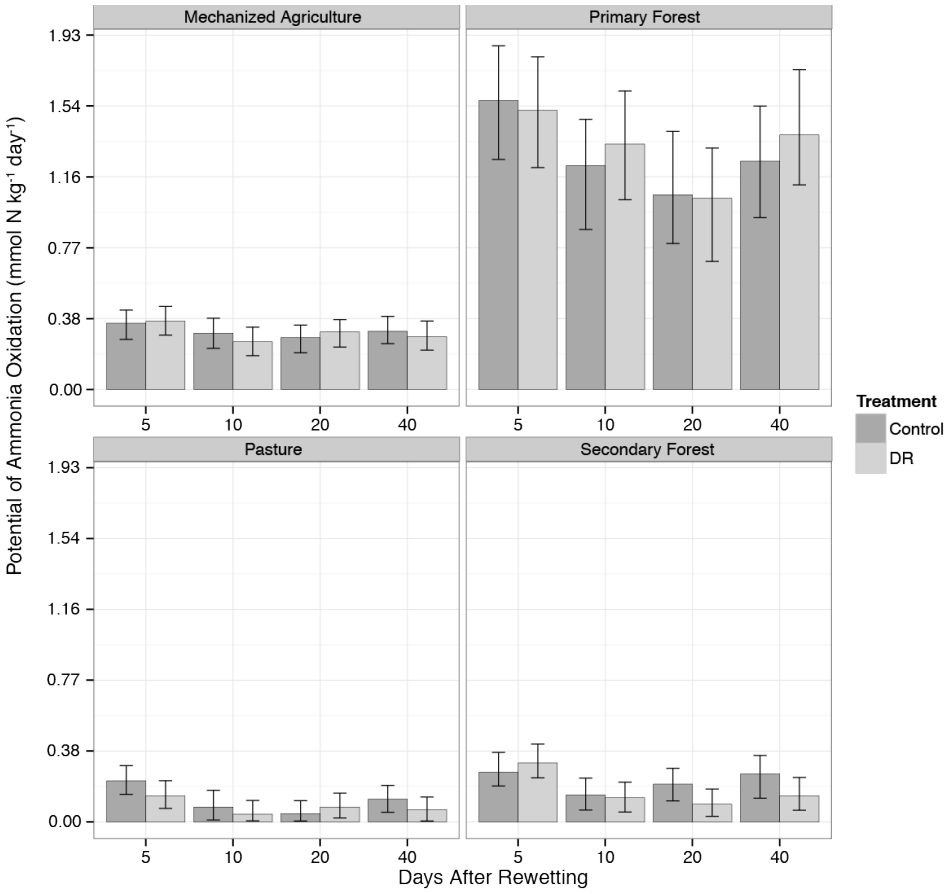

Figure 5. Potential of ammonia oxidation for Amazonian soils, under four land uses, submitted to dry/rewet treatment (DR) or kept at constant soil moisture (control). The column height indicates the mean (n = 5) and the bars indicate the 95% credible interval.

PAPER 3

Standards of the journal - Bioinformatics

A new approach for designing primers for functional genes with emphasis on gene-targeted metagenomics

ABSTRACT

Motivation: Design of primers is a critical step to adequately assess the diversity of specific components of soil diversity using gene-targeted metagenomics. Although some databases with thousands of sequences for genes of functional relevance are currently available, tools that easily allow their usage for primer design are still limited because they can not handle large datasets, are very computationally intensive, do not use a flexible approach to design degenerate primers, and/or do not allow primer design on specific regions of a multiple alignment of sequences to ensure a minimum overlap between amplicons.

Results: We present a fast, thermodynamics-based approach to select primers from a multiple alignment of sequences, which is specifically suited for designing primers for gene-targeted metagenomics. Instead of designing degenerate primers, as in most commonly used tools, this algorithm uses a more flexible approach based on the selection of two sets of thermodynamically compatible forward/reverse sub-primers that can be multiplexed in a single polymerase chain reaction (PCR). We tested this tool on the archaeal *amoA* gene (encoding for the alpha-subunit of ammonia monooxygenase) and obtained a set of primers that are predicted to

specifically amplify sequences from the *Nitrososphaera* subgroup. When tested with DNA extracted from Amazonian soils, the designed primers allowed the amplification of fragments with the predicted length, indicating their suitability for PCR in difficult samples, like soil.

Availability: The source code is freely available, under the MIT license at www.github.com/teodecarvalho/PrimerDesign/blob/master/functions_definition.py
Instructions and examples of usage are available at www.github.com/teodecarvalho/PrimerDesign

INTRODUCTION

Gene-targeted metagenomics offers the possibility of studying specific components of microbial diversity at an unprecedented scale. It is especially useful for studying genes of functional relevance because of its higher sensitiveness compared to other metagenomic approaches like shotgun-metagenomics (Iwai et al., 2010). However, because polymerase chain reaction (PCR) is used in this technique, it can be highly biased toward the sequences used to design primers. Consequently, when used for studying the ecology of functional groups of microorganisms, especially in soil, gene-targeted metagenomics is strongly dependent on primer design to capture the diversity of these genes.

Many software have been proposed for primer design, but their application for gene-target metagenomics is still limited because they can not handle large databases (Kaderali et al., 2002; Li et al., 2008; Gervais et al., 2010; Untergasser et al., 2012); are very computationally intensive and do not allow specification of positions for primer enumeration (Gans et al., 2012); or can only design degenerate primers in the sense of Linhart and Shamir (2002), where a set of sub-primers are combined into one degenerate primer for which only some bases are allowed to vary (Figure 1), restricting the primer design to conserved regions (Boyce et al., 2009; Najafabadi et al., 2008; Brodin et al., 2013; Yoon and Leitner, 2014).

Here we present an approach that helps to circumvent these limitations by enumerating sub-primers along regions of a multiple alignment without trying to combine them into a single degenerate primer. Instead, this tool allows the sub-primers to be enumerated from several positions and to vary in length.

ALGORITHM AND IMPLEMENTATION

This tool was primarily written in the python programming language using IPython (Perez and Granger, 2007), with minor portions of the code written in R (R Development Core Team, 2014). It runs in the IPython Notebook and is compatible with Unix/Linux systems. It is divided in three main sections as detailed below.

Section 01: screening of the alignment

In this section, which requires a FASTA file containing a set of aligned nucleotide sequences, a screening is performed to evaluate regions of the alignment that are suitable for the enumeration of compatible forward and reverse primers, taking into account the coverage (number of sequences detected) and predicted melting temperatures (T_m).

At each position (column) of the alignment, all valid primers (according to a set of user-defined constraints as illustrated in **Table S1**) are enumerated, and a greedy algorithm is performed to select the optimal set of primers (up to the maximal degeneracy allowed by the user). After that, the proportion of sequences detected (coverage), the number of valid primers, and the range of T_m is calculated for the selected set of primers. This process is repeated in parallel using multiprocessing for every position in the alignment, and the final result is plotted as in **Figure S1 and S2**, allowing for a visual inspection of the alignment.

Nearest-neighbor thermodynamic models are used to calculate the predicted T_m of candidate primers, primer-primer interactions (hetero and homodimers), and the stability of potential hairpin structures (**Figure 2**), according to the parameters suggested in SantaLucia and Hicks (2004) using

primer3 (Untergasser et al., 2012). These models have higher accuracy than those based only on nucleotide composition because they take into account the actual stacking of neighboring base pairs in the sequence (SantaLucia, 1998).

Section 02: primer enumeration and selection

Based on the results of the first section, the user must choose two regions of the alignment that will be used for forward and reverse primer design. The main steps performed in this section are as follows:

1. For every position of the two chosen regions (forward and reverse) in the alignment, all possible candidate primers are enumerated (**Figure 3**) in parallel using multiprocessing, and filtered based on user-defined constraints as those illustrated in **Table S2**.
2. To reduce the number of candidate sub-primers to be tested, they are combined into groups of redundant candidates (those occurring in the same sequences in the alignment).
3. A list containing one representative candidate for each group from the previous step is formed for forward and reverse candidate primers.
4. They are further combined in all possible forward/reverse candidate pairs.
5. A greedy algorithm is performed to select the combination of compatible pairs that maximizes the coverage.

