THE BIOGEOGRAPHY AND FUNCTIONAL ECOLOGY OF TROPICAL SOIL MICROORGANISMS

by

KYLE MATTHEW MEYER

A DISSERTATION

Presented to the Department of Biology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2016

DISSERTATION APPROVAL PAGE

Student: Kyle Matthew Meyer

Title: The Biogeography and Functional Ecology of Tropical Soil Microorganisms

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Biology by:

Scott Bridgham	Chairperson
Brendan Bohannan	Advisor
William Cresko	Core Member
Jorge Rodrigues	Core Member
Daniel Gavin	Institutional Representative

and

Scott L. Pratt	Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded December 2016

DISSERTATION ABSTRACT

Kyle Matthew Meyer

Doctor of Philosophy

Department of Biology

December 2016

Title: The Biogeography and Functional Ecology of Tropical Soil Microorganisms

Tropical ecosystems are some of the most diverse and productive ecosystems on the planet. These ecosystems are also some of the most threatened worldwide and this is largely driven by agricultural expansion. Predicting biotic responses to such forms of environmental change is a challenge that requires an increased understanding of the factors structuring these communities in both pristine environments as well as environments undergoing environmental change. Studying patterns in the spatial structure of communities can provide important insights into ecological and evolutionary processes structuring communities. Combining such approaches with analyses of the distribution of activity and the genomic content of communities can help us better understand relationships between community structure and function. I explore the topics of microbial spatial scaling, activity, and gene content in both pristine tropical rainforest environments and tropical regions undergoing agricultural conversion. I first pose a fundamental question in microbial spatial ecology, *i.e.* why do microorganisms tend to show weaker spatial patterns than macro-organisms? I show that trees and soil

microorganisms differ in the rates at which their communities change over space. I test the hypothesis that low rates of spatial turnover in microbial communities are an artifact of how we assess the community structure of microbial communities and show that sampling extent is likely the main driver of these differences. Next, I examine a Central Africa ecosystem undergoing agricultural conversion. I show that there are numerous indications of biotic homogenization in these soil microbial communities and that the active fraction of the community shows a more pronounced response to environmental change. Finally, I examine two microbial processes in the Amazon Basin that have been reported to change following agricultural conversion: methane production and methane consumption. I investigate changes to the genes and taxa involved in these processes and propose a new conceptual framework for how these processes might be changing. Work in this thesis contributes to a broader understanding of the spatial and functional ecology of tropical microorganisms and offers perspectives useful for predicting and mitigating the impacts of environmental change on these communities.

This dissertation includes previously submitted and co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Kyle Matthew Meyer

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon

The University of Auckland, Auckland, New Zealand

University of Wisconsin – Milwaukee, Milwaukee, Wisconsin

DEGREES AWARDED:

Doctor of Philosophy, Biology, 2016, University of Oregon

Master's of Science, Biology, 2011, The University of Auckland

Bachelor's of Science, Biology & French, University of Wisconsin-Milwaukee

AREAS OF SPECIAL INTEREST:

Microbial Ecology

Community Ecology

GRANTS, AWARDS, AND HONORS:

Donald Wimber Award - 2016

Best Student Talk UO Biology Student Retreat - 2015

William R. Sistrom Award - 2015

School of Biological Sciences First Class Graduating Honors 2011

J. William Fulbright Scholarship – New Zealand 2010-11

Phi Beta Kappa Membership – 2008

Ruth Walker Graduating Senior Award - 2008

Botanical Society of America – Young Botanist of the Year, 2007

PUBLICATIONS:

K.M. Meyer, P. Deines, P. Schupp, M. Taylor. Impact of explantation techniques on the microbiota of the marine sponge *Ecionemia alata*. 2016. Journal of Experimental Marine Biology and Ecology.

A. A. Navarrete, A. M.Venturini, **K.M. Meyer**, A.M. Klein, B.J.M. Bohannan, K. Nüsslein, S.M. Tsai, J.L.M. Rodrigues. Differential response of the Acidobacteria to forest-to-pasture conversion and their biogeographic patterns in the western Brazilian Amazon. 2015. Frontiers in Microbiology.

S.B. Hoot, **K.M. Meyer**, J.C. Manning. Phylogeny and Reclassification of *Anemone* (Ranunculaceae) with an Emphasis on Austral Species. 2012. Systematic Botany.

K.M. Meyer, S.B. Hoot, M.T.K. Arroyo. Phylogenetic Affinities of South American *Anemone* (Ranunculaceae), Including the Endemic Segregate Genera, *Barneoudia* and *Oreithales*. 2010. International Journal of Plant Sciences.

M.D. Rausher, Y. Lu, and **K.M. Meyer**. Variation in Constraint *vs.* Positive Selection as Explanation for Evolutionary Rate Variation among Anthocyanin Genes. 2008. Journal of Molecular Evolution.

Contribution acknowledged in:
S.A. Mangan, S.A. Schnitzer, E.A. Herre, K. Mack, M. Valencia, E. Sanchez, and J.D. Bever. Negative plant-soil feedback predicts relative species abundance in a tropical forest. 2010. Nature.

ACKNOWLEDGEMENTS

This thesis is the culmination of my development as a scientist over the last four and a half years. There are a number of truly inspiring people who have pushed me to expand the way I see the world and I would like to thank them for this.

Brendan Bohannan has not only acted as an exceptional mentor, but also as a friend. Brendan has always been an excellent example of a leader, a collaborator, a diplomat, and a critical thinker. He has also helped me realize how to balance my life to remain continually inspired.

My girlfriend, Johanna Tomorsky, has never ceased to amaze me and to challenge me to be a better person. We have grown so much over our years together and will continue to do so in years to come.

My parents, Doug and Nancy Meyer, have instilled in me the value of a well-balanced life filled with education, curiosity, and adventure. Without their love and support none of this would have been possible.

Ann Klein has been an excellent mentor and friend ever since I joined the institute. She has enthusiastically helped me with many computational and conceptual issues and has always been available to talk about microbes.

Anya Hopple and Scott Bridgham have instilled in me the wonder of the soil environment and have helped me navigate its many challenges.

Jorge Rodrigues has been an exceptional collaborator and has always been available for feedback and questions.

I would also like to thank the members of my committee, Scott Bridgham, Jorge Rodrigues, Dan Gavin, and Bill Cresko for inspiring me to be a more rigorous scientist.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

Tropical rainforests: model ecosystems for studying ecological interactions in a changing world

Tropical ecosystems are some of the most diverse and productive ecosystems worldwide. Because of the extraordinary variety of life in the tropics, many scientists have turned to these regions to develop or test theories in ecology and evolution. Tropical rainforests exert influence over the planet by producing oxygen, filtering freshwater, and sequestering carbon. Yet despite their global importance, these ecosystems are some of the most threatened on the planet (Dirzo & Raven, 2003). Environmental change- largely driven by agricultural development- is occurring disproportionately faster in the tropics than in any other region, driving increased rates of species loss as well as dramatic changes to the ways these ecosystems function (Laurance et al., 2014). There is considerable uncertainty in how these communities will respond to the growing human pressures imposed on them, largely due to the fact that these communities are so understudied. Directing our attention to better understand these threatened ecosystems is a priority and it is the only way forward if we are to predict and mitigate changes from mounting human pressures. This requires studying communities in intact environments and environments undergoing change, as well as navigating the numerous historical, ecological, and socioeconomic dimensions that make tropical environments unique.

The global importance of the soil environment and the many challenges it presents

All terrestrial life is intimately connected with the soil environment. Predicting the fate of soils in a changing environment is therefore an issue that is of global importance. Soils change over time through interactions with the environment and biotic communities. Changes in nutrient inputs, precipitation regimes, average temperatures, and land cover are just a few of the ways in which human activities can impact the soil environment and these can lead to changes to the communities of organisms inhabiting soil. Microbial communities in the soil are some of the most diverse communities on the planet and this diversity is mirrored by the tremendous breadth of metabolic activities performed by these communities (Fierer & Jackson, 2006). Soil bacteria, archaea, and fungi are directly involved in processes that can influence the assembly and productivity of aboveground communities (Wardle et al., 2004; van der Heijden et al., 2008), and they also govern ecosystem functions such as transforming soil chemicals (Falkowski et al., 2008) and cycling greenhouse gases (Schimel & Gulledge, 1998) that can impact Earth's climate. The exceptional diversity of these communities makes understanding their ecology and predicting their responses to environmental change one of the ultimate challenges in microbial ecology. Studying these communities and the processes they perform, however, can allow us to test predictions in ecology and uncover novel mechanisms for survival in such a dynamic environment.

The soil environment can harbor very high levels of microbial dormancy (*i.e.* taxa that are in a prolonged state of physiological inactivity) and some have

estimated that between 40 and 97% of bacterial cells present in soil can be dormant (Lennon & Jones, 2011; Blagodatskaya & Kuzyakov, 2013). Very little is known about the factors governing microbial resuscitation from dormancy and this presents a unique challenge to microbial ecologists interested in predicting community processes and responses to change. This problem is compounded by the fact that most of the traditional methods used to identify microbial taxa in the environment (*e.g.* DNA-metabarcoding) do not distinguish between active and dormant members. Since dormant individuals are likely avoiding negative interactions (either with their environment or other community members), it is only the active fraction of the community that is interacting with the environment and performing metabolic processes. If the active fraction of the community is more intimately interacting with the environment, it may be structured differently than the "seedbank" of dormant taxa and it may also be more susceptible to environmental change. From the few studies that look specifically at the active microbial taxa, it has been shown that this fraction may be more responsive to seasonal changes (Barnard et al., 2013) and may be more directly influenced by variation in the environment (Zhang et al., 2014). However, there has been little effort to measure the response of the active fraction to environmental change, and there has been even less effort to do so in tropical ecosystems.

Spatial and functional ecology as a framework to study environmental change

The distribution of taxa through space and time is a fundamental topic in biology and has led to profound insights into the processes driving species

evolution. Both microbial and macro-organismal communities show spatial patterns in their taxonomic composition, yet there appear to be differences in these patterns. For example, the decay of community similarity with geographic distance (*i.e.* the distance-decay relationship) and the accumulation of taxa with increasing area (*i.e.* the taxa-area relationship) - two well-studied relationships in spatial ecology- tend to be much weaker for microorganisms relative to macro-organisms (Green et al., 2004; Horner-Devine et al., 2004). Some have suggested that unicellular organisms display different spatial patterns because they are more capable of long-distance dispersal, have the potential to be dormant for long periods, and tend to be much more abundant (Finlay, 2002; Finlay et al., 2004)- *i.e.* they are fundamentally different from multi-cellular macro-organisms. Others have suggested that these differences may simply be an artifact of how we measure microbial communities- since our species delineations and means of surveying communities tend to be so different from those used for macro-organisms (Martiny et al., 2006; Hanson et al., 2012). Understanding whether micro- and macro-organisms are indeed fundamentally different in their spatial patterning not only presents an interesting opportunity to investigate general principles in ecology (*i.e.* those shared by *all* organisms), but it is also an important consideration for ecosystem management and conservation planning (which depend on spatial relationships to estimate extinction probabilities) since society depends on many microbial natural products and ecosystem functions.

Human impacts on the environment are intensifying globally and there is considerable uncertainty over how microbial communities will respond. Spatial

variation in microbial communities can be used as one measure of how a community responds to a changing environment. The effects of species diversity are thought to increase at larger spatial scales (Dimitrakopoulos & Schmid, 2004; Cardinale et al., 2012); thus a loss of spatial diversity could have negative consequences on the resilience of community functions across a landscape and could imply lower richness in the pool of species from which future communities will assemble. Species loss in microbial communities is a topic few have addressed largely due to the fact that microbial taxa have been generally assumed to be ubiquitous and hence not susceptible to extinction. Demonstrating an extinction event in a microbial community is also a considerable challenge because of the high abundance and diversity of microbial communities. Numerous alterations to ecosystems, however, have been shown to drive the local loss of certain microbial taxa and functional groups (Rodrigues et al., 2013; Navarrete et al., 2015b; Hamaoui et al., 2016) and this is of concern because the loss of soil microbial functions could have devastating consequences. Elucidating whether losses in spatial or local diversity follow general trends could help us better identify the drivers of change in these systems. Understanding microbial responses to environmental change can help us mitigate our impacts on these communities and it can help us to better understand how these poorly understood taxa interact with their environment.

Microbial communities can be altered by environmental change in many ways and this makes the task of predicting community responses all the more difficult. A change in the composition of a community, for example, does not necessarily imply a change in function since many microbial taxa are capable of

exchanging ecologically relevant genes. In this regard a change in the gene content of the community (*i.e.* the presence and abundance of genes that code for cellular and metabolic processes, aka the metagenome) may offer an alternative perspective (Fierer et al., 2014). For example, the loss of genes that encode certain processes (*e.g.* methane consumption) could be a strong indication of a loss or perturbation to that process. Certain combinations of genes could also represent life history strategies employed by microbial communities in a given environment and could lend insights into the environmental conditions as experienced by the community (Barberán et al., 2012). For example, an increase in genes related to temperature stress tolerance, dormancy, or spore formation could be indicative of a stressful (or unpredictable) environment- a perspective that we would not necessarily gain from taxonomy or abiotic measurements alone. Some have also argued that the processes that shape the taxonomic composition of a community differ from the processes that shape the gene content of the community (Louca et al., 2016). Thus, if we are interested in how community function may respond to environmental change, measuring gene content may be a more direct approach.