The outcome of these steps is a set of compatible candidate forward and reverse primers that can be combined, behaving as a degenerate primer, except that the sub-primers do not necessarily come from the same position in the alignment and do not necessarily have the same length as illustrated in Figure 1.

Section 03: test of the primers

In the final section, the specificity of the set of primers selected in section 02 is assessed by *in silico* polymerase chain reaction (PCR) using MFEprimer-2.0 (Qu

et al., 2012). At this point, the user can check the selected primers against target or non-target sequences provided in a FASTA file. By evaluating the primers against target sequences, it is possible to further reduce the total number of selected sub-primers because some redundancy (detection of the same sequences by different primers) will arise due to mismatches with low impact on T_m. If non-target sequences are amplified, the user can go back to step 02 and choose a different region of the alignment or change the constraints for the selection of a new set of primers.

TEST OF THE TOOL

Database

The archaeal *amoA* gene (encoding for the alpha-subunit of ammonia monooxygenase) was chosen to illustrate the usage of this tool because of the functional relevance of ammonia-oxidizing archaea Pester et al. (2012).

A total of 21,663 nucleotide sequences classified as archaeal *amoA*, and their respective amino acid sequences, were retrieved from the fungene repository (Fish et al., 2013) using a threshold of 400 for the hidden markov model (HMM) score. The amino acid sequences were aligned using hmmalign from HMMER (Eddy, 2009) with the HMM profile available in the repository. Nucleotide sequences were dereplicated and aligned based on the protein alignment with in-house Python scripts, resulting in 15,400 unique sequences.

These nucleotide sequences were classified into five major groups of ammonia-oxidizing archaea (*Nitrosopumilus*, *Nitrosotalea*, *Nitrosocaldus*, *Nitrososphaera*, and *Nitrososphaera* sister cluster) by matching them against the database provided by Pester et al. (2012) using a similarity threshold of 97% in USEARCH (Edgar, 2013). In total, 96% of all sequences were unambiguously classified into those five groups.

Primers were designed only for the *Nitrososphaera* group, containing 6,824 unique sequences in this database, whereas those sequences classified in other groups (8,576 unique sequences) were used to test the specificity of the primers.

All tests were performed in a desktop computer with a quad-core i7 2.6 GHz processor and 16 Gb of RAM.

Polymerase chain reaction

The designed primers were further tested on DNA extracted from soils collected in the eastern Amazon (*see* Paper 01 for sampling details). The PCR was performed on 20 μL solutions containing 2 μL of a equimolar (10 μM) mixture of forward/reverse primers, 10 μL of KAPA2G Robust (Kapa Biosystems) mastermix, 1 μL of the template, and 7 μL of PCR grade water. The PCR conditions were as follows: initial denaturation at 95 ˚C for 3 min; 40 cycles of denaturation at 95 ˚C for 15 s, annealing at 60 ˚C for 15 s, and extension at 72 ˚C for 15 s. A final extension at 72 ˚C for 10 min was also performed.

RESULTS AND DISCUSSION

Using this tool, we were able to design a set of primers for the archaeal *amoA* gene that were specific for the *Nitrososphaera* group. The time required to complete the analysis, not including database preparation, was 20 min. The designed primers are predicted to amplify fragments (amplicons) with about 235-280 base pairs (bp); and the amplicons are predicted to have a minimum overlap of 235 bp based on the multiple alignment. This overlap is relevant for sequencing because it allows comparisons among amplicons, which is necessary for clustering them into operational taxonomic units (Huse et al., 2010).

In total, 12 forward and 13 reverse sub-primers were selected, being equivalent to two degenerate primers with 12 and 13-fold degeneracy, respectively. However, in contrast to degenerate primers in the sense of Linhart and Shamir (2002) (Figure 1), these primers produce amplicons with varied length depending on the combination of forward/reverse sub-primer.

Considering only the nucleotide identities, they are predicted to detect 82% (allowing 1 mismatch) and 96% (allowing 2 mismatches) of all sequences available in the database used for primer design (**Figure S3**). When considering thermodynamic parameters, these primers are expected to detect 70% of these sequences. In those simulations, they did not amplify any *amoA* sequences from other ammonia-oxidizing archaea groups, indicating high specificity of these primers for the *Nitrososphaera* group. Unfortunately, we can not assess the coverage of the most commonly used set of primers for archaeal *amoA* (Francis et al., 2005) because the positions where these primers are expected to anneal are missing for most sequences available in the fungene database.

In the PCR performed on DNA extracted from Amazonian soils, the designed primers allowed the amplification of fragments with the predicted sample size, with no sign of nonspecific amplification (**Figure 4**). Polymerase

chain reaction on DNA extracted from soil is particularly difficult because of the high microbial diversity in soil, which increases the opportunity for loss of specificity (Gans et al., 2012), and because of the presence of PCR inhibitors that reduces the efficiency of the reaction (Vishnivetskaya et al., 2014).

LIMITATIONS AND FUTURE WORK

The two most important limitations of the current implementation of this approach are as follows:

Specificity is only checked after the primers are selected. This is potentially a problem for closely related genes. Extensions of this algorithm could allow the user to specify a set of (previously tested) primers to be excluded during the selection step. Another approach would be the inclusion of specificity checking during primer selection, but this would considerably increase the processing time.

Cost of primer synthesis. Despite the higher flexibility of the enumeration method, the resulting primers usually can not be combined into single degenerate primers, what increases their synthesis cost. Modifications of this algorithm during the selection step could favor small sets of sub-primers differing in only a few bases so that they could be combined into several degenerate primers with 2 or 3-fold degeneracy.

CONCLUSIONS

We presented a new approach to designing primers for functional genes, which is especially suited for gene-targeted metagenomics. When tested on archaeal *amoA* gene fragments, this tool provided a fast method to design primers that: have high predicted coverage; have desirable thermodynamic properties; produce overlapping amplicons; and are suitable for complex samples like DNA extracted from soil.

REFERENCES

Boyce, R., Chilana, P., Rose, T.M., 2009. iCODEHOP: a new interactive program for designing COnsensus-DEgenerate Hybrid Oligonucleotide Primers from multiply aligned protein sequences. Nucleic Acids Research 37, 222-228.

Brodin, J., Krishnamoorthy, M., Athreya, G., Fischer, W., Hraber, P., Gleasner, C., Green, L., Korber, B., Leitner, T., 2013. A multiple-alignment based primer design algorithm for genetically highly variable DNA targets. BMC Bioinformatics 14, 255-255.

Eddy, S.R., 2009. A new generation of homology search tools based on probabilistic inference. Genome Informatics 23, 205-211.

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature methods 10, 996-998.

Fish, J.A., Chai, B., Wang, Q., Sun, Y., Brown, C.T., Tiedje, J.M., Cole, J.R., 2013. FunGene: the functional gene pipeline and repository. Frontiers in Microbiology 4.

Gans, J.D., Dunbar, J., Eichorst, S.A., Gallegos-Graves, L.V., Wolinsky, M., Kuske, C.R., 2012. A robust PCR primer design platform applied to the detection of Acidobacteria Group 1 in soil. Nucleic Acids Research 40, e96-e96.

Gervais, A.L., Marques, M., Gaudreau, L., 2010. PCRTiler: automated design of tiled and specific PCR primer pairs. Nucleic Acids Research 38, W308-312.

Huse, S.M., Welch, D.M., Morrison, H.G., Sogin, M.L., 2010. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology 12, 1889-1898.

Iwai, S., Chai, B., Sul, W.J., Cole, J.R., Hashsham, S.A., Tiedje, J.M., 2009. Gene-targeted-metagenomics reveals extensive diversity of aromatic dioxygenase genes in the environment. ISME Journal 4, 279-285.

Kaderali, L., Schliep, A., 2002. Selecting signature oligonucleotides to identify organisms using DNA arrays. Bioinformatics 18, 1340-1349.

Li, K., Brownley, A., Stockwell, T.B., Beeson, K., McIntosh, T.C., Busam, D., Ferriera,

S., Murphy, S., Levy, S., 2008. Novel computational methods for increasing PCR primer design effectiveness in directed sequencing. BMC Bioinformatics 9, 191-191.

Linhart, C., Shamir, R., 2002. The degenerate primer design problem. Bioinformatics 18 Suppl 1, S172-S181.

Najafabadi, H.S., Torabi, N., Chamankhah, M., 2008. Designing multiple degenerate primers via consecutive pairwise alignments. BMC Bioinformatics 9, 55-55.

Perez, F., Granger, B.E., 2007. IPython: A System for Interactive Scientific Computing. Computing in Science and Engineering 9, 21-29.

Pester, M., Rattei, T., Flechl, S., Gröngröft, A., Richter, A., Overmann, J., Reinhold-Hurek, B., Loy, A., Wagner, M., 2012. amoA-based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of amoA genes from soils of four different geographic regions. Environmental Microbiology 14, 525-539.

Qu, W., Zhou, Y., Zhang, Y., Lu, Y., Wang, X., Zhao, D., Yang, Y., Zhang, C., 2012. MFEprimer-2.0: a fast thermodynamics-based program for checking PCR primer specificity. Nucleic Acids Research 40, W205-208.

SantaLucia, J., 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proceedings of the National Academy of Sciences 95, 1460-1465.

SantaLucia, J., Jr., Hicks, D., 2004. The thermodynamics of DNA structural motifs. Annual Review of Biophysics and Biomolecular Structure 33, 415-440.

Team, R.C., 2014. R: A Language and Environment for Statistical Computing, Vienna, Austria.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3-new capabilities and interfaces. Nucleic Acids Research 40, e115-e115.

Vishnivetskaya, T.A., Layton, A.C., Lau, M.C.Y., Chauhan, A., Cheng, K.R., Meyers, A.J., Murphy, J.R., Rogers, A.W., Saarunya, G.S., Williams, D.E., Pfiffner, S.M., Biggerstaff, J.P., Stackhouse, B.T., Phelps, T.J., Whyte, L., Sayler, G.S., Onstott, T.C., 2014. Commercial DNA extraction kits impact observed microbial community composition in permafrost samples. FEMS Microbiology Ecology 87, 217-230.

Yoon, H., Leitner, T., 2014. PrimerDesign-M: a multiple-alignment based multiple-primer design tool for walking across variable genomes. Bioinformatics.

LIST OF FIGURES

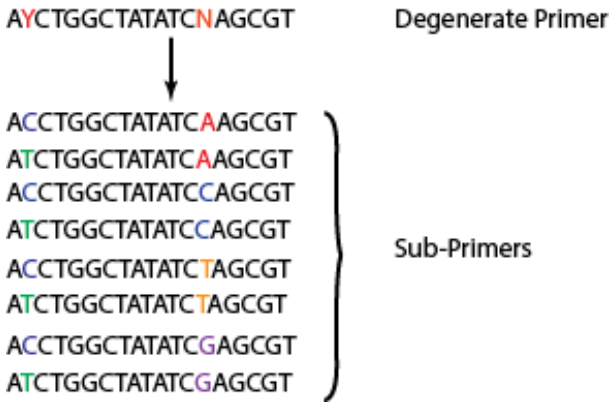

Figure 1. Hypothetical degenerate primer with 8-fold degeneracy. Degenerate bases are used to indicate possible variations in the positions 2 and 13. In this representation, the degenerate primer is composed by a equimolar mixture of all possible variants (sub-primers).

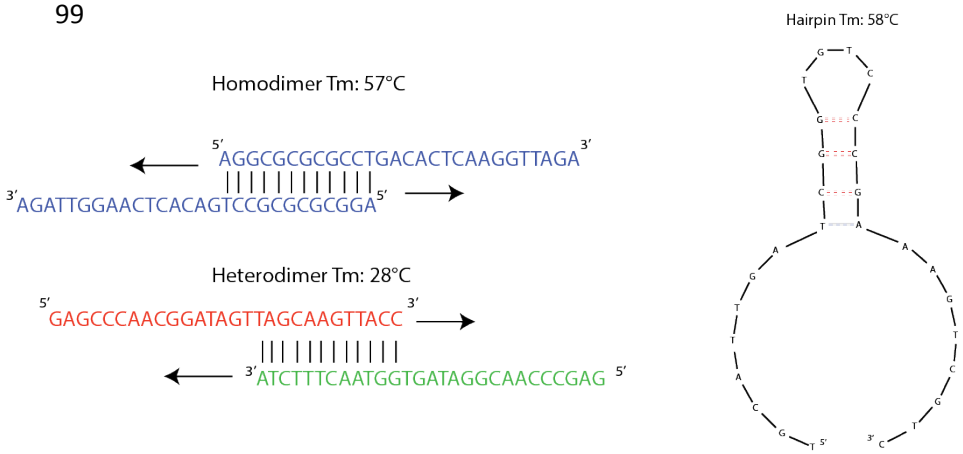

Figure 2. Examples of homodimer, heterodimer, and hairpin structures of oligonucleotides.

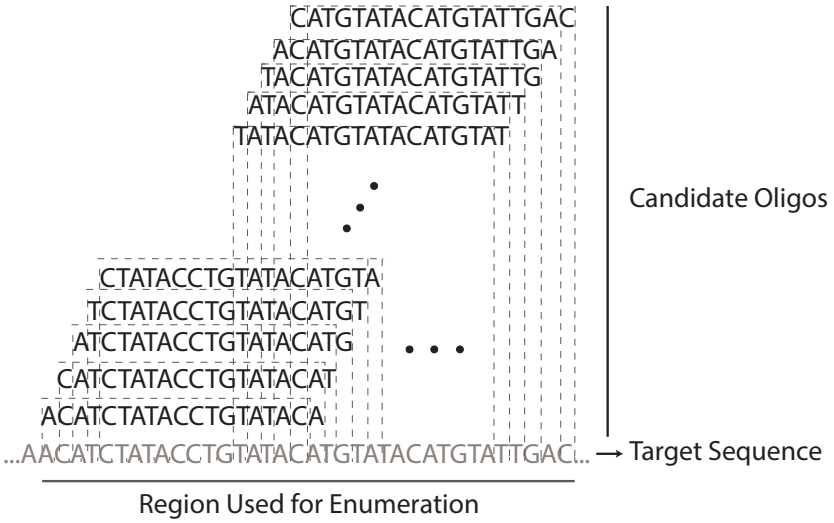

Figure 3. Illustration of the primer enumeration process. All possible oligonucleotides (oligos) are enumerated from each position of a given target sequence. For the Section 02 of the algorithm, this process is repeated for all sequences in the alignment and the length of the oligos varies according to a range specified by the user (usually 18-28 nucleotides).

Figure 4. Validation of the primers designed for archaeal *amoA* gene from *Nitrososphaera* group. Agarose gel (3%) stained with SYBR Green.