The spatial and functional ecology of tropical soil microorganisms

This dissertation addresses the topic of spatial and functional ecology of microbial communities in tropical soils. In Chapter II, I examine a classic paradigm in the field of microbial spatial ecology: that is, why do microbial communities tend to show much weaker spatial patterns than macro-organisms such as plants and animals. I compare soil bacterial spatial patterns to those of the tree community in a

Central African rainforest that is renowned for its high levels of plant and animal diversity. I test the hypothesis that low rates of spatial turnover in microbial communities are an artifact of how we assess the community structure of microbial communities. In Chapter III, I look at a Central African ecosystem that is a mosaic of land undergoing agricultural conversion. I survey lands representative of Central African agricultural conversion including a forested plot, a plot that had very recently been deforested and burned, a manioc/banana plantation plot, as well as an abandoned plantation, all within close proximity to one another. I test for indications of biotic homogenization in the microbial communities following ecosystem conversion to agriculture, and ask whether these changes share commonalities with other tropical ecosystems undergoing conversion. I also distinguish the active fraction of the microbial community from the inactive taxa to examine whether these members are disproportionately impacted by land use change. Finally in Chapter IV, I examine two microbial processes in the Amazon Basin that have been reported to change following conversion to agriculture: methane production and methane consumption. I take a metagenomic approach to investigate changes to the genes and taxa that are involved in these processes and examine several explanations for how these processes might be changing.

This dissertation includes previously submitted and unpublished co-authored material. In Chapter II Hervé Memighe, Lisa Korte, David Kenfack, Alfonso Alonso, and Brendan Bohannan are included as co-authors. In Chapter III Brendan Bohannan is a co-author. Chapter IV is currently in review at Molecular Ecology and

Ann Klein, Jorge Rodrigues, Klaus Nüsslein, Babur Mirza, Susanne Tringe, James Tiedje, and Brendan Bohannan are co-authors.

CHAPTER II

WHY DO MICROBES EXHIBIT WEAK BIOGEOGRAPHIC PATTERNS?

This chapter is co-authored by myself, Hervé Memiaghe, Lisa Korte, David Kenfack, Alfonso Alonso and Brendan Bohannan. Data analysis and writing were primarily performed by myself. Hervé Memiaghe, Lisa Korte, David Kenfack, and Alfonso Alonso performed the tree community census and Hervé provided the data for analysis. Brendan Bohannan filled the advisory role on this project and as such aided in the conception of the design as well as contributed edits to the text. Supplementary material for this chapter can be found in Appendix A.

INTRODUCTION

Biogeography describes the distribution of taxa over space and time, and it has led to fundamental insights into the mechanisms maintaining and generating species diversity. Numerous studies have established that microbial communities can exhibit biogeographic patterns, and in many cases these patterns are qualitatively similar to those of macro-organisms (Hillebrand et al., 2001; Green et al., 2004; Horner-Devine et al., 2004). Microbial biogeographic patterns, however, tend to be much weaker than those of macro-organisms. For example, the accumulation of taxa with increasing area and the decay of community similarity with geographic distance (two very well studied biogeographical patterns) tend to be lower for microorganisms than for plants and animals (Hillebrand et al., 2001; Green et al., 2004; Horner-Devine et al., 2004; Zhou et al., 2008). It is as yet unclear why this occurs.

Understanding why microorganisms differ quantitatively from plants and animals in their distribution is important for several reasons. First, biogeographic patterns can provide insight into the fundamental processes that determine biodiversity. Quantitative differences in biogeographic patterns could suggest that these fundamental processes are different for microbes and larger organisms. Second, biogeography forms a foundation for conservation and environmental management, including bioprospecting. Understanding whether or not microbial and plant/animal biogeography are governed by different rules is important for designing effective management and conservation strategies (Diamond, 1975; Simberloff & Abele, 1982; Soule & Simberloff, 1986).

Some have suggested that microbes have weak biogeographic patterns because they are fundamentally different in ways that alter their biogeography; for example, due to high abundance, longevity, or dispersal abilities (Finlay, 2002). Others however, have suggested that these differences are artifacts of how microbial biogeography is studied (Woodcock et al., 2006; Martiny et al., 2011). These artifacts could include: 1) that the operational taxonomic units (OTUs) used for characterizing microbes are not an appropriate analog to plant or animal species (Tiedje, 1995; Fuhrman & Campbell, 1998; Horner-Devine et al., 2004; Storch & Šizling, 2008), 2) microbial communities tend to contain high numbers of inactive individuals and most microbial surveys do not distinguish active from inactive individuals (Lennon & Jones, 2011; Blagodatskaya & Kuzyakov, 2013), 3) the spatial scales over which biogeographic patterns are assessed differ between microbial and plant/animal studies (*e.g.* Hillebrand et al., 2001), and 4) microbial communities

tend to be of much higher diversity than plant/animal communities, and thus more prone to severe undersampling, which in turn may result in under-estimating rates of taxonomic turnover (Woodcock et al., 2006). We consider the implications of each of these potential artifacts below.

How taxonomic groups are defined strongly differs between macro-organisms and microorganisms. For microbial taxa, morphological traits are rarely useful for separating lineages, and the physiological measurements necessary to distinguish taxa are possible only for the minority of taxa that can be grown in culture. Thus, researchers commonly delineate taxa using the sequence similarity of marker genes (most commonly ribosomal genes (Pace, 1997)). This sequence similarity is used to create operational taxonomic units (OTUs), defined by an arbitrary sequence similarity cutoff (*e.g.* 97%). It has been suggested that OTUs defined at 97% sequence similarity tend to contain much higher levels of diversity than typical plant or animal species, and thus may be more comparable to a higher taxonomic level, *e.g.* a genus or family (Horner-Devine et al., 2004; Hanson et al., 2012). It has been demonstrated that the choice of OTU similarity cutoff can impact diversity patterns (Storch & Šizling, 2008) including biogeographic patterns (Horner-Devine et al., 2004).

Not all microbial taxa are active in a given place and time (Lennon & Jones, 2011). Numerous microbial taxa are capable of entering a state of dormancy (*i.e.* physiological inactivity), and the percentage of microbial cells in this state can be as high as 80-97 percent in certain environments(Lennon & Jones, 2011; Blagodatskaya & Kuzyakov, 2013). This pool of inactive taxa has been likened to a

seed bank in that member taxa may emerge into a state of activity/growth in response to various biotic or abiotic cues much like plant seeds in the soil. The typical DNA-based surveys used to assess microbial community membership do not distinguish between active and inactive taxa. Locey (2010) argued that if dormancy increases the rate of immigration (by allowing immigrants to avoid initial adverse conditions) and decreases the rate of extinction (by allowing taxa to avoid death), then microbial communities containing dormant taxa should exhibit lower temporal turnover since the likelihood of a newcomer being a new species would decrease over time (Locey, 2010). The same argument could be used for spatial turnover, *i.e.* that over time the seed bank should tend to accumulate most regional taxa regardless of whether they are suited to the local environment. Thus, including inactive taxa in our surveys could decouple community turnover from environmental turnover and result in an underestimation of rates of community turnover.

It is well established that biogeographic patterns can change quantitatively with spatial scale. This is true for both microbes (Franklin & Mills, 2003; Martiny et al., 2011) and larger organisms (Preston, 1960; Nekola & White, 1999; Condit et al., 2002; Tuomisto et al., 2003; Soininen et al., 2007). It has been suggested that environmental filtering is a more important driver of biogeographic patterns at smaller spatial scales (Preston, 1960; Martiny et al., 2006, 2011) while dispersal limitation and/or diversification are more important drivers of large-scale spatial patterns (Papke et al., 2003; Whitaker et al., 2003; Green et al., 2004)- although dispersal limitation can also play a role at local scales as well (Bell, 2010; Martiny et

al., 2011). Microbial and plant/animal biogeographic surveys are often performed at different spatial scales and this could potentially confound our interpretations of how the diversity of these groups scales quantitatively. Including, for example, more small-scale spatial comparisons in a survey could make rates of community turnover appear lower when compared to a survey comprised mainly of large-scale comparisons.

Finally, incomplete sampling (aka *under-sampling*) of communities is a problem that exists throughout ecology (Cam et al., 2002) but is particularly pronounced for microbial communities, which tend to be especially diverse. By under-sampling a community our surveys tend to be biased against rare community members. Rare members are often more restricted in range and hence could be important in determining biogeographic patterns. Woodcock et al. (2006) showed that the rate at which microbial species richness increases with area can be strongly influenced by the intensity of sampling effort (Woodcock et al., 2006). However, it has also been suggested that rare taxa exert a relatively minimal effect on microbial biogeographic patterns compared to the effects of species abundances and levels of population aggregation (Nekola & White, 1999; Morlon et al., 2008). The impacts of under-sampling on biogeography in environmental surveys has rarely been assessed and, to our knowledge, never in the context of accounting for the differences between microbial and plant/animal biogeographic patterns.

Here we compare the rates of the decay of taxonomic similarity over geographic distance between the soil microbial community and the tree community at the Rabi plot, Gabon, a research site established for the purposes of studying

spatial ecology and terrestrial biogeography. The distance decay of community similarity is a fundamental pattern in the biogeography of plant/animal (Nekola & White, 1999; Condit et al., 2002; Tuomisto et al., 2003; Soininen et al., 2007; Morlon et al., 2008) and microbial (Hillebrand et al., 2001; Green et al., 2004; Horner-Devine et al., 2004; Fierer & Jackson, 2006; Bell, 2010; Zinger et al., 2014) taxa. Our design allows us to compare this relationship across spatial scales ranging from centimeters to 100s of meters. We test the following hypotheses: 1) microbial species definitions will influence the rate at which microbial community similarity changes over space, 2) excluding inactive microbial taxa will result in the steepening of microbial distance-decay patterns, 3) microbial and tree distance decay patterns will become more similar when compared at the same spatial scales, and 4) the effects of under-sampling a community can account for the differences between microbial and tree distance-decay rates.

MATERIALS AND METHODS

Experimental Design

Samples were taken at the Smithsonian Center for Tropical Forest Science's (CTFS) 25 ha plot located near the Rabi oil field in Southwestern Gabon, adjacent to the Gamba Complex of Protected Areas (Lee et al., 2006) at the end of the dry season in September 2013. This site is ideal for testing questions in microbial biogeography. It was established for the purposes of studying spatial ecology and terrestrial biogeography; for example, it is extensively mapped, and it has a spatial grid system (with permanent ground markers) that extends across the entire site.

Moreover, the CTFS plot at Rabi is particularly advantageous in that the tree community has been completely censused (Memiaghe et al., 2016), which allows for direct comparisons between spatial patterns of trees and microbes in the same landscape.

Within the 25 ha plot, we sampled using a spatially explicit nested sampling scheme (Supp. Fig. 1a) whereby three 100 m^2 quadrats were established with 10 m^2, 1 m^2, 0.1 m^2, and 0.01 m^2 quadrats nested within each. This design has been previously implemented to measure spatial turnover (Rodrigues et al., 2013), and gives high coverage of a range of spatial scales. Soil cores were taken from the corners of each quadrat giving a total of 39 samples. Soil cores were taken using standard coring methods to a depth of 15 cm, following the removal of the litter layer. For each sampling point three representative soil cores were taken, homogenized, then either subsampled and preserved for molecular analysis (described below) or kept on ice and transported back to the US for soil chemical analysis (described below).

Tree census data were obtained for all individuals greater than 1 cm diameter at breast height (dbh) for all areas of the 25 ha plot overlapping with the soil bacterial census (Memiaghe et al., 2016, Supp. Fig. 1b). To assess tree community turnover, the composition of each of the 20 m x 20 m plots included in the study were compared.

Molecular analysis

From each set of homogenized soil cores, 3 ml (approximately 1 g) of soil was added to 9 ml Lifeguard solution (Mobio, California, USA) in the field, then shipped cold and stored at -80° C in order to stabilize nucleotides for later extraction. Soil DNA and RNA were co-extracted from each sample using MoBio's Powersoil RNA Isolation kit with the DNA Elution Accessory Kit (MoBio, California, USA) following manufacturer's instructions, using 3 ml of the soil:Lifeguard mixture (~ 0.25 g). Extractions were quantified using Qubit (Life Technologies, USA). RNA was reverse transcribed to cDNA using Superscript III first-strand reverse transcriptase and random hexamer primers (Life Technologies, USA).

The V3 and V4 region of the 16S rRNA gene of the DNA and cDNA were PCR amplified using the primers 319F and 806R (Fadrosh et al., 2014). Sequencing libraries were prepared using a 2-step PCR with a dual-indexing approach (Kozich et al., 2013; Fadrosh et al., 2014). In short, the first round of amplification consisted of 22 cycles with Phusion HiFi polymerase using an equimolar blend of six forward and reverse primers. The same blend of all 6 forward and reverse primers for each sample were used in order to control for any potential bias that could be introduced by adding the heterogeneity spacers immediately upstream of the 319F and 806R primer sequences. Variable length spacers were added in order to avoid problems that arise when sequencing libraries of similar amplicon length on the Illumina platform. Round 1 products were cleaned using Agencourt AMPure XP (Beckman Coulter, California, USA), then amplified for an additional 6 cycles using Phusion HiFi. Step 2 adds the sequences required for cluster formation on the Illumina

flowcell. The final library was sent to the Dana-Farber Cancer Institute Molecular Biology Core Facilities for 300bp paired-end sequencing on the Illumina MiSeq platform.