SUPPLEMENTARY MATERIAL

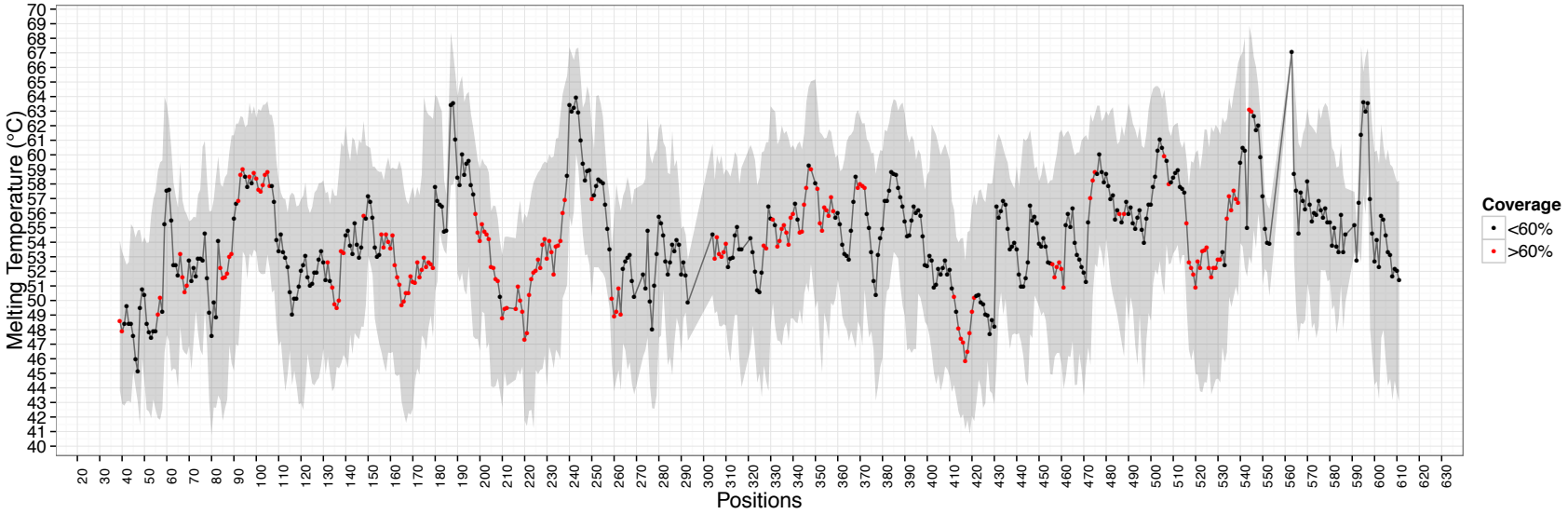

Figure S1. Predicted melting temperature for the selected set of sub-primers (up to 60) at each position of the alignment (fragments of archaeal *amoA* gene from *Nitrososphaera* group). The points indicate the median melting temperature at each position. Point colors indicate if the coverage is above (red) or below (black) 60% at each position. The shaded area represents the region between the 10th to the 90th percentile for the melting temperature of the best set of primer selected at each position.

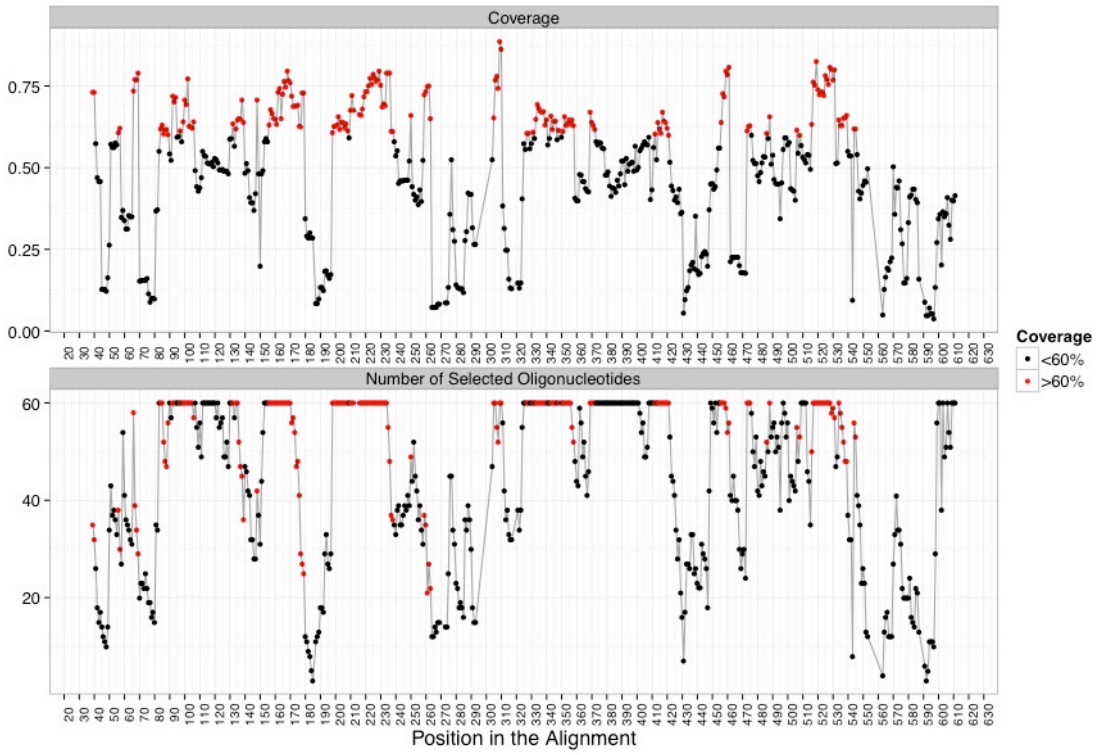

Figure S2. Proportion of sequences detected (coverage) and number of selected oligonucleotides (oligos) for each set of candidate primers selected along all positions in the alignment (fragments of archaeal *amoA* gene from *Nitrososphaera* group). These results depend on the set of constraints specified by the user (see Table S1).

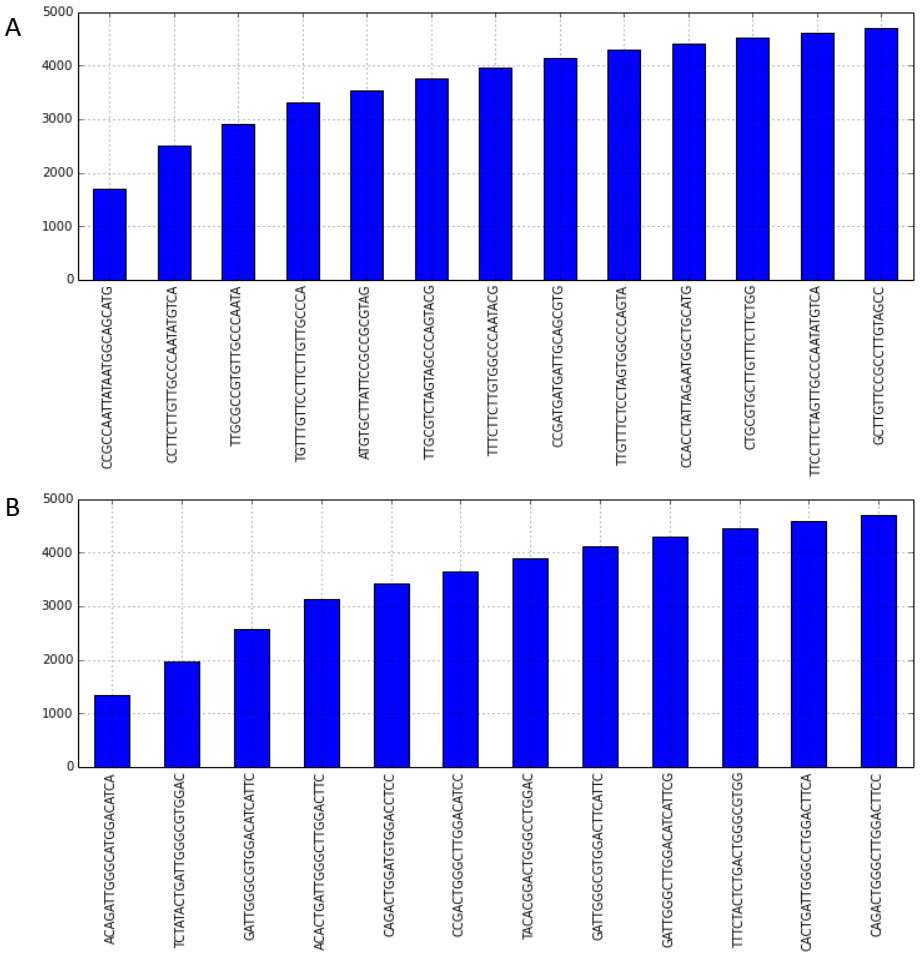

Figure S3. Cumulative number of sequences detected (out of 6,824 in total) for each forward (A) and reverse (B) oligonucleotide added to the selected set of sub-primers.

Table S1. Parameters for alignment screening in the Nitrososphaera test.

Parameter	**Value**
k-mer length	18-28
Step size	1
Maximal hairpin T_m (ºC)	35
Maximal homodimer T_m (ºC)	35
Minimal oligo T_m (ºC)	40
Maximal oligo T_m (ºC)	70
Minimal oligo occurrence	60
Maximal degeneracy	60
Total primer concentration (nM)	200
Monovalent ion concentration (mM)	50
Divalent ion concentration (mM)	1.5

Table S2. Parameters for primer enumeration in the Nitrososphaera test.

Parameter	**Value**
k-mer length	20-28
Maximal hairpin T_m (ºC)	35
Maximal homodimer T_m (ºC)	30
Maximal heterodimer T_m (ºC)	35
Maximal delta T_m (ºC)	5
Minimal oligo T_m (ºC)	55
Maximal oligo T_m (ºC)	60
Minimal oligo occurrence	50
Maximal degeneracy	60
Total primer concentration (nM)	200
Monovalent ion concentration (mM)	50
Divalent ion concentration (mM)	1.5
Forward region	90-105
Reverse region	340-370
No-3-T[a]	Yes
No-poly-run[b]	Yes
No-poly-3-GC[c]	Yes

[a]Three consecutives thymines are not allowed at the 3' end of the primer.

[b]Consecutive repetitions of four or more bases are not allowed.

[c]Repetitions of 3 guanines or cytosines are not allowed at the 3' end of the primer.

PAPER 4

Standards of the journal – Soil Biology and Biochemistry

Pitfalls in a commonly used index for measuring resistance and resilience of ecosystem processes in soil

ABSTRACT

Ecosystem stability and its dependence on biodiversity are currently among the main topics in environmental sciences. Here a commonly used index for measuring resistance and resilience in soil is explored with regard to some properties that can lead to erroneous conclusions about stability. We use Monte Carlo simulations to show that the index proposed for resistance is sensitive to the coefficient of variation (ratio of the underlying standard deviation to mean) of the process, potentially leading to differences in this index between systems, even when there is no disturbance effect. Similarly, the resilience index is sensitive to the standard deviation of the measured process, especially at lower levels of initial disturbance. As a result, the index may show differences in resilience for fully recovered systems if they differ in their resistance. Based on those observations, it is advisable that: (i) the resistance index should only be compared among systems with similar coefficient of variation in the ecosystem process of interest; (ii) it should only be interpreted based on the expected value under the null model (no disturbance effect); (iii) the resilience index should only be compared between systems with similar standard deviations and initial disturbance effects because differences may be otherwise explained by differences in their resistance or in the standard deviations of the process; and (iv) anyone using these index should be aware that they introduce purely

artifactual trade-offs between resilience and resistance that have no biological meaning.

INTRODUCTION

The functional stability of ecosystems and its dependence on biodiversity are among the main concerns about the sustainability of human activities. The concept of stability is divided into resistance (amount of change in a system process caused by a disturbance) and resilience (ability of a system to recover after a disturbance) (Pimm, 1984). For soil, several indices have been proposed for measuring resilience and resistance (Griffiths and Philippot, 2013). Among them, the two indices proposed by Orwin and Wardle (2004) are frequently used and have been recently highlighted (Shade et al 2012; Griffiths and Philippot, 2013) because they were formulated in such a way to be easily interpretable and to have the following properties: monotonic increase with increases in resistance and resilience; disturbance is measured as absolute difference to the control; it is bounded by +1 and -1, with +1 indicating no effect of disturbance/full recovery, for resistance and resilience, respectively; zero can never appear in the denominator; and the resistance is standardized by the control, whereas resilience is standardized by the initial change caused by disturbance.

The index proposed for resistance (RS) is:

$$RS(t_0) = 1 - \frac{2|C_0 - P_0|}{(C_0 + |C_0 - P_0|)}$$

where C_0 is the value of the response variable in the control and P_0 is the respective value in the disturbed soil at the end of the disturbance.

For resilience (RL), the index is calculated for a given time t_x ($x > 0$) as follows:

$$RL(t_x) = \frac{2|C_0 - P_0|}{(|C_x - P_x| + |C_0 - P_0|)} - 1$$

where C_0 and P_0 are as above, and C_x and P_x are the values of the response variable at the time point (t_x) chosen to measure resilience, in the control and disturbed soil, respectively.

Based on the formulation of these indices, we hypothesized that they have the following problems: (i) the resistance index does not necessarily reach +1 (no disturbance) when random variation is considered; (ii) the resistance index calculated on any process is influenced by the mean and standard deviation of its underlying distribution for the control treatment, even when no disturbance effect occurs; (iii) consequently, this index can falsely indicate differences in resistance between any two systems if the underlying distributions of the measured process for the control treatments differ between them; (iv) the resilience index will always be zero if no disturbance is present at any time; (v) this index does not necessarily reach +1 (full recovery) when random variation is considered; (vi) it varies with the standard deviation of the underlying distribution of the measured process in the control treatment; and (vii) it introduces an artifactual negative relationship between resistance and resilience, i.e., for any two equally resilient systems, this index will be higher for the one with lower resistance.

MATERIAL AND METHODS

Resistance

Simulations of normally distributed random variables corresponding to hypothetical ecosystem processes (hereafter referred as process) were conducted using Monte Carlo methods in R (R Development Core Team, 2015).

To test the resistance index, two treatments were simulated:

- C_0: Control treatment at the end of the disturbance.
- P_0: Disturbed treatment at the end of the disturbance.

All combinations of a sequence of standard deviations (SD) ranging from 0.1 to 4, and seven means (30, 50, 100, 150, 300, and 500) were used as parameters for normal distributions from which one set of 1,000 samples were drawn in each simulation for C_0. To further evaluate how the index varies with disturbance effects, six sets of P_0 were simulated for each of the controls described above by combining three levels of disturbance effects on the mean: no effect, 10, and 20% increase relative to the respective mean in the control; and two levels of disturbance effects on the standard deviation: no effect and an increase by 10% of the mean (i.e. SD in P_0 increased by 10% of the mean in C_0). Finally, the resistance index was calculated on the resulting dataset.

Resilience

To test the resilience index, C_0 and P_0 (defined as above) were used with two additional simulated treatments:

- C_x: Control treatment at any ($x > 0$) time point chosen to measure resilience.
- P_x: Disturbed treatment at any ($x > 0$) time point chosen to measure resilience.

Fully Resilient Systems

For the treatments C_0, C_x, and P_x, the process was simulated by drawing three sets of 1,000 samples in each normal distribution with mean 50 and standard deviations along a sequence ranging from 0.1 to 8. Because values for P_x and C_x were simulated from identical distributions, no disturbance effect was introduced in the hypothetical time chosen to measure resilience, apart from random variation.

To introduce a disturbance effect in P_0, the simulated values for this treatment were drawn as above, but with means multiplied by 1, 1.03, 1.05, 1.1, 1.3, 1.5, 2, 4, and 6 to simulate a respective increase of 0, 2, 5, 10, 30, 50, 100, 300, and 500% in the mean of the process in P_0 as compared to the respective mean in C_0.

Not Fully Resilient Systems

To evaluate the resilience index in a scenario where the system does not fully recover from disturbance, the simulations were performed as described above for fully resilient systems, with the exception that the mean in P_x was calculated as the mean in C_x multiplied by 1.1 or 2 in order to introduce a respective increase (disturbance effect) of 10% and 100% at the hypothetical time chosen to measure resilience. The resilience index was then applied to the resulting dataset.

The script for all these analyses is given in the supplementary material.

RESULTS AND DISCUSSION

Resistance

The simulated values of the resistance index monotonically decreased as the coefficient of variation in the control treatment increased. This relationship was more pronounced when no disturbance effect was introduced, but was still present at small effects (**Figure 1 and S1**). The susceptibility of the index to variation in the control treatment is potentially very misleading, because two systems that differ in the control means but have the same variances will be (and probably have been) considered to differ in their response to disturbance even when there is no effect of disturbance at all. Therefore, the resistance index should not be directly compared between any two systems unless the estimated coefficient of variation for the control treatments do not differ between them, especially for positive values of this index. Otherwise, simulations as those performed here would be necessary to verify whether or not differences in this index are due to those differences in the coefficient of variation.

Our simulations also indicate that when random variation is taken into account, the index can not be interpreted in the scale +1 (no disturbance effect) to -1 (maximal disturbance) as originally suggested. Even when no disturbance effect is present, the index will only approach +1 if the coefficient of variation is nearly zero. Consequently, this index should only be interpreted in relation to its expected value under the null model (no disturbance effect), taking into account the estimated mean and variance in the control treatment.