Soil Chemical Analysis

Soil chemical parameters were measured from each soil core to estimate the impact on microbial community composition by A & L Western Agricultural Lab (Modesto, CA, USA). In total, percent organic matter (loss on ignition (Dean, 1974)), extractable phosphorus (Weak Bray (Kamprath & Watson, 1980) & sodium bicarbonate (Olsen, 1954)), extractable cations (K, Mg, Ca, Na), sulfate-S (Fox et al., 1964), pH, buffer pH, cation exchange capacity (CEC, (Chapman, 1965)), and percent cation saturation were measured.

Data processing and statistical analysis

Paired end reads were joined then demultiplexed in QIIME (Caporaso et al., 2010) before quality filtering. Primers were removed using a custom script. UPARSE was used to quality filter and truncate sequences (416bp, EE 0.5) (Edgar, 2013). Sequences were retained only if they had an identical duplicate in the database. Operational taxonomic units (OTUs) were clustered de novo at 97% using USEARCH (Edgar, 2010). OTUs were checked for chimeras using the gold database in USEARCH. We used a custom script to format the UCLUST output for input into QIIME. To assign taxonomy, we used repset from UPARSE in QIIME using greengenes version 13_5 (RDP classifier algorithm). Finally, we rarefied to 3790

observations per sample in QIIME to achieve approximately equal sampling depth, which excluded 4 samples.

After quality filtering, demultiplexing, and OTU clustering, statistical analyses were performed in the R platform (R Development Core Team, 2010). Canberra pairwise community distances were calculated for both the bacterial and tree communities using the vegdist function in the package 'vegan' (Oksanen et al., 2015). Canberra was chosen because of its incorporation of abundance data and sensitivity to rare community members (Jost et al., 2011). Turnover was estimated for both the bacterial and tree communities by regressing pairwise similarity against pairwise geographic distance (Nekola & White, 1999). Mantel tests were used to test for significant associations between geographic and community distance in base R. Distance-decay slopes were compared using the function diffslope in the package 'simba' (Jurasinski & Retzer, 2012), which employs a random permutation approach to the distances and calculates the difference in slopes. The p-values computed are the ratio between the number of cases where the differences in slope exceed the difference in slope of the initial configuration and the number of permutations.

The relative impacts of the environment and geographic distance on microbial community dissimilarity were assessed using multiple regression on distance matrices as implemented by the MRM function in the 'ecodist' package (Goslee & Urban, 2007) in R. Environmental dissimilarity was calculated using the Gower general dissimilarity coefficient (Gower, 1971) as implemented by the function daisy in the 'cluster' package (Maechler et al., 2016) in R. The influence of

individual soil parameters on community dissimilarity was assessed using a redundancy analysis as implemented by the rda function in 'vegan' (Oksanen et al., 2015) following Hellinger transformation of the community data.

OTU clustering experiment

To test whether species definition impacts biogeographic patterns, OTUs were clustered at 95%, 97%, 99%, and 100% similarity thresholds, each time using the aforementioned bioinformatic pipeline. Clustering at these levels resulted in 1179, 2243, 6611, and 14,864 OTUs, respectively. OTU tables were rarefied to 3696, 3100, 3324, and 2521 observations per sample (the minimum number of observations per sample that would allow us to retain all samples), respectively, to achieve approximately equal sampling depth. Linear models of community turnover (described above) were compared against the tree community turnover linear model for each OTU threshold using the random permutation approach described above.

RNA- vs *DNA-inferred community comparison*

To ask whether distinguishing the active bacterial community members from the inactive members would impact biogeographic patterns, we inferred bacterial community membership using two molecular methods: analysis of community RNA and analysis of community DNA. By inferring community membership *via* RNA we enrich for taxa that are active, whereas communities inferred *via* DNA will tend to include a higher proportion of inactive members. Distance-decay linear regression

slopes were compared between the RNA- and DNA-inferred communities clustered at the 97% OTU similarity threshold using the aforementioned permutation approach.

Spatial scale

To assess whether bacterial community distance-decay rates more closely resemble tree community distance-decay at the same spatial scale, we subset the bacterial community to only include comparisons at the same spatial scale as trees. We also asked whether bacterial distance-decay patterns differed at different spatial scales by subsampling our data to include only small- to medium-scale comparisons (tens of centimeters to tens of meters) and medium- to large-scale comparisons (tens of meters to hundreds of meters).

Effects of undersampling

We used rarefication to assess the impact of undersampling on biogeographic patterns for both tree and bacterial communities. We wrote a custom R function (provided in the supplementary code) that repeatedly subsamples (1000 times) a community at a given depth and computes a distance-decay linear regression for each sampling event. For this study we used a 97% OTU cutoff for the DNA-inferred community.

RESULTS & DISCUSSION

Community similarity (1- Canberra dissimilarity) significantly decreased with geographic distance for both the bacterial (Mantel r = 0.55, $p = 0.001$) and tree (Mantel r = 0.47, $p = 0.001$) communities (Fig. 1). The soil chemical environment showed slight spatial autocorrelation over the distances covered (Mantel r = 0.11, $p < 0.01$), but was relatively uniform. Variation in the soil chemical environment overall was not significantly correlated with bacterial community turnover (Partial Mantel r = 0.14, $p = 0.10$) after having controlled for the effects of distance. There were, however, certain soil chemical attributes that were significantly associated with bacterial community structure (pH and S).

The rate at which community similarity decayed over space differed significantly between the bacterial and tree communities (difference in slope: 0.02, $p < 0.001$) with the tree community exhibiting a sharper rate of turnover (-0.0358 ± 0.001) than the bacterial community (Fig. 1, -0.0148 ± 0.0007). This finding is consistent with many other observations that microbial communities often exhibit more gradual distance-decay patterns (Hillebrand et al., 2001) and accumulate new taxa at a lower rate over space (Green et al., 2004; Horner-Devine et al., 2004) relative to plant and animal communities.

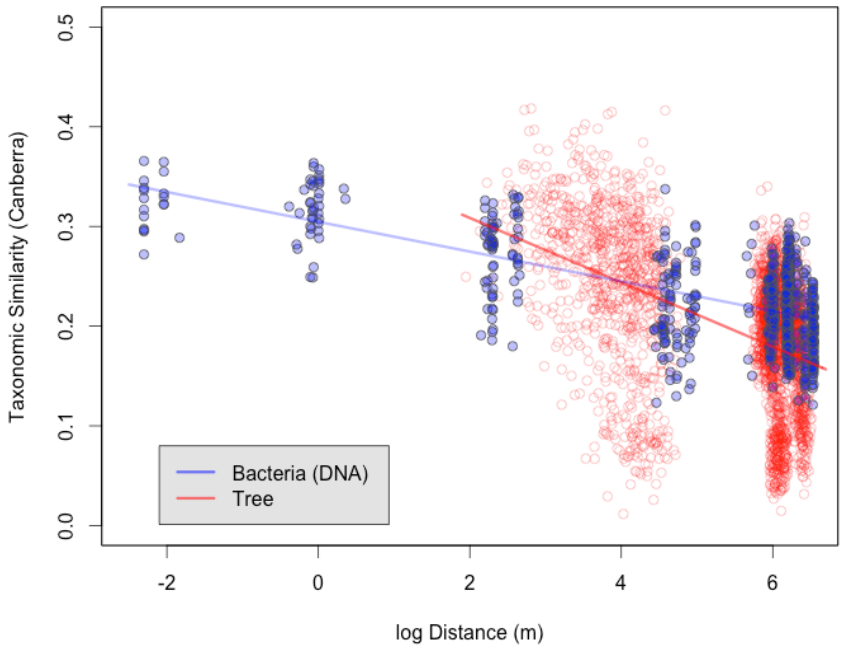


Figure 1: Distance-decay plot of the bacterial community (inferred from DNA, OTU cutoff = 97%) *versus* the tree community on the Rabi plot, Gabon.

The impact of OTU clustering

Various studies have suggested that broadening taxonomic resolution (for example, by comparing genera or families, rather than species) can decrease the strength of biogeographic patterns (Horner-Devine et al., 2004; Storch & Šizling, 2008; Hanson et al., 2012), although not always (Green & Bohannan, 2007). We asked whether altering the sequence similarity cutoff we used to define our taxa (analogous to moving from subspecies to species to genera and families) could impact the rate of bacterial community turnover in our data. Neither broadening (*i.e.* to 95%) nor narrowing (*i.e.* to 99 and 100%) the sequence similarity cutoffs altered the rate of community turnover (Fig. 2). The range of taxonomic similarity values,

however, did change with taxonomic definition, as expected. Broader cutoffs tended to exhibit higher levels of taxonomic similarity while narrower cutoffs exhibited lower ranges of taxonomic similarity.

Our results are in contrast to Horner-Devine *et al.* (2004) who reported that narrowing the sequence similarity cutoff for taxon definition resulted in a steeper bacterial distance-decay slope. There are a number of potential explanations for why we did not observe this in our study. Our findings might be different because the contribution of environmental variation to bacterial community turnover was lower in our study than that reported by Horner-Devine *et al.* (2004). If the distance decay of community similarity is driven strongly by the distance decay of environmental similarity, and if narrowing taxonomic resolution results in groups with narrower environmental tolerances, then a steeper distance decay pattern should result. Another possibility is that the traits required for survival under any given set of environmental conditions were strongly phylogenetically conserved in the taxa in our study. This would result in less of an impact of changing taxonomic (*i.e.* phylogenetic) resolution on the breadth of environmental tolerances (and ultimately, the rate of distance-decay).

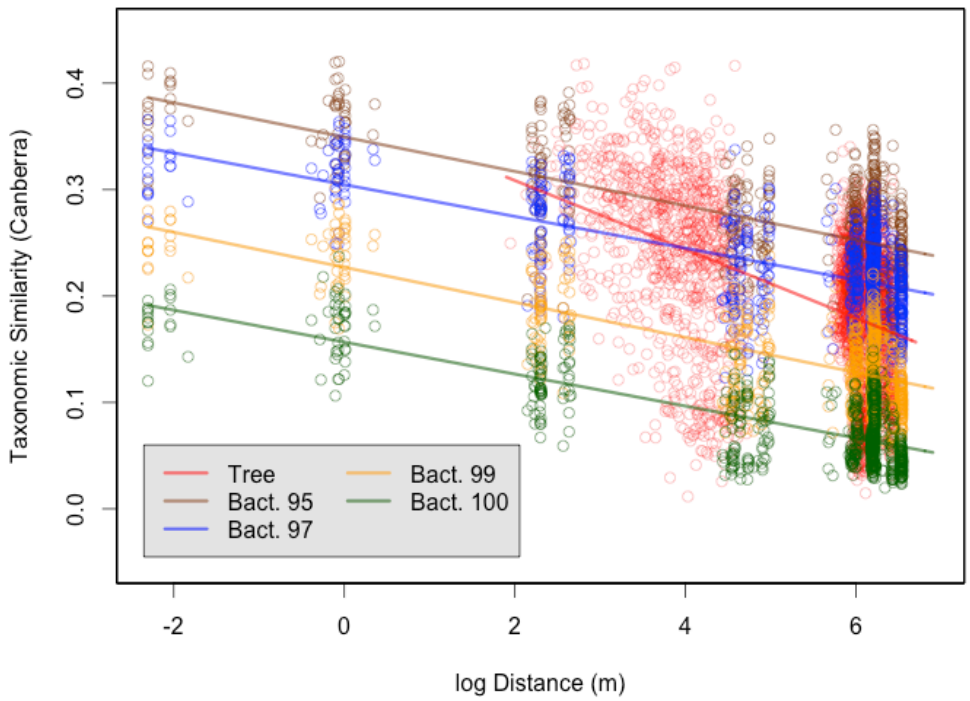


Figure 2: The impacts of changing OTU threshold on distance-decay patterns of soil bacterial community at the Rabi plot, Gabon.

Excluding inactive taxa

The soil environment contains especially high proportions of physiologically inactive (*i.e.* dormant) microbial taxa (Lennon & Jones, 2011; Blagodatskaya & Kuzyakov, 2013) and most DNA-based microbial surveys include both active and inactive taxa. In contrast, biogeographic surveys of plants and animals rarely include dormant individuals (*e.g.* seeds). Given that dormancy can allow taxa to persist outside of optimal environmental conditions, the inclusion of inactive taxa could decouple microbial community turnover from environmental turnover. We hypothesized that if landscape level distance-decay relationships are largely driven by environmental turnover, then including inactive taxa in a microbial survey would

flatten the distance-decay slope. Thus, by excluding the inactive taxa (and focusing solely on the active taxa) we would expect the microbial distance-decay slope to become steeper and that this could – at least in part- account for the differences in biogeographic patterns between trees and microbes in our study

Excluding inactive taxa, however, did not result in a steeper distance-decay slope in our study (Fig. 3). The RNA-inferred (active) distance-decay slope (-0.0116 ± 0.001) was significantly flatter than the DNA-inferred (active + inactive) distance-decay slope (-0.0149 ± 0.0007, Difference in slope = 0.0033, $p = 0.01$) and both community distance-decay rates were lower than the tree community distance-decay rate (-0.0358 ± 0.0012). For both communities, geographic distance was a more important predictor of community variation than turnover in the soil chemical environment. Variation in the DNA-inferred community structure was more predictable overall by our meta-data (geographic distance and soil chemical environment) than the RNA-inferred community. In fact, variation in the soil chemical environment was not a significant predictor of variation in the RNA-inferred community.

Our observation is at odds with our expectation that the RNA-inferred community would more closely track the soil chemical environment than the DNA-inferred community, since the DNA-inferred community could include dormant taxa that are not interacting with their local environment. This hypothesis does, however, rely on the assumption that the environmental factors responsible for microbial activity are spatially autocorrelated. Alternatively, if climatic variables such as rainfall events -which tend to be relatively uniform over a landscape- are

stronger determinants of soil activity, then we would expect the active community to be more uniform over space, which is what we observed.

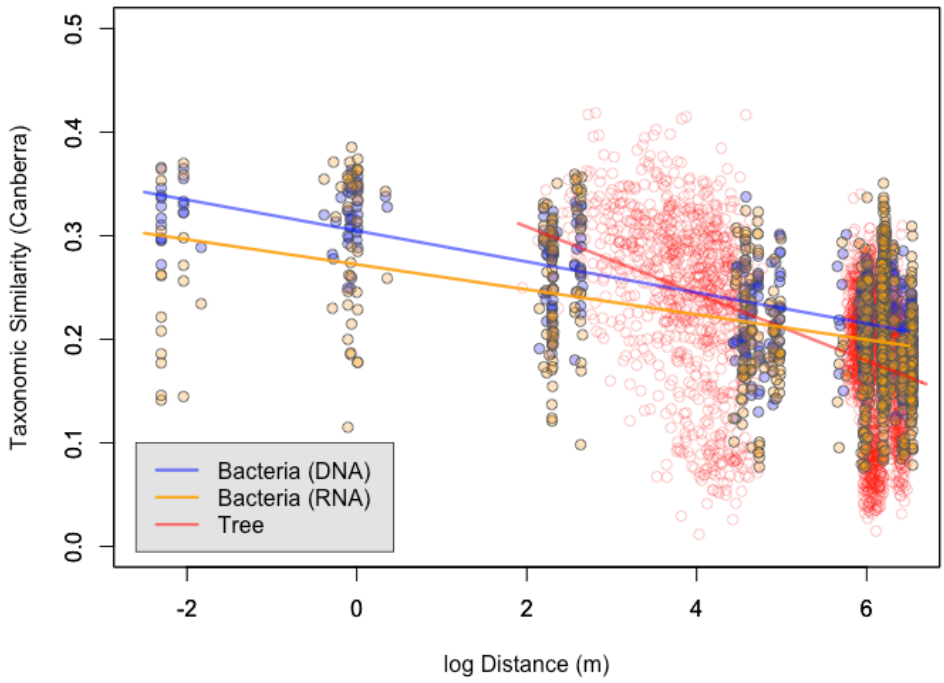


Figure 3: Distance-decay patterns of DNA- and RNA-inferred bacterial communities at the Rabi plot, Gabon.

Spatial scale

Both plant/animal and microbial communities have been reported to have different drivers of biogeographic patterns at different spatial scales (Preston, 1960; Whitaker et al., 2003; Soininen et al., 2007; Bissett et al., 2010; Martiny et al., 2011; Hanson et al., 2012). Studies of microbial biogeography are often conducted at smaller spatial scales than those of plants and animals (although not always (Horner-Devine et al., 2004)), and this could result in differences in the relative

strength of the biogeographic patterns observed. We asked first whether comparing microbial and tree communities at the same spatial scale might account for the discrepancy between tree and microbial distance-decay patterns and second whether there was an alternate spatial scale at which the bacterial distance-decay slope might resemble more closely that of trees. However microbial distance-decay slope across all scales did not significantly differ from the slope derived from the subset of spatial distances shared with trees (difference in slope = 0.0006, p = 0.27, Fig. 4). Thus, when compared at the same spatial scales, the microbial distance-decay slope was still significantly shallower than the tree distance-decay slope (difference in slope = 0.022, $p < 0.001$). At the small (centimeters to meters) scale subset, the distance-decay slope was not significantly different from zero, although it tended to be shallower than the distance-decay slope calculated from the entire dataset. At the largest subset (hundreds of meters) the slope was slightly shallower than the slope derived from the entire dataset (difference in slope = 0.0058, p = 0.001). It has previously been reported that distance-decay rates at smaller spatial scales tended to be lower than those calculated from datasets spanning a larger range of spatial scales (Franklin & Mills, 2003; Morlon et al., 2008; Martiny et al., 2011). Martiny *et al.* (2011) also showed that larger spatial scales tended to exhibit steeper distance-decay slopes than slopes derived from the entire dataset. Although this was not the case for our largest spatial subsets, our largest subset was still at a smaller spatial scale and spanned less spatial scales than their survey. Thus adjusting for differences in scale does not seem to account for the differences in microbial and tree distance-decay slopes in our study.

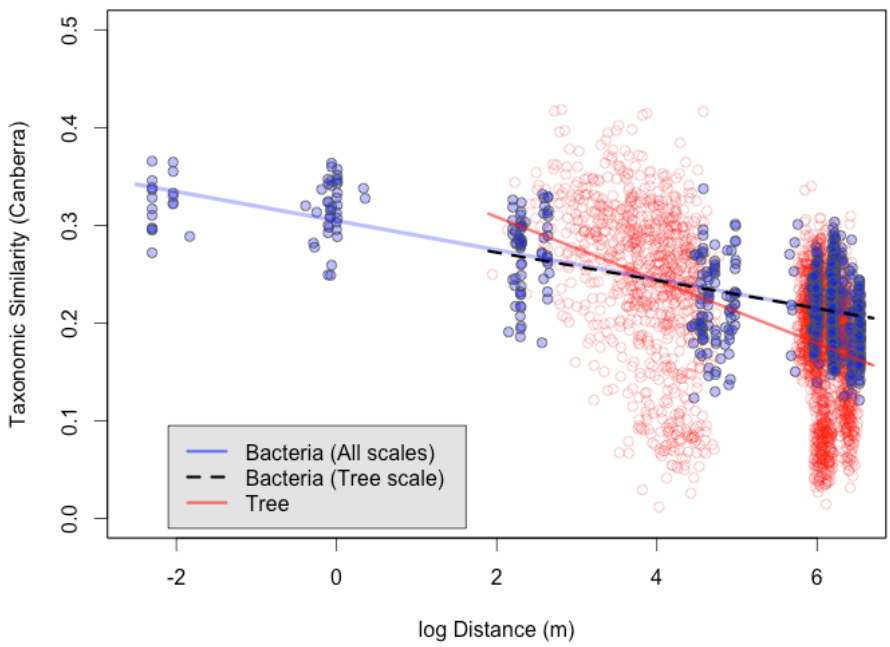


Figure 4: The distance-decay slope of soil bacterial communities considered at spatial subsets.

Sampling effort

Undersampling communities is a problem that exists throughout ecology (Cam et al., 2002). This problem is particularly pronounced in microbial ecology where exhaustively sampling any environment can be impractical if not impossible. In most studies of microbial communities collector's curves are far from saturation, and unique taxa continue to accumulate with increased sampling effort (Rosenzweig, 1995; Woodcock et al., 2006). Undersampling can lead to a weakening of biogeographic patterns if taxa have a positive frequency-abundance relationship (Sloan et al., 2006; Woodcock et al., 2006), whereby abundant community members tend to be more widespread and less abundant taxa tend to be more restricted in distribution. This occurs because undersampling results in decreased detection of

low abundance taxa (with restricted distributions) and the community will thus appear to have less taxonomic turnover across space. Both microbial and plant/animal communities have been reported to have positive frequency-abundance relationships (Thompson et al., 1998; Sloan et al., 2006; Östman et al., 2010), and indeed both the tree and bacterial communities in our study showed such a pattern (Supp. Fig. 2a,b).

We simulated the effects of under-sampling on the distance-decay relationship by using rarefication on both the tree and bacterial communities. For both communities we saw the same trend; the more thoroughly sampled a community was, the steeper the distance-decay rate (Fig. 5a). We then asked whether the effects of sampling effort could account for the differences in the distance-decay slope between trees and microbes. We found that if we sampled the bacterial community as deeply as we could, the distance-decay slope was within the 95% CI range of the tree community when the tree community was dramatically undersampled (Fig. 5b). Therefore if the tree community were as dramatically undersampled as most bacterial surveys tend to be, the tree distance-decay slope would be flattened to within the range of most microbial surveys. This finding is congruent with results reported by Woodcock *et al.* (2006) where it was shown that lower sampling effort could flatten the slope of the taxa-area relationship (although they did not compare these slopes to those of plant or animal communities). Thus sampling effort alone can sufficiently explain the discrepancy between the distance-decay slopes of trees and soil microbial communities in our study.

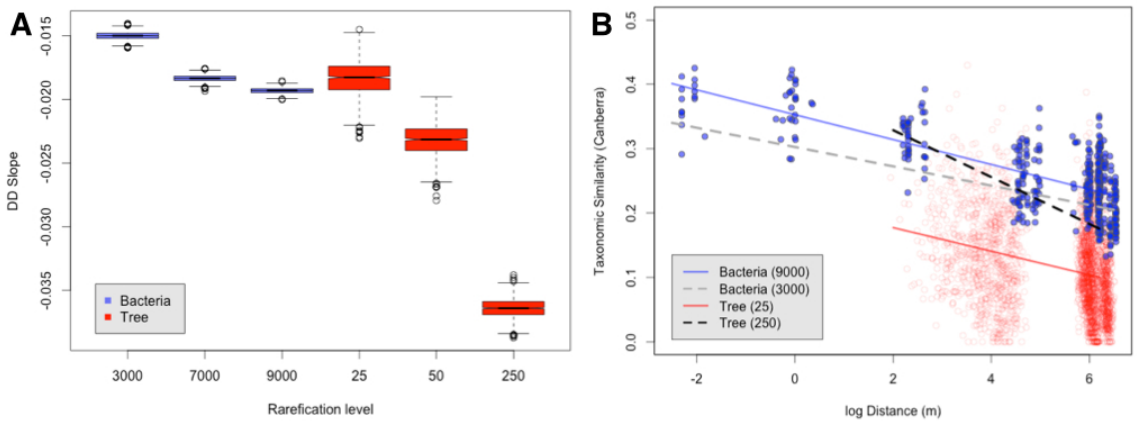


Figure 5: Sampling effort impacts the distance-decay slope in bacterial and tree communities. A) The range of distance-decay slopes derived from different levels of sampling intensity for the bacterial and tree communities. Results shown represent 1000 sampling efforts at each level of rarefication. B) Sampling effort can account for the differences in distance-decay rate between bacteria and trees.

CONCLUSION

Microbial biogeographic patterns are frequently observed to be weaker than the same patterns in plant and animal communities, and a number of explanations have been proposed. Here we demonstrate that in the same landscape, microbial and tree communities showed different community turnover rates (*i.e.* distance-decay relationships with different slopes), with the tree community turnover considerably steeper. We first asked whether the DNA sequence similarity cutoff used to define bacterial taxa had an influence on the rate at which community similarity changed over space. The range of community similarity values changed with the similarity cutoff used, but the slope of the distance-decay relationship did

not, indicating that taxonomic resolution in this system does not account for our observed differences in distance-decay slope. We next tested the hypothesis that microbial distance-decay relationships are flattened by the inclusion of inactive taxa in microbial surveys. Excluding the majority of inactive taxa from our survey by inferring the community *via* RNA (rather than DNA) sequence analysis also did not account for the difference between tree and microbial distance-decay rates. We next asked whether only comparing microbial and tree communities over the same spatial scales could account for the differences in distance decay slope. At the same spatial scales, the two communities continued to show the same level of difference in distance-decay slope as they did at different spatial scales. Finally, we asked whether sampling effort could account for the discrepancy between the tree and bacterial community turnover. We show that by sampling the bacterial community as deeply as our data allow and under-sampling the tree community, we can arrive at two statistically indistinguishable slopes. Hence of all the hypotheses tested, our data seem to only support the idea that sampling effort is driving the differences between the distance-decay relationships of soil microbes and trees in our study system.

Whether our findings are generalizable across other environments, taxonomic groups, or spatial scales remains untested, but since frequency-abundance relationships are so common across taxa it seems likely that the influence of sampling effort on biogeographic patterns will be common across other systems. Our results underpin the importance of deeper sampling if we are to learn about the ecology of endemic microbial taxa. Furthermore, our findings support the

idea that microbial taxa not only qualitatively fit the same biogeographic patterns as plants and animals, but they may do so quantitatively as well. Indeed more intensive sampling efforts of microbial taxa may reveal that the spatial scaling of microbial diversity is not so fundamentally different from that of plants and animals.