Resilience

Because the definition of resilience given by Pimm (1984) (how fast a system recovers from disturbance) was cited in the original work describing this index, one can assume that the index was proposed to measure resilience under

that definition. In that case, it makes sense that any two systems should be equally resilient if they fully recover (the process in the disturbed treatment does no differ from that in the control) from a disturbance at the same time, regardless of the magnitude of the initial disturbance. In this context, the index would correctly represent resilience if and only if there is no random variation in the system (**Proof 1**). In that case, if there is no final disturbance effect ($|C_x - P_x| = 0$), the index will necessarily yield +1, independently of the initial disturbance.

However, that does not hold true when the stochastic variation of natural systems is taken into account (**Figure 2**). In all simulations of fully resilient systems, differences in the response variable between control and disturbed treatments at the hypothetical time used to measure resilience (t_x) are only due to randomness because their respective values were drawn from identical normal distributions (same means and standard deviations). Consequently, one would expect this index to indicate a complete recovery of the system (+1) in all cases. On the contrary, the simulations show that the index not only fails to do that, but actually yield lower values of resilience for the more resistant systems, despite the fact that they are all equally resilient. The simulations for not fully resilient systems also indicate the same negative relationship between the resilience index and the resistance of the system. This is not only counter-intuitive, it is also very misleading because it introduces an artifactual trade-off between resistance and resilience that can be easily confused with that discussed in Pimm (1984).

Furthermore, it should be pointed out that this index is undefined if there is no disturbance effect at the initial time (t_0) and at the time chosen (t_x) to measure resilience (**Proof 2**). Although this is not mentioned in the original paper describing the index, it should not be calculated unless an evidence of disturbance is found. On the other hand, if randomness is taken into account, the index will yield values near to zero for fully resistant systems (Figure 2). Although this is coherent with the idea that the index should always yield 0 if

the final (absolute) disturbance effect equals the initial (absolute) disturbance effect, there is no point in measuring resilience if there is no initial response to disturbance. Therefore, values of zero for this index should be carefully interpreted as they can either indicate a fully resistant system (which does not make sense) or an unresilient system for which the (absolute) disturbance effect did not change between t_0 and t_x. This ambiguity can be easily avoided by only calculating this index after testing for and verifying a disturbance effect.

Finally, another aspect of the index that deserves attention is its dependence on the standard deviation, especially at lower (absolute) initial disturbance effects. For positive values of this index, it is negatively related to the standard deviation in the control, whereas for negative values the relationship is inverted. Therefore, this index should not be compared among systems with different standard deviations for the process in the control treatments. Moreover, to be coherent with the definition of resilience as the ability of a system to recover to pre-disturbance states, this index is only comparable between any two systems if: (i) the initial or the final disturbance effects are not null for both of them; (ii) they have similar initial disturbance effects and similar standard deviations in the controls. Otherwise, any difference in this index between systems should not be attributed to differences in their ability to recover from a disturbance because it may be equally explained by differences in the parameters of the underlying distribution of the measured process in the control treatments or by differences in their resistances. As recommended for the resistance index, the use of simulations may help to refute those hypothesis.

CONCLUDING REMARKS

Our results do not invalidate the two stability indices proposed by Orwin and Wardle (2004), but rather give a better understanding of their properties under

naturally stochastic conditions, which will avoid misinterpretations and erroneous conclusions. We showed that comparing these indices between systems is not a trivial task and should be performed carefully. Monte Carlo simulations, like those presented here, can be useful to assess whether differences in these indexes among systems are really due to disturbance effects or just the result of differences in the intrinsic properties of the systems.

REFERENCES

Pimm SL (1984) The complexity and stability of ecosystems. *Nature* **307**: 321-326.

Orwin KH & Wardle DA (2004) New indices for quantifying the resistance and resilience of soil biota to exogenous disturbances. *Soil Biology and Biochemistry* **36**: 1907-1912.

Shade A, Peter H, Allison SD, *et al*. (2012) Fundamentals of microbial community resistance and resilience. *Frontiers in Microbiology* **3**: 1-19.

Griffiths BS & Philippot L (2013) Insights into the resistance and resilience of the soil microbial community. *FEMS Microbiology Reviews* **37**: 112-129.

R Development Core Team (2014) R: A Language and Environment for Statistical Computing. Vienna, Austria.

LIST OF FIGURES

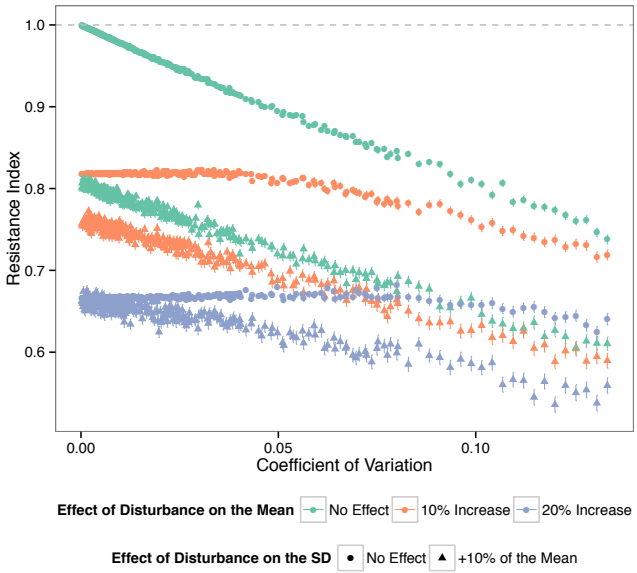

Figure 01. Simulated values for the resistance index with varying disturbance effects and coefficients of variation of the measured process. Points represent the mean (1,000 simulations) and bars indicate the standard error of the mean for each simulated pair of control and disturbed treatment. Point colors indicate the disturbance effect on the mean and the symbols indicate the disturbance effect on the standard deviation of the process according to the legend at bottom. Note that the y-axis does not start at zero.

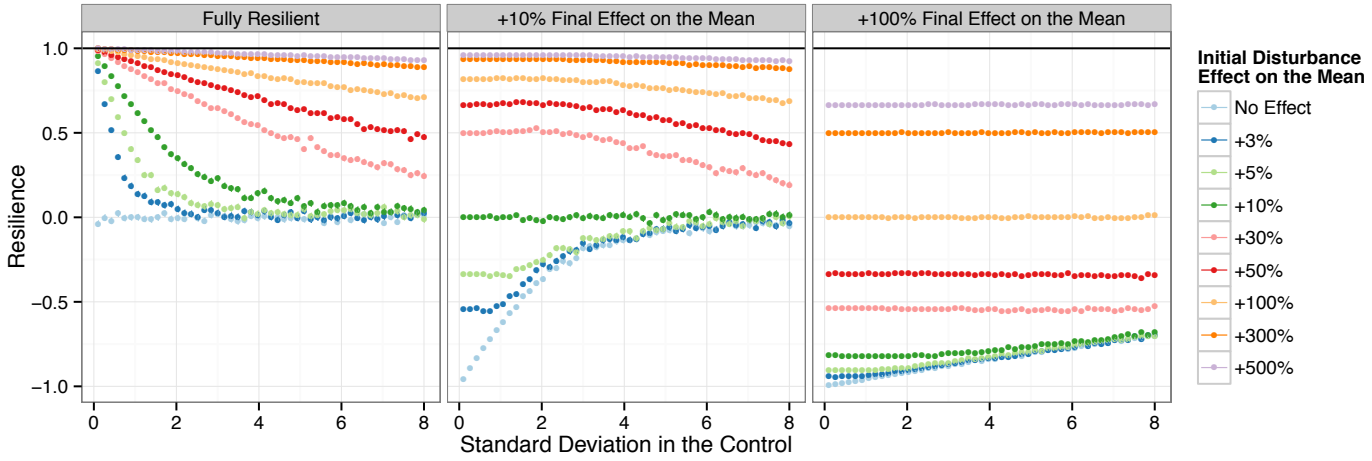

Figure 02. Simulations of the influence of increasing disturbance effects and standard deviations of the measured process on the resistance index. Each panel represents one level of resilience as follows: Fully resilient (no disturbance effect at the time chosen to measure resilience – t_x); +10% final effect on the mean (the mean in the disturbed treatment is 10% higher than that in the control at t_x); +100% final effect on the mean (the mean in the disturbed treatment is twice that in the control at t_x). Points represent the mean (1,000 simulations) and bars indicate the respective standard error for each set of simulated treatments. Point colors indicate the initial disturbance effect on the mean (increase in the mean of the process in the disturbed treatment relative to the mean in the control at the end of disturbance – t_0). The x-axis represents the standard deviation of the process in the control, also used for the other treatments.