BRIDGE

In Chapter II my results indicate that microbial communities may not necessarily scale differently in space than tree communities if we equalize sampling efforts across communities. This suggests that principles in the spatial ecology of macro-organisms (*e.g.* the decay of community similarity over space or the increase in species richness with area) not only qualitatively apply to microbial communities, but they may do so quantitatively as well. My data also demonstrate that soil microbial communities in tropical rainforests tend to show a positive frequency-abundance relationship, whereby abundant taxa tend to be more widespread spatially and low abundance taxa tend to have more spatially-restricted ranges. Incorporating spatial concepts such as these has played an instrumental role in conservation efforts and management strategies of natural reserves. Low abundance taxa with geographically-restricted ranges are generally considered to be threatened in fragmented habitats or environments undergoing change since they tend to be more ecologically specialized and prone to stochastic extinction events due to their low abundance. Thus what Chapter II suggests, but does not directly address, is that numerous low abundance (spatially restricted) taxa in rainforest soils could be vulnerable to the effects of environmental change. In the

following chapter I explore the concept of microbial responses to environmental change in a similar environment to Chapter II. The area where I conducted my study is in a state of rapid conversion to agriculture and thus contains a mosaic of sites representative of the land conversion process. I selected sites in this area that were in close proximity to one another and that were representative of each of the primary stages in the conversion of forest to agriculture (*i.e.* a forest site, a recently burned site, a plantation site, and an abandoned plantation site). In this study I look for indications of biotic homogenization in the soil microbial communities and I distinguish the physiologically active taxa from the dormant taxa to test the hypothesis that the active fraction of the community is more susceptible to the impacts of environmental change.

CHAPTER III

RESPONSE OF SOIL BACTERIAL COMMUNITIES TO CONVERSION OF CONGO BASIN RAINFOREST TO AGRICULTURE

This chapter is co-authored by myself and Brendan Bohannan. I primarily performed the sample processing, data analysis, and writing. Brendan Bohannan filled the advisory role on this project and as such aided in the conception of the design and the sample collection in the field, as well as contributed edits to the text. Supplementary material for this chapter can be found in Appendix B.

INTRODUCTION

Land use change is occurring disproportionately faster in the tropics relative to the rest of the world (Houghton, 1994; Dirzo & Raven, 2003). The primary motivation for this change is agricultural development (Foley et al., 2005; Laurance et al., 2014). The tropical ecosystems under threat tend to be exceptionally diverse and play critical roles in global biogeochemical cycles. Ecosystem conversion for agriculture has been shown to have a pronounced impact on both above- and below-ground biotic communities (Hooper et al., 2012; Rodrigues et al., 2013; Mueller et al., 2014) and understanding the response of communities to such environmental changes is a global priority (Cardinale et al., 2012). Understanding the response of microbial communities to ecosystem conversion is especially challenging. Microbial communities are diverse, abundant and, because most microbial taxa cannot be easily distinguished morphologically or through laboratory culture, they must be

studied indirectly through methods such as metabarcoding (*i.e.* surveying taxa by surveying gene regions amplified from environmental samples; Pace, 1997). Such methods often do not distinguish between actively growing and dormant or dead individuals, which is problematic because soil microbial communities are characterized by high levels of dormancy (Lennon & Jones, 2011; Blagodatskaya & Kuzyakov, 2013). To date, there has been little effort to compare microbial responses to land use change across the tropics, and even less effort to investigate the response of the active fraction of the community to environmental change.

Of the numerous studies documenting microbial responses to tropical land use change, few have focused on Africa (Laurance et al., 2014). Most studies have focused instead on the neotropics-especially the Amazon Basin. While these regions both contain large tracts of tropical rainforest and share a relatively similar climate, they differ in numerous ways such as the predominant forms of agriculture (Laurance et al., 2014), land management practices (Naughton-Treves & Weber, 2001), and geologic history (Livingstone, 2001) including parent material of the soil. Each of these factors has been shown to alter soil microbial community composition (Drenovsky et al., 2004; Berg & Smalla, 2009; Lauber et al., 2013; Ofek et al., 2014). Thus, in order to gain a broader understanding of biotic responses to tropical environmental change, it is imperative that our studies be expanded to other regions of the tropics such as Africa.

Tropical deforestation has been linked to a number of pronounced changes to soil microbial communities and microbially-mediated ecosystem functions. In the Amazon, this includes a loss of endemic taxa and a decrease in community spatial

turnover (Rodrigues et al., 2013; Hamaoui et al., 2016; Navarrete et al., 2015b), as well as various alterations to ecosystem functions including those in the nitrogen and methane cycles (Neill et al., 1997b, 2005; Fernandes et al., 2002). We are aware of only two studies to focus on microbial responses to African land use change (Bossio et al., 2005; Sul et al., 2013), and both reported alterations to microbial community composition following land use change and/or implementation of different agricultural management practices. Each of these studies utilized DNA-barcoding to survey microbial communities, and it is well established that DNA-based surveys do not distinguish active cells from inactive (dormant or dead) cells. Thus it is unclear from these studies whether the changes they report are occurring uniformly across all taxa present, or whether the active fraction of the community is disproportionately impacted. Making this distinction could be important for understanding how environmental change impacts microbially-mediated functions, since the active fraction of the community is more intimately interacting with the environment and contributing directly to ecosystem function.

We asked whether the responses of soil microbial communities to land use change in the Central African nation of Gabon were similar to those reported for other tropical regions. We used the distance-decay of community similarity to approximate spatial turnover (Nekola & White, 1999) and estimated alterations to the spatial ecology of soil microbial communities following land use change. We surveyed the community *via* RNA in order to distinguish the active fraction of taxa from the inactive fraction. We also surveyed communities *via* DNA in order to directly compare our findings to past studies. Gabonese forests are included in the

Guineo-Congolian regional center of endemism, making these ecosystems relatively representative of the Congo Basin at large (White, 2001). We chose sites that are broadly representative of the slash-and-burn cycle in Central Africa (Laurance et al., 2006), including a forested plot, a plot that had very recently been deforested and burned, an established manioc/banana plantation plot, and a limited number of samples from an abandoned plantation. We tested the following hypotheses: 1), converted sites will have numerous indications of biotic homogenization in the soil microbial community 2) the recently burned site will have a more pronounced change in microbial community diversity patterns than the plantation site, and 3) the community surveyed *via* RNA (*i.e.* the more physiologically active fraction of the community) will exhibit a stronger response to land use change than the community surveyed *via* DNA (the "total" community).

MATERIALS & METHODS

Sampling site

Central Africa contains up to 1.8 million km^2 of contiguous tropical moist forest, making it the second largest block of tropical moist forest in the world, after the Amazon Basin (Wilkie & Laporte, 2001). Central African rainforest is renowned for its exceptionally high levels of biodiversity and endemism (White, 2001; Lee et al., 2006; Butler & Laurance, 2008) and is rapidly being deforested and transformed (Naughton-Treves & Weber, 2001). The nation of Gabon contains more than 10% of the contiguous tropical moist forest in Africa (Wilkie & Laporte, 2001; Lee et al., 2006), and the majority of these forested areas are either currently leased as long-

term logging concessions or are at risk from agricultural conversion (Collomb et al., 2000; Laurance et al., 2006; Lee et al., 2006).

Our study was performed in southwestern Gabon near the Gamba Complex of Protected Areas (Lee et al., 2006). Agricultural conversion in this region follows the typical slash-and-burn practices of most tropical regions whereby forests are selectively logged and the remaining vegetation is burned. The following season, the plantation crops (typically manioc, yucca, or banana) are planted and harvested for 1-3 years. Following the last harvest, plantations are abandoned and secondary forest develops. We selected sites representative of this cycle including a recently burned site, an active manioc and banana plantation, an abandoned plantation, and an adjacent intact forest. Including a burned plot allowed us to better discern the impact of ecosystem conversion (*i.e.* logging and burning) from the impact of agriculture (planting and managing cropland). We chose to survey one site across each of four land types, rather than performing higher levels of replication on fewer land types. The sites we selected were in close proximity (all within <6 km of each other) to one another, allowing us to minimize the confounding effects of distance on community comparisons. The soil type (entisol) was consistent across our sites and tended to be sandy in texture - which is typical of Southwestern Gabon (Delègue et al., 2001). Sites are found at the following coordinates: burned site (2° 44' 48" S, 10° 8' 54" E), plantation (2° 44' 58" S, 10° 8' 51" E), abandoned plantation (2° 45' 39" S, 10° 8' 49" E), and adjacent forest (2° 44' 46" S, 10° 8' 52" E).

Sampling Design

Soil samples were taken at the end of the Gabonese dry season (September 24-27, 2013). We established plots within each of the aforementioned sites. Each plot consisted of a nested sampling scheme (Rodrigues et al., 2013) where a 100-m^2 quadrat was established with 10-m^2, 1-m^2, 0.1-m^2, and 0.01-m^2 quadrats nested within. Soil cores were taken to a depth of 15 cm (after removal of leaf litter) from the corners of each quadrat (N=13 samples per site, N=4 samples in abandoned plantation). For each point, 3 cores were taken, homogenized, and then subsampled. From the homogenized mixture, 3 ml (approximately 1 g) of soil was added to 9 ml Lifeguard solution (Mobio, California, USA) in the field, then shipped cold and stored at -80° C in order to stabilize nucleotides for later extraction. Our spatially explicit design allows for the estimation of species richness (local and regional) as well as spatial turnover (beta diversity)(Anderson et al., 2011).

Extraction, PCR, and Sequencing

Soil DNA and RNA were co-extracted using MoBio's Powersoil RNA Isolation kit with the DNA Elution Accessory Kit (MoBio, California, USA) following manufacturer's instructions. Extractions were quantified using Qubit (Life Technologies, USA). RNA was reverse transcribed to cDNA using Superscript III first-strand reverse transcriptase and random hexamer primers (Life Technologies, USA).

The V3 and V4 region of the 16S rRNA gene of the DNA and cDNA were PCR amplified using the primers 319F and 806R (primarily targeting Bacteria and limited coverage of Archaea). Sequencing libraries were prepped using a 2-step PCR

with dual-indexing approach (Kozich et al., 2013; Fadrosh et al., 2014). In short, the first round of amplification consisted of 22 cycles with Phusion HiFi polymerase. Round 1 products were cleaned using Agencourt AMPure XP (Beckman Coulter, California, USA) then amplified for an additional 6 cycles using Phusion HiFi to add the sequences required for cluster formation on the Illumina flowcell. The final library was sent to the Dana-Farber Cancer Institute Molecular Biology Core Facilities for 300 PE sequencing on the Illumina MiSeq platform.

Soil Chemical Analysis

Soil chemical parameters were measured from each soil core to estimate the impact on microbial community composition by A & L Western Agricultural Lab (Modesto, CA, USA). In total, percent organic matter (loss on ignition (Dean, 1974)), extractable phosphorus (Weak Bray (Kamprath & Watson, 1980) & sodium bicarbonate (Olsen, 1954)), extractable cations (K, Mg, Ca, Na), sulfate-S (Fox et al., 1964), pH, buffer pH, cation exchange capacity (CEC, (Chapman, 1965)), and percent cation saturation were measured. Pearson's correlation tests were performed on all pairs of chemical parameters to test for autocorrelation and reduce the number of chemical variables used in our models. Pairs of variables that were highly correlated ($R^2 > 0.6$, $P < 0.05$) were reduced to a single variable.

Data processing and statistical analysis

Paired end reads were joined then demultiplexed in QIIME (Caporaso et al., 2010) before quality filtering. Primers were removed using a custom script.

UPARSE was used to quality filter and truncate sequences (416bp, EE 0.5) (Edgar, 2013). Sequences were retained only if they had an identical duplicate in the database. Operational taxonomic units (OTUs) were clustered *de novo* at 97% similarity using USEARCH (Edgar, 2010). OTUs were checked for chimeras using the gold database in USEARCH. We used a custom script to format the UCLUST output for input into QIIME. To assign taxonomy, we used repset from UPARSE in QIIME using greengenes version 13_5 (RDP classifier algorithm). Finally, we rarefied to 3790 in QIIME to achieve approximately equal sampling depth, which excluded 4 samples.

Statistical analyses were performed in the R platform (R Development Core Team, 2010). Canberra pairwise community distances were calculated using the vegdist function in the package 'vegan' (Oksanen et al., 2015). Canberra was chosen because of its incorporation of abundance data and sensitivity to rare community members (Jost et al., 2011). Ordinations were created using non-metric multidimensional scaling (NMDS) using the function metaMDS in the 'vegan' package. Vectors of environmental variables were projected in the ordinations using the envfit function in the 'vegan' package (Oksanen et al., 2015) using only significantly-associated ($P < 0.05$) variables. OTU richness was used for the estimation of alpha diversity; however results were qualitatively similar using the Shannon or Simpson indices. The Chao1 estimator was used to approximate gamma (*i.e.* landscape-level) taxonomic richness- error bars indicate standard error measurements (Chao et al., 2009). Turnover rates of each land type were estimated by regressing pairwise similarity (1- Canberra distance) against pairwise

geographic distance between samples (Nekola & White, 1999). We excluded the abandoned plantation from this analysis because of the low number of samples. Mantel tests were used to test for significant associations between geographic and community distance and partial Mantel tests were used to estimate the relative contribution of environmental heterogeneity and geographic distance on variation in community dissimilarity in the 'vegan' package in R. Distance-decay slopes were compared using the function diffslope (package 'simba'), which calculates the difference in slope and employs a random permutation approach to calculate a *P*-value (Jurasinski & Retzer, 2012). The *P*-values computed are the ratio between the number of cases where the differences in slope exceed the difference in slope of the initial configuration and the number of permutations. Figures were either created using base R or the 'ggplot2' package (Wickham, 2009).

In order to identify differentially active taxa, we used the DESeq2 function (Love et al., 2014) as implemented in QIIME (Caporaso et al., 2010). Low abundance samples were excluded prior to performing the analysis and the abandoned plantation was excluded because of its limited number of samples. Our application of this function tests for enrichment of taxa in the RNA-inferred community relative to the DNA-inferred community in each land type using an estimate of dispersion and log-fold change from a negative binomial generalized linear model. Taxa that were deemed differentially active (*i.e.* those with a positive log-fold change in the RNA and $P_{adj} < 0.05$) were then subsetted from the rarefied DNA-inferred community matrix in order to compare patterns in their abundance, diversity, and spatial patterning across land types. To approximate the contribution of soil

chemical variables in structuring the differences in composition of the differentially active taxa, community matrices were first Hellinger transformed, then tested against all soil chemical variables using a redundancy analysis (RDA) using the functions decostand and rda, respectively, in the 'vegan' package. The most explanatory model was selected using stepwise model selection based on permutation tests using the function ordistep in the 'vegan' package.

RESULTS

The DNA-inferred community is altered by land use

Soil microbial community composition differed among land types ($F_{3,39}$ = 2.13, R^2 = 0.15, P <0.001). The four land types (forest, burned, plantation, and abandoned plantation) clustered separately in NMDS space (Fig. 1a); the burned and plantation sites, however, showed more overlap relative to the other sites, indicating higher taxonomic similarity. Various soil chemical attributes were significantly associated with differences in community structure. For example, forest communities tended to be associated with levels of soil organic matter (OM), which were higher in forest soils than the converted sites (Supplementary Table 1). The burned communities tended to be associated with nitrate-nitrogen (NO_3-N) and sulfur (S)- both of which tended to be highest in the burned site. The plantation communities tended to be associated with phosphorus (P) and potassium (K), both of which were highest in the plantation site. Finally, the abandoned plantation communities tended be associated with elevated magnesium (Mg) levels. The forest-to-burned transition showed a decrease in the relative abundance of the phyla

Acidobacteria, *Proteobacteria, Firmicutes*, and *AD3*, as well as an increase in *Actinobacteria* (Supp. Figure 1a). The burned-to-planted conversion was associated with a decrease in *Actinobacteria,* back to within the range of relative abundance of the forest samples, as well as an increase in *Acidobacteria* back to the forest levels. The forest site had a higher number of endemic taxa (400) relative to the burned site (~175) and the plantation site (~300, Supp. Fig. 2a). The taxa that were unique to the forest tended to be much more restricted in range relative to the taxa that were shared between the forest and burned sites or between the forest and plantation sites (Supp. Fig. 2b).

Alpha diversity (OTU richness within samples) varied by land type (One-way ANOVA $F_{3,39} = 4.9$, $P < 0.01$, Fig. 1b) with the most pronounced change being a loss of richness in the burned site relative to the other sites (Tukey's HSD $P_{adj} < 0.05$). Levels of alpha diversity were statistically indistinguishable between the forest, plantation, and abandoned plantation sites. The same trend was found for gamma (*i.e.* landscape-level) diversity (Fig. 1c), where the forest site showed higher estimated OTU richness (1891 ± 61 OTUs) than the burned site (1621 ± 63 OTUs), and the plantation site showed levels of diversity similar to the forest (1845 ± 68 OTUs).

Levels of pairwise similarity among forest microbial communities exhibited the strongest relationship with spatial distance of all sites (Mantel r = 0.78, $P < 0.01$), as well as the steepest distance-decay slope (slope = -0.02 ± 0.002, $P < 0.001$, Fig. 1d). Levels of community similarity in the burned site had a weaker relationship with spatial distance (Mantel r = 0.41, $P = 0.05$) as well as a significantly weaker

distance-decay slope relative to the forest (slope = -0.004 ± 0.001, difference in slope = 0.016, P = 0.001). The relationship between community similarity and distance was also weaker in the plantation (Mantel r = 0.39, P = 0.05) relative to forest and showed a weaker distance-decay slope compared to the forest site (slope = -0.007 ± 0.001, difference in slope = 0.013, P = 0.001). The distance-decay slope of plantation communities was slightly steeper than that of the burned site (difference in slope = 0.003, P = 0.05).

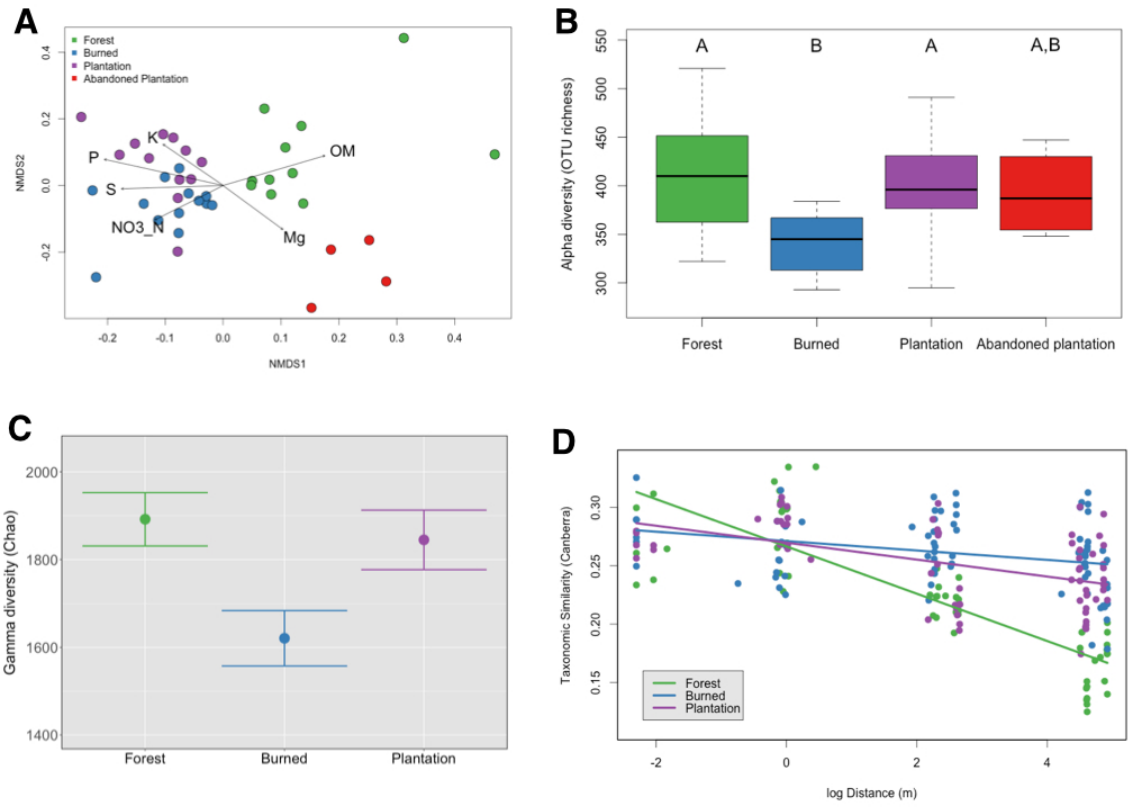


Figure 1: Land use change impacts soil microbial community diversity in the Central African nation of Gabon. A) Microbial community composition clusters by land use, non-metric multidimensional scaling (NMDS) of Canberra dissimilarities

showing significant ($P < 0.05$) environmental vectors. OM: organic matter, Mg: magnesium percent saturation, NO_3-N: nitrate nitrogen, S: sulphate-sulfur, P: weak bray phosphorus, and K: potassium percent saturation. B) OTU richness across conversion sites, significant differences assessed using Tukeys HSD of one-way ANOVA. C) Gamma (landscape-level) diversity from Chao1 estimator with standard error bars. D) The rate of distance-decay of community similarity (1- Canberra dissimilarity) significantly differs ($P < 0.05$) by land use. Analyses for the abandoned plantation were excluded from panels C and D due to lack of statistical power from low sample numbers.

The response of the RNA-inferred community to land use is stronger than that of the DNA-inferred community

We distinguished active microbial taxa from the total pool of microorganisms using two techniques: 1) by inferring community structure *via* RNA-barcoding of 16S rRNA, and 2) by identifying the members that are differentially abundant in the RNA-inferred community relative to the DNA-inferred community. The former assumes that by surveying the community *via* RNA-barcoding, we are enriching for the taxa that are actively growing and transcribing the rRNA gene (heretofore referred to as the *active community*), while the latter distinguishes the taxa that are differentially active (heretofore referred to as the *differentially active taxa*) given their abundance in the RNA pool relative to the DNA pool.

The changes to the active (RNA-inferred) community across our sites were qualitatively similar to the changes observed in the DNA-inferred community-

although the effect size of changes tended to be stronger in the active community. Taxonomic membership of the active community varied significantly across the four sites ($F_{3,38}$ = 2.25, R^2 = 0.15, P < 0.001). There was also slight variation in community structure between the active community and the DNA-inferred community ($F_{1,38}$ =1.01, R^2=0.05, P < 0.001). The four sites clustered distinctly in NMDS space (Supp. Fig. 3a). The forest-to-burn conversion was associated with a decrease in the relative abundance of the phyla *Actinobacteria*, *Acidobacteria*, and *Firmicutes* as well as an increase in *Proteobacteria*, while the burned-to-plantation transition resulting in the reverse of all four of these phylum-level changes (Supp. Fig. 1). The compositional shifts in the active community across sites were predicted by the same soil chemical factors as the DNA-inferred community, but with two additional significant predictors: cation exchange capacity (CEC; associated with the burned site communities), and soil pH (associated with the plantation site communities). The alpha diversity (OTU richness) of the active community varied significantly by site ($F_{3,38}$ = 9.1, P < 0.001), and was significantly lower in the burned site relative to the other sites (Tukey's HSD on one-way ANOVA P < 0.001), which were indistinguishable from one another (Supp. Fig. 3b). The gamma diversity followed a similar qualitative trend to the DNA-inferred community (*i.e.* a loss of diversity in the burned site and some recovery in the plantation, Supp. Fig. 3c), but with several exceptions: 1) the diversity estimate for the forest site was higher in the active community than in the DNA-inferred community (2057 ± 76 OTUs *versus* 1891 ± 61 OTUs), 2) the magnitude of change between the forest and burned sites was larger in the active community relative to the DNA-inferred community (a loss of ~654

active OTUs *versus* a loss of ~270 OTUs), and 3) the estimated richness of the plantation site was below the range of forest richness (1739 ± 56 OTUs) and hence did not fully "recover" to the level of richness in the forest as it did in the DNA-inferred community.

The forest active community showed a significant distance-decay trend in community similarity over space (Mantel r = 0.84, P = 0.001, slope = -0.021 ± 0.001, Supp. Fig. 3d). Similar to the DNA-inferred community, this distance-decay pattern was notably weaker in the burned site (slope = -0.003 ± 0.001), so much so that the relationship was no longer significant (Mantel r = 0.27, P = 0.11). The magnitude of the change in distance-decay slope between the forest and burned sites was slightly larger in the active community than the DNA-inferred community (Difference in slope = 0.0174 *versus* 0.0163). Community similarity showed a marginally significant distance-decay relationship in the plantation (Mantel r = 0.40, P = 0.056, slope = -0.008 ± 0.001), and this slope was significantly steeper than the burned site (difference in slope = 0.004, P =0.012) and significantly weaker than the forest site (difference in slope = 0.013, P = 0.001). The average distance to centroid – a measure of beta diversity- tended to decrease across the chronosequence ($F_{3,38}$ = 3.14, P = 0.036, Supp. Fig. 4). This trend was qualitatively-similar in the DNA-inferred community, but not significant ($F_{3,36}$ = 2.63, P = 0.065).

When we removed the effect of geographic distance, partial mantel tests showed that DNA-inferred communities in the forest were more similar under similar environmental conditions (Mantel r = 0.52, P = 0.055, Table 1), and when we removed the effect of environmental similarity, the effects of geographic distance on

community similarity were no longer significant (r = 0.52, *P* = 0.302). The opposite was the case for the active community, showing a significant relationship with distance after removing the effect of environmental similarity (r = 0.65, *P* = 0.007), but not vice versa (r = 0.31, *P* = 0.143). The active communities in the burned site tended to be more similar under similar environmental conditions once we removed the effect of geographic distance (Mantel r = 0.52, *P* = 0.004), and active community similarity was no longer correlated with geographic distance after removing the effect of environmental similarity (r = 0.16, *P* = 0.198). Neither relationship was significant for the DNA-inferred community in the burned site. Lastly, community similarity was related to environmental similarity (after removing the effect of geographic distance) in the plantation for both the active and the DNA-inferred community; however the relationship was stronger for the active community (r_{DNA} = 0.41, *P* = 0.025, r_{RNA} = 0.63, *P* = 0.001). After removing the effect of environmental similarity in the plantation, the distance-decay relationship was no longer significant for the DNA-inferred community or the active community.