SUPPLEMENTARY MATERIAL

Proof 1

The resilience index proposed by Orwin and Wardle (2004) can be written as

$$RL(t_x) = \frac{2|C_0 - P_0|}{(|C_x - P_x| + |C_0 - P_0|)} - 1$$

where C_0 and P_0 are respectively the values of the response variable for the control and disturbed treatments at the end of the disturbance; and C_x and P_x are the values of the response variable at the time point (t_x) chosen to measure resilience, in the control and disturbed soil, respectively.

Rewriting $|C_0 - P_0|$ as D_0 and $|C_x - P_x|$ as D_x, the index can be written:

$RL(t_x) = 2D_0 /(D_0 + D_x) - 1$

In a deterministic scenario (with no random variation, i.e. null variance), if there is no disturbance effect, $D_x = 0$, yielding:

$RL(t_x) = 2D_0 /(D_0 + 0) - 1$

$RL(t_x) = 2D_0 /D_0 - 1$

$RL(t_x) = 2 - 1$

$RL(t_x) = 1$

In that case, the index would always yield 1 for any fully resilient system ($D_x = 0$), regardless of the initial disturbance D_0. That is true because ($D_0 + D_x$) necessarily equals D_0 and cancel D_0 in the numerator.

However, considering random variation, if for simplicity we assume that C_x and P_x are two normally and independently distributed random variables with means μ_c and μ_P, and standard deviations σ_c and σ_P, respectively; and denoting $C_x - P_x$ as a new random variable D, it necessarily follows that D is also normally distributed with mean and standard deviation as follows

$\mu_D = \mu_c - \mu_P$ (1)

$\sigma_D = \sqrt{(\sigma_c^2 + \sigma_P^2)}$ (2)

Consequently, following Leone (1961), |*D*| will be a random variable with a "folded normal distribution" with the following expectation:

$$E(|D|) = \sigma_D\sqrt{2/\pi}\, e^{-\mu_D^2/2\sigma_D^2} + \mu_D\left[1 - 2\Phi\left(-\frac{\mu_D}{\sigma_D}\right)\right] \qquad (3)$$

where $\phi()$ denotes the cumulative distribution function of a standard normal distribution.

If there is no disturbance effect: C_x and P_x will be identically distributed. Therefore, from (1) and (2) it follows that $\mu_D = 0$ and $\sigma_D = \sigma_c^2\sqrt{2}$, respectively. Substituting that in (3)

$$E(|D|) = \sigma_c\sqrt{2}\sqrt{2/\pi}\, e^0 + 0$$

$$E(|D|) = \sigma_c 2\sqrt{\pi}$$

As a consequence, when there is no disturbance effect, the expected value for D_x is dependent on the standard deviation and will equals zero if, and only if, $\sigma_c = 0$, that is, if there is no random variation. Only under that (unrealistic) condition, the index would give 1 for all equally resilient systems, regardless of the initial disturbance.

Under this scenario, when the standard deviation of the process in the control is not null (what is naturally true for any real process), the resilience index vary with the standard deviation and with the initial disturbance, even for fully resilient systems.

Proof 2

Not accounting for random variation, if there is no disturbance effect at any time, it follows that

$$|C_0 - P_0| = 0$$

$$|C_x - P_x| = 0$$

Therefore

$RL(t_x) = 0 / 0 - 1$

Which is undefined because of the zero division.

This is not true when randomness is taken into account for the reasons presented in Proof 1.

Reference

Leone FC, Nelson LS & Nottingham RB (1961) The Folded Normal Distribution. *Technometrics* **3**: 543-550.

R Script Used for the Simulations

Resistance

```
library(dplyr)
library(magrittr)
library(reshape2)
library(ggplot2)
library(stringr)

quartz()
# To allow full reproducibility, I set the
# seed to day when this analysis was done
set.seed(03132015)

n.sims <- 1000

##### RESISTANCE

####  Simulation of resistance under varying means and standard deviations

# Function to calculate resistance
calc_resistance <- function(ctrl, trat){
  if(any(c(ctrl, trat) < 0)) stop("Values cannot be negative!")
  D0 <- abs(ctrl - trat)
  RS <- 1 - 2 * D0 / (ctrl + D0)
  return(RS)
}

# Function to simulate the resistance given the means and standard deviations
# of the control
simulate <- function(mu_c, mu_p, sigma_c, sigma_p, n.sims, return_summary = TRUE){
  trat <- rnorm(n.sims, mean = mu_p, sd = sigma_p)
  ctrl <- rnorm(n.sims, mean = mu_c, sd = sigma_p)
  resistance <- calc_resistance(ctrl, trat)
  if(return_summary){
    return(c(Mean = mean(resistance),
             SE = sd(resistance) / sqrt(n.sims),
              mu_c, mu_p, sigma_c, sigma_p))
  } else {
    return(resistance)
  }
}
```

```
# Simulation of the effect of varing means, standard deviations, and effect sizes
# on resistance
levels_sigma <- c("No Effect", "+10% of the Mean")
levels_mean <- c(`1` = "No Effect", `1.1` = "10% Increase", `1.2` = "20% Increase")
means <- c(30, 50, 100, 150, 300, 500)
sigmas <- seq(.1, 4, length.out = 50)
resistance <- expand.grid(mu_c = means,
                          sigma_c = sigmas) %>%
 do({rbind(mutate(., effect_mean = 1, effect_sigma = 0),  # No effect on mean or SD
           mutate(., effect_mean = 1.1, effect_sigma = 0), # +10% on mean, no effect on SD
           mutate(., effect_mean = 1.1, effect_sigma = 1), # +10% on mean, +10% of mean on SD
           mutate(., effect_mean = 1, effect_sigma = 1),   # No effect on mean, +10% of mean on SD
           mutate(., effect_mean = 1.2, effect_sigma = 1), # +20% on mean, +10% of mean on SD
           mutate(., effect_mean = 1.2, effect_sigma = 0))}) %>% # +20% on mean, no effect on SD
 mutate(mu_p = mu_c * effect_mean,
        sigma_p = sigma_c + effect_sigma * 0.1 * mu_c) %>%
 apply(1, function(row){
   simulation <- simulate(mu_c = row["mu_c"],
                          mu_p = row["mu_p"],
                          sigma_c = row["sigma_c"],
                          sigma_p = row["sigma_p"],
                          n.sims = n.sims)
                 return(simulation)
 }) %>%
 t() %>%
 data.frame() %>%
 mutate(Effect_Sigma = ifelse(sigma_c/sigma_p == 1, levels_sigma[1], levels_sigma[2]),
        Effect_Mean = levels_mean[as.character(round(mu_p/mu_c, 1))]) %>%
 mutate(Effect_Sigma = factor(Effect_Sigma, levels = levels_sigma),
        Effect_Mean = factor(Effect_Mean, levels = levels_mean))

ggplot(resistance) +
 aes(y = Mean,
     x = (sigma_c / mu_c),
     colour = Effect_Mean,
     shape = Effect_Sigma) +
 geom_point() +
 geom_errorbar(aes(ymin = Mean - SE,
                   ymax = Mean + SE),
               width = 0,
               size = .3) +
 scale_colour_brewer(palette = "Set2") +
 geom_hline(y = 1, linetype = "dashed", colour = "grey") +
 theme(panel.grid = element_blank(),
       legend.position = "bottom") +
 labs(x = "Coefficient of Variation in the Control",
      y = "Resistance Index",
      colour = "Effect of Disturbance on the Mean",
      shape = "Effect of Disturbance on the SD")
```

Resilience

```
################## RESILIENCE

####  Simulation of resistance under varying means and standard deviations
calc_resilience <- function(ctrl0, trat0, ctrl1, trat1){
 if(any(c(ctrl0, ctrl1, trat0, trat1) < 0)) stop("Values cannot be negative!")
 D0 <- abs(ctrl0 - trat0)
 Dx <- abs(ctrl1 - trat1)
 RL <- (2 * D0 / (Dx + D0)) - 1
 return(RL)
}

simulate_resilience <- function(mu_c0, mu_p0, mu_c1, mu_p1,
                          sigma_c0, sigma_p0, sigma_c1,
                          sigma_p1, n.sims, return_summary = TRUE){
 trat0 <- rnorm(n.sims, mean = mu_p0, sd = sigma_p0)
 ctrl0 <- rnorm(n.sims, mean = mu_c0, sd = sigma_p0)
 trat1 <- rnorm(n.sims, mean = mu_p1, sd = sigma_p1)
 ctrl1 <- rnorm(n.sims, mean = mu_c1, sd = sigma_p1)
 resilience <- calc_resilience(ctrl0 = ctrl0, trat0 = trat0,
                          ctrl1 = ctrl1, trat1 = trat1)
 if(return_summary){
  return(c(Mean = mean(resilience),
         SE = sd(resilience) / sqrt(n.sims),
         mu_c0, mu_p0, mu_c1, mu_p1,
         sigma_c0, sigma_p0, sigma_c1,
         sigma_p1))
 } else {
  return(resilience)
 }
}
levels_recovery <- c("Fully Resilient", "+10% Final Effect on the Mean", "+100% Final Effect on the Mean")
levels_effect <- c("No Effect", "+3%", "+5%", "+10%", "+30%", "+50%", "+100%", "+300%", "+500%")
sigmas <- seq(.1, 8, length.out = 50)
resilience <- expand.grid(mu_c0 = 50,
            sigma_c0 = sigmas,
            d0_mean = c(1, 1.03, 1.05, 1.1, 1.3, 1.5, 2, 4, 6),
            d1_mean = c(1, 1.1, 2)) %>%
  mutate(mu_c1 = mu_c0,
        mu_p1 = mu_c1 * d1_mean,
        mu_p0 = mu_c0 * d0_mean,
        sigma_c1 = sigma_c0,
        sigma_p0 = sigma_c0,
        sigma_p1 = sigma_c0) %>%
```

```
  apply(1, function(row){
    simulation <- simulate_resilience(
                    mu_c0 = row["mu_c0"],
                    mu_p0 = row["mu_p0"],
                    mu_c1 = row["mu_c1"],
                    mu_p1 = row["mu_p1"],
                    sigma_c0 = row["sigma_c0"],
                    sigma_p0 = row["sigma_p0"],
                    sigma_c1 = row["sigma_c1"],
                    sigma_p1 = row["sigma_p1"],
                    n.sims = n.sims)
    simulation_resistance <- simulate(mu_c = row["mu_c0"],
                        mu_p = row["mu_p0"],
                        sigma_c = row["sigma_c0"],
                        sigma_p = row["sigma_p0"],
                        n.sims = n.sims)
    return(c(simulation, Resistance = simulation_resistance["Mean"]))
  }) %>%
  t() %>%
  data.frame() %>%
  mutate(Effect_D0 = 100 * (mu_p0/mu_c0 - 1),
       Effect_D0 = sprintf("+%d%%", round(Effect_D0)),
       Effect_D0 = ifelse(Effect_D0 == "+0%", "No Effect", Effect_D0),
       Recovery = round(100 * (mu_p1/mu_c1 - 1)),
       Recovery = sprintf("+%d%% Final Effect on the Mean", Recovery),
       Recovery = ifelse(Recovery == "+0% Final Effect on the Mean",
                   levels_recovery[1], Recovery)) %>%
  mutate(Recovery = factor(Recovery, levels = levels_recovery),
       Effect_D0 = factor(Effect_D0, levels = levels_effect))

#quartz()

theme_set(theme_bw())
ggplot(resilience) +
  aes(x = sigma_c0,
     y = Mean,
     colour = factor(Effect_D0)) +
  facet_wrap(~Recovery, ncol = 3) +
  geom_point(size = 1.5) +
  scale_colour_brewer(palette = "Paired") +
  geom_errorbar(aes(ymin = Mean - SE,
                ymax = Mean + SE),
            width = 0,
            size = .3) +
  geom_hline(y = 1) +
  labs(x = "Standard Deviation in the Control",
     y = "Resilience",
     colour = "Initial Disturbance\nEffect on the Mean")
```

APPENDIX A: Other outcomes

Philosophical transactions of the Royal Society of London. Series B, Biological sciences

A social and ecological assessment of tropical land uses at multiple scales: the Sustainable Amazon Network

Toby A. Gardner , Joice Ferreira , Jos Barlow , Alexander C. Lees , Luke Parry , Ima Célia Guimarães Vieira , Erika Berenguer , Ricardo Abramovay , Alexandre Aleixo , Christian Andretti , Luiz E. O. C. Aragão , Ivanei Araújo , Williams Souza de Ávila , Richard D. Bardgett , Mateus Batistella , Rodrigo Anzolin Begotti , Troy Beldini , Driss Ezzine de Blas , Rodrigo Fagundes Braga , Danielle de Lima Braga , Janaína Gomes de Brito , Plínio Barbosa de Camargo , Fabiane Campos dos Santos , Vívian Campos de Oliveira , Amanda Cardoso Nunes Cordeiro , Thiago Moreira Cardoso , Déborah Reis de Carvalho , Sergio André Castelani , Júlio Cézar Mário Chaul , Carlos Eduardo Cerri , Francisco de Assis Costa , Carla Daniele Furtado da Costa , Emilie Coudel , Alexandre Camargo Coutinho , Dênis Cunha , Álvaro D'Antona , Joelma Dezincourt , Karina Dias-Silva , Mariana Durigan , Júlio César Dalla Mora Esquerdo , José Feres , Silvio Frosini de Barros Ferraz , Amanda Estefânia de Melo Ferreira , Ana Carolina Fiorini , Lenise Vargas Flores da Silva , Fábio Soares Frazão , Rachel Garrett , Alessandra dos Santos Gomes , Karoline da Silva Gonçalves , José Benito Guerrero , Neusa Hamada , Robert M. Hughes , Danilo Carmago Igliori , Ederson da Conceição Jesus , Leandro Juen , Miércio Junior , José Max Barbosa de Oliveira Junior , Raimundo Cosme de Oliveira Junior , Carlos Souza Junior , Phil Kaufmann , Vanesca Korasaki , Cecília Gontijo Leal , Rafael Leitão , Natália Lima , Maria de Fátima Lopes Almeida , Reinaldo Lourival , Júlio Louzada , Ralph Mac Nally , Sébastien Marchand , Márcia Motta Maués , Fátima M. S. Moreira , Carla Morsello , Nárgila Moura , Jorge Nessimian , Sâmia Nunes , Victor Hugo Fonseca Oliveira , Renata Pardini , Heloisa Correia Pereira , Paulo Santos Pompeu , Carla Rodrigues Ribas , Felipe Rossetti , Fernando Augusto Schmidt , Rodrigo da Silva , Regina Célia Viana Martins da Silva , Thiago Fonseca Morello Ramalho da

Silva , Juliana Silveira , João Victor Siqueira , **Teotonio Soares de Carvalho** , Ricardo R. C. Solar , Nicola Savério Holanda Tancredi , James R. Thomson , Patrícia Carignano Torres , Fernando Zagury Vaz-de-Mello , Ruan Carlo Stulpen Veiga , Adriano Venturieri , Cecília Viana , Diana Weinhold , Ronald Zanetti , Jansen Zuanon

Abstract

Science has a critical role to play in guiding more sustainable development trajectories. Here, we present the Sustainable Amazon Network (*Rede Amazônia Sustentável*, RAS): a multidisciplinary research initiative involving more than 30 partner organizations working to assess both social and ecological dimensions of land-use sustainability in eastern Brazilian Amazonia. The research approach adopted by RAS offers three advantages for addressing land-use sustainability problems: (i) the collection of synchronized and co-located ecological and socioeconomic data across broad gradients of past and present human use; (ii) a nested sampling design to aid comparison of ecological and socioeconomic conditions associated with different land uses across local, landscape and regional scales; and (iii) a strong engagement with a wide variety of actors and non-research institutions. Here, we elaborate on these key features, and identify the ways in which RAS can help in highlighting those problems in most urgent need of attention, and in guiding improvements in land-use sustainability in Amazonia and elsewhere in the tropics. We also discuss some of the practical lessons, limitations and realities faced during the development of the RAS initiative so far.