Table 1: The influence of geographic distance and habitat heterogeneity on DNA-inferred and RNA-inferred bacterial communities. Partial mantel test summary statistics showing 1) the effect of environmental similarity after removing the effect of geographic distance and 2) the effect of geographic distance after removing the effect of environmental similarity. *P* values estimated from 1000 permutations.

	Env. Simil.		**Geog. Dist.**	
	r	*P*	r	*P*
Forest DNA	**0.52**	**0.055**	0.12	0.302
Forest RNA	0.31	0.143	**0.65**	**0.007**
Burned DNA	0.24	0.146	0.34	0.07
Burned RNA	**0.52**	**0.004**	0.16	0.198
Plantation DNA	**0.41**	**0.025**	0.21	0.192
Plantation RNA	**0.63**	**0.001**	0.11	0.252

The differentially active taxa increase in abundance and richness in impacted sites

We distinguished the differentially active taxa by testing for enrichment of OTUs in the RNA-inferred community relative to the DNA-inferred community in each land type. Taxa that were deemed differentially active (*i.e.* those with a positive log-fold change in the RNA and $P_{adj} < 0.05$) were then subsetted from the rarefied DNA-inferred community matrix. The proportion of differentially active taxa in the DNA-inferred community differed by site ($F_{2,33} = 24.4$, $P < 0.001$), with the forest and burned sites showing significantly lower proportions of differentially active taxa relative to the plantation (Tukey's HSD $P < 0.001$, Fig. 2a). Of these taxa, 22 (11%) were shared between all sites, 18 (9%) were unique to the forest, 12 (6%)

were unique to the burned site, and 46 (23%) were unique to the plantation site (Fig. 2a). The average within-sample alpha diversity of the differentially active community was significantly less diverse in the forest than the burned and plantation sites ($F_{2,33}$ = 46.7, P < 0.001, Tukey's HSD P < 0.001 for each site comparison, Fig. 2b). The variability in the community composition of differentially active taxa in the forest was best explained by soil cation exchange capacity (CEC, $F_{1,11}$ = 2.2, P = 0.005), while the compositional variation of the differentially active taxa in the plantation was best explained by soil CEC and pH ($F_{1,11}$ = 2.2, P = 0.05, $F_{1,11}$ = 2.4, P = 0.005, respectively). Compositional shifts in the differentially active burned site community were not related to any soil chemical factors. The forest-to-burned transition consisted of an increase in the relative abundance of the phyla *Actinobacteria* and *Acidobacteria* as well as a decrease in *Planctomycetes* (Supp. Fig. 5). The burned-to-planted transition consisted of a decrease in the relative abundance of the phyla *Actinobacteria* (back to within the range of the forest) and an increase in *Planctomycetes*.

The differentially active community in the forest exhibited significant spatial turnover (Mantel r = 0.72, P = 0.002, slope = -0.028, Fig. 2c). The rates of spatial turnover of the differentially active community in the burned and plantation sites were substantially weaker (slope = -0.003 & -0.008, respectively), and neither of these relationships were significant (Mantel r_{burned} = 0.12, P = 0.26, Mantel $r_{plant.}$ = 0.12, P = 0.27).

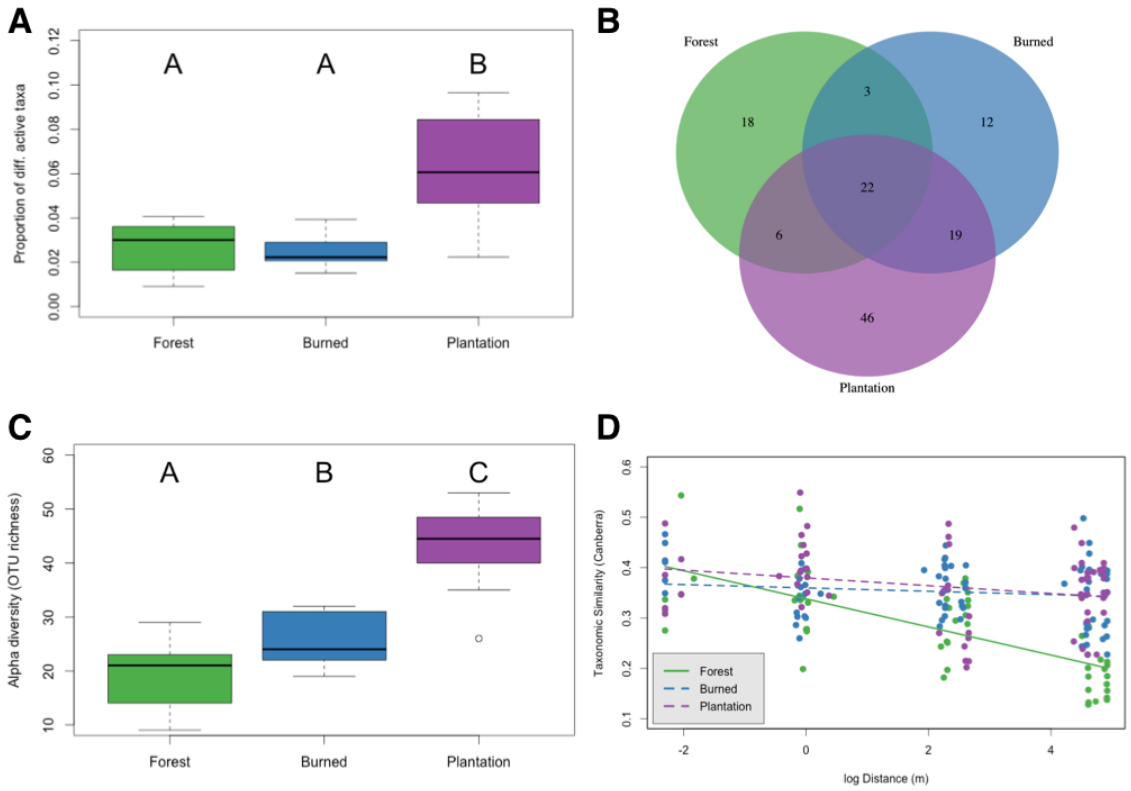


Figure 2: The abundance, composition, and diversity of the differentially active taxa vary by land use. A) The proportion of differentially active taxa in the DNA-inferred is highest in the plantation B) Plantations have the highest number of unique differentially active taxa, C) the within-sample OTU richness of the differentially active taxa is highest in disturbed (burned and plantation) sites, and D) the distance-decay of community similarity of the differentially active community is strongest in the forest, dashed lines indicate a non-significant trend ($P > 0.05$).

DISCUSSION

Land use change is occurring at a disproportionately faster rate in the tropics than in other parts of the world (Houghton, 1994; Dirzo & Raven, 2003; Laurance et al., 2014). Understanding biotic responses to land use change in these highly diverse regions is a priority if we are to predict and/or mitigate changes to these communities. While numerous studies have estimated microbial responses to land use change in the neotropics (particularly in the Amazon Basin) few have focused on changes happening in the African tropics. My results indicate that bacterial communities in Central African rainforest soils respond strongly to ecosystem conversion. We documented biotic homogenization in response to conversion, as well as other changes to bacterial community attributes. Furthermore, we observed that the active fraction of the community (those taxa detected using RNA metabarcoding) tended to respond more strongly to ecosystem conversion. Moreover, if we distinguish the differentially active taxa- the taxa that are likely to be highly active- we see an additional story emerging; namely that the highly active taxa tend to increase in proportion and richness in impacted sites, but decrease in spatial diversity.

Biotic homogenization can take many forms including changes to diversity, life history strategies, and the spatial patterning of taxa, as well as a loss of endemics (Olden & Poff, 2003; Olden et al., 2004; Smart et al., 2006). These trends have been reported for soils in the Amazon Basin undergoing agricultural transformation. Here we report that these responses also appear to be elicited under African land use change. While it is not necessarily surprising that converting a rainforest to

agriculture adversely impacts the native biota, testing whether microbial responses follow a similar trend across other regions of the tropics is an important step in predicting the effects of environmental change. Rodrigues *et al*. (2013) reported a weakening of microbial spatial patterning (namely a decrease in the distance-decay rate) in Amazonian cattle pasture relative to nearby primary rainforest. Our data show a similar trend: the distance-decay of community similarity has a significantly shallower slope in the converted sites (burned and planted) relative to the nearby rainforest. This trend could be driven by several alterations to the landscape. On the one hand, deforestation could be facilitating higher rates of microbial dispersal at the landscape level, decreasing the rate at which communities differentiate over space. On the other hand, the conversion process could be homogenizing the soil environment, which in turn could be homogenizing the microbial communities. My data provides evidence for the latter explanation: environmental distance-decay rates followed the same trends as the communities (*i.e.* the forest soil exhibited the highest rate of distance-decay and the burned and plantation soils were significantly lower). Our partial mantel tests also suggested that variation in communities in the burned and plantation site was largely driven by environmental variation.

Our design allowed me to distinguish the impact of ecosystem conversion (slash-and-burn) from plantation management, giving us a number of new insights from this comparison. In nearly all cases, the burned site appeared to be more impacted than the plantation site. For example, the distance-decay slope was much shallower in the burned site than in the plantation site. Species richness both within samples and at the landscape-level was also lowest in the burned site. This

observation suggests that the act of slash-and-burn conversion may be stronger than the impact of planting and growing crops. Rodrigues *et al.* (2013) showed an increase in microbial species richness in cattle pasture relative to the rainforest. Forest and plantation microbial richness were indistinguishable in our study. This difference could be specific to the form of agriculture. Amazonian cattle pasture are planted with African grasses such as *Urochloa* (Boddey et al., 2004), which provide the soil environment with high levels of labile carbon in the form of root exudates (Dias-Filho et al., 2001). Gabonese plantation crops tend to be less dense, with large patches of exposed soil in between individual plants, which could reduce the belowground carbon input from agriculture (and indeed we did not observe an increase in soil carbon in the plantations, unlike the Amazonian pastures). Hence the increase in richness under cattle pasture could be driven by higher plant density and/or increased belowground productivity. My study suggests that ecosystem conversion (as opposed to planting) has a greater impact on the soil bacterial community, and that the impacts of planting and agricultural management on the soil bacterial community may be crop-specific.

There have been a number of reports of changes in soil bacterial community composition in response to land use change (Bossio et al., 2005; Jesus et al., 2009; Rodrigues et al., 2013; Alele et al., 2014; Navarrete et al., 2015b, 2015c). One notable change reported from cattle pasture conversion sites in the Amazon is a shift in the dominance of oligotrophic taxa such as *Acidobacteria* to copiotrophic taxa such as *Firmicutes* and *Actinobacteria* (Rodrigues et al., 2013; Navarrete et al., 2015b, 2015c). Our RNA and DNA-inferred communities both show a decrease in

Acidobacteria from the forest to the burned site, followed by a recovery in the plantation. The recovery in the Gabonese plantation system (where levels of organic matter decrease) and not in the Brazilian pasture system (where carbon increases) suggests that carbon may be playing a role in these shifts. Navarrete et al. (2015b) show that in a DNA-inferred community, *Actinobacteria* increase in relative abundance in converted cattle pasture in the Amazon. Interestingly we see the same pattern in our DNA-inferred community, but see the opposite trend in the RNA-inferred community. *Actinobacteria* are known spore-formers and could therefore be tolerating the ecosystem conversion in a dormant state but not necessarily be increasing in activity. Thus it appears that some of the compositional shifts following conversion may be generalizable across tropical regions and that this may be driven by a common response to changes in soil organic matter.

Distinguishing active taxa from the total pool of community members is an important step toward linking community structure and ecosystem function (Nannipieri et al., 2003), especially in environments that contain high proportions of inactive taxa (Lennon & Jones, 2011). Since active members are more likely to interact with their environment, our expectation was that environmental change would elicit a stronger response in the active (RNA-inferred) community than in the DNA-inferred community (which includes inactive taxa). We thus hypothesized that the active community would exhibit larger changes to diversity and show trends that were undetectable in the total pool. Nearly all changes to diversity showed similar qualitative trends between the active (RNA-inferred) community and the DNA-inferred community. This included a separation of community structure by

land type, a loss of alpha diversity in the burned site, a loss of gamma (landscape) diversity in the burned site, as well as a weakening of the community distance-decay rate in the burned and plantation sites. The soil chemical factors structuring the active community were also largely similar to those structuring the DNA-inferred community, which has been shown to be the case in northern hardwood forest soils as well (Romanowicz et al., 2016). From a quantitative perspective, the magnitude of change tended to be larger in the active community than the DNA-inferred community. For example, the change of alpha diversity across land types exhibited a higher effect size in the active than the DNA-inferred community ($F_{3,38}$ = 9.1 in RNA *versus* $F_{3,39}$ = 4.9 in DNA). The difference in community distance-decay slope across sites was also slightly larger in the active community than the DNA-inferred community (Difference in slope = 0.0174 in RNA *versus* 0.0163 in DNA in the burned site). Moreover, gamma diversity had a higher proportional change across land types in the active community than the DNA-inferred community. The active community tended to show a decrease in beta diversity (average distance to centroid); while this was qualitatively similar in the DNA-inferred community, it was only statistically significant in the active community. Previous work has shown that the active community was more responsive to soil drying and rewetting than the DNA-inferred community (Barnard et al., 2013). Work on aquatic microbial communities in the South China Sea has also shown that variation in the active community is more predictable by environmental variables, suggesting that the active community is slightly more sensitive to the environment relative to the DNA-inferred community (Zhang et al., 2014). Thus while changes to the active

community in our system may be qualitatively similar to changes in the DNA-inferred community, the changes tend to be more pronounced in the active community, and this is consistent with reports from other systems.

It has been suggested that rRNA may not be a reliable indicator of microbial activity because cellular concentrations of rRNA do not scale with growth rate uniformly across lineages, and that certain taxa may be producing rRNA even in relative states of metabolic inactivity (Blazewicz et al., 2013). Although we were conservative in our inferences about how changes to the RNA-inferred community might translate to activity alterations, we took our analysis one step further by using a probabilistic model to distinguish the differentially active taxa. This approach allowed us to identify taxa that were overly enriched in the RNA given their abundance in the DNA. Using this approach we were able to detect new trends and develop future hypotheses regarding how environmental change impacts microbial communities. The proportion of differentially active taxa tended to vary by land type, with the highest proportions found in the plantation site. The plantation site also tended to have more unique differentially active taxa than all other sites. Levels of richness in the differentially active taxa were also higher in impacted sites (burned and plantation) relative to the forest. A number of studies have suggested that tropical rainforest soils tend to be an oligotrophic environment favoring life history strategies of slow growth, stress tolerance, and high substrate affinity, and that agricultural soils tend to favor strategies related to faster growth and low substrate affinity (Rodrigues et al., 2013; Navarrete et al., 2015a; Mueller et al., 2016). Thus, if agricultural soils are overall a less stressful environment that is more

amenable to fast growing taxa, then our finding that these sites harbor a higher proportion and richness of differentially active taxa may not be surprising. We see support for this trend in the increase in relative abundance of members of the *Actinobacteria* (a phylum reported to contain copiotrophic taxa, Naverette et al. 2015b) in the differentially active community in the burned and plantation sites relative to the forest. Our data also suggest that the spatial turnover of the differentially active taxa may follow a similar trend under land use change as seen for the community as a whole. Thus while the differentially active taxa may increase in proportion and richness in converted sites, these taxa may also be susceptible to spatial homogenization.

Tropical ecosystems are home to tremendously high levels of biological diversity and are intimately involved in global biogeochemical cycling. It is therefore imperative that our understanding of the impact of environmental change in these regions continues to expand. Our study illustrates that biotic homogenization following land use change can follow similar trends in the African tropics as it does in the neotropics, even if the land management practices and crops differ. While this finding is of concern, the fact that microbial responses may follow predictable trends presents an opportunity for those interested in mitigating the impacts of environmental change on the soil community. Our results also show that the active fraction of the community follows similar - but more pronounced - trends as the DNA-inferred community, indicating that predicting alterations to this fraction of the community may not require an alternate conceptual framework. Finally, by distinguishing the taxa that are likely to be especially active, we have revealed that

these taxa are proportionately more abundant and species rich in impacted sites than in the forest, indicating a potential shift in microbial life history strategies. Our work contributes towards a broader understanding of the numerous ways in which land use change can impact microbial communities and may help guide future land management strategies towards minimizing impacts.

BRIDGE

In Chapter III I showed that Central African rainforest soil communities are susceptible to biotic homogenization and that these communities follow trends that are largely similar to other tropical soil communities undergoing conversion to agriculture. The fact that microbial communities in the soil contain such high proportions of dormant taxa can make the task of linking variation in microbial community attributes to variation in the functions they mediate exceptionally challenging. I distinguished the active fraction of the community from the total pool of community members –which includes dormant and some dead taxa. My results revealed that the active fraction exhibits a more pronounced response to environmental change, which could be an indication of a disruption to community functions. In the next chapter, I explore the concept of community structure-function relationships in a region of the Amazon undergoing conversion to cattle pasture. Soils in this region have been shown to shift from a methane sink (exhibiting net methane consumption) to a methane source (exhibiting net methane emission) following conversion to cattle pasture. Methane flux is primarily governed by two microbial processes: methanogenesis (methane production) and

methanotrophy (methane consumption). It has remained unclear whether the shift in methane flux observed in the Amazon Basin is driven by physico-chemical alterations to the soil (*e.g.* changes in compaction or O_2 levels) or changes to the community of microorganisms that govern these processes. Since microbial taxa are capable of horizontally exchanging genes, the task of linking variation in microbial community processes to variation in individual organisms can be difficult. For this reason, some have proposed that measuring the gene content of a community (*i.e.* the presence and abundance of genes encoding the cellular machinery that perform a given function) may be a more direct link to community processes. I measured the microbial gene content in Amazonian forest soils and cattle pasture soils with the specific goal of investigating changes to the collection of microorganisms that mediate methane flux. Taking this approach provides me with a unique view of a microbial community undergoing dramatic changes from agricultural conversion and it allows me to develop a conceptual model for how these community changes could translate into changes in ecosystem function.

CHAPTER IV

Conversion of Amazon rainforest to agriculture alters community traits of methane-cycling organisms

This chapter is co-authored by myself, Ann Klein, Babur Mirza, Susanne Tringe, James Tiedje, Jorge Rodrigues, Klaus Nüsslein, and Brendan Bohannan. I primarily performed the data analysis, and writing along with help from Ann Klein. Brendan Bohannan filled the advisory role on this project and as such aided in the conception of the design and the sample collection in the field, as well as contributed edits to the text. Supplementary material for this chapter can be found in Appendix C.

INTRODUCTION

Land use change poses one of the largest threats to global biodiversity (Foley et al., 2005). In tropical regions, this process is occurring disproportionately faster than in any other region worldwide (Dirzo & Raven, 2003; Laurance et al., 2014). In the Amazon Basin, the primary motivation for land use change is conversion to cattle pasture. This process has been shown to alter soil chemistry (de Moraes et al., 1996; Herpin et al., 2002; Rodrigues et al., 2013), as well as soil microbial biodiversity and functional traits (Jesus et al., 2009; Rodrigues et al., 2013; Mueller et al., 2014; Paula et al., 2014), which may be responsible for the alteration of a number of ecosystem processes governed by soil microbes such as methane emission (Verchot et al., 2000; Fernandes et al., 2002).

Methane is a potent greenhouse gas with a global warming potential that is

34 times higher than CO_2 (over a 100-year time frame) (Myhre et al., 2013). It is well established that forest soils throughout the Amazon Basin generally act as methane sinks, but when forests in the Amazon Basin are converted to cattle pasture, the underlying soils can shift from methane sink to source (Steudler et al., 1996; Verchot et al., 2000; Fernandes et al., 2002; Carmo et al., 2012). It is not known what factors are responsible for this shift. Most of the proposed explanations have focused on physico-chemical alterations to the soil (*e.g.* increased water-filled pore space and decreased O_2 diffusion) driving increased methanogenesis (Steudler et al., 1996; Fernandes et al., 2002), yet few have investigated how the conversion process alters the communities and traits of microorganisms responsible for these processes.

Two counteracting microbial processes control biogenic methane emission: methanogenesis and methanotrophy. These processes are both governed by a suite of phylogenetically conserved community traits (Martiny et al., 2013). Methane flux rates have been associated with the community composition (Seghers et al., 2003; Maxfield et al., 2008; Bodelier et al., 2013; McCalley et al., 2014), abundance and activity (Freitag & Prosser, 2009; Freitag et al., 2010) of both methanogens and methanotrophs. Thus, there is precedent to suggest that shifts in methane-cycling community traits could alter rates of methane flux. Moreover, each of these functional groups can be further divided by differences in specific traits. Within the methanotrophs, methanotrophic *Gammaproteobacteria* (also referred to as Type I methanotrophs), *Alphaproteobacteria* (also referred to as Type II methanotrophs), and *Verrucomicrobia* (also referred to as Type III methanotrophs) differ in their

physiology, substrate affinity, and life history strategies (Hanson & Hanson, 1996; Ho et al., 2013; Knief, 2015). Within the methanogens, acetoclastic, hydrogenotrophic, and methylotrophic taxa display different life history strategies, utilize different substrates for methanogenesis, and can generate methane at different rates (Conrad, 1999; Hedderich & Whitman, 2013).

Applying trait-based approaches to microbial ecology can provide an alternative perspective on community responses to environmental change (Green et al., 2008). Such approaches can reveal, for example, shifts in functional potential (*i.e.* gene content), taxonomy (*e.g.* diversity or composition), or life history strategies (Barberán et al., 2012; Krause et al., 2014). Life history strategies generally refer to an organism's investments in survival, growth, and reproduction. Documenting changes in life history strategies has played an important role in understanding how plant and animal communities respond to environmental changes, but this approach has rarely been applied to microbial communities (Fierer et al., 2007), mainly due to the difficulty of cultivating the majority of microbial taxa. Ho *et al*. (2013) classified methanotrophic microorganisms into the Competitor-Stress tolerator-Ruderal life history framework (Grime, 1977) using physiological measurements of cultured representatives and habitat range data. This framework divides taxa among three primary strategies: "Competitors" (exhibiting fast growth under high nutrient or substrate conditions), "Stress Tolerators" (tolerating low or variable substrate availability), and "Ruderals" (performing optimally in frequently disturbed sites). Under this system, *Alphaproteobacteria* methanotrophs are primarily classified as Stress Tolerators –performing better under conditions of low or variable methane

or O_2 availability, while *Gammaproteobacteria* methanotrophs are more variable, spanning from Competitor to Ruderal- implying that they perform better under conditions of high substrate availability or in frequently disturbed sites. Using a trait-based framework such as this provides a new way to assess microbial responses to environmental change, and will help contribute to our understanding of the relationship between community attributes and ecosystem function.

Here, we apply a trait and life history-based framework to ask how methane-cycling communities differ between primary rainforest and cattle pasture derived from primary rainforest in the Western Amazon. We use environmental metagenomics to provide a more comprehensive assessment of microbial community traits than can be obtained by culture-dependent methods or culture-independent methods which rely on sequencing of individual target genes. First, we investigate changes in the abundances of methane-cycling taxa, their functional traits, and their life history strategies. Second, we compare the abundance of genes involved in methane-cycling pathways. Finally, we discuss how changes to these community traits are consistent with the shift from methane-sink to source previously reported from this site.

MATERIALS AND METHODS

Site Description and Sampling

Our study was performed at the Amazon Rainforest Microbial Observatory (ARMO) site (10°10'5'' S and 62°49'27'' W). This site was selected as representative of the current agricultural expansion in the Western Amazon. It is located in the

Brazilian state of Rondônia, which has experienced the highest percentage of forest loss of any state in the Brazilian Amazon (Rodrigues et al., 2013). Agricultural conversion in this region typically follows the following stages: 1) selective logging of valuable timber, 2) slash-and-burn deforestation of the remaining vegetation, and 3) aerial seeding of members of the non-native fast-growing grass genera *Urochloa* (formerly *Brachiara*) or *Panicum* in order to establish pasture for cattle ranching. Pastures may be burned periodically in order to control the invasion of weeds. Herbicides, tillage, or chemical fertilizers are not commonly used. The vegetation and soil characteristics at this site have been described in detail elsewhere (Neill et al., 1997a; Feigl et al., 2006).

Ten soil cores were collected from ARMO in April 2010 (5 soil cores from primary rainforest and 5 from a 38 year-old converted pasture). Soil was sampled to a depth of 10 cm (after removal of the litter layer) using standard coring methods and homogenized. Samples were frozen on the spot, transported on dry ice, and stored at -80° C until extraction.

Soil DNA Extraction and Sequencing

DNA was extracted from five soil subsamples per core (*i.e.* 50 extractions per 10 soil cores) following the same protocol described in Mirza *et al*. (2014). DNA from the subsamples was pooled, and 3-5 μg of DNA were used from each sample. Metagenomic libraries were constructed from 10 samples using the Illumina TruSeq kit with ~270 bp insert sizes. Sequencing of 150 bp paired-end reads was performed on the Illumina HiSeq platform. In total, 21 lanes were sequenced to

produce 6.4 billion paired-end reads, resulting in an average of 636 million (±12%) reads per sample.

Bioinformatics and Statistics

Functional and taxonomic annotations were obtained using the MG-RAST pipeline (Meyer et al., 2008). Raw sequences were uploaded to MG-RAST, and paired-end reads were joined using fastq-join as part of the MG-RAST pipeline. Single end reads that could not be joined were retained. After merging paired-end reads, a total of 6.3 billion sequences with an average length of 171 bp were processed through the MG-RAST pipeline. All other pipeline options were left as default (*i.e.* trimming of low quality bases, removal of artificial replicate sequences, and filtering of sequences with greater than 5 ambiguous bases). Hierarchical functional annotations were generated using the SEED subsystems (Overbeek et al., 2014) and organismal annotations were obtained via the MG-RAST M5RNA database (Wilke et al., 2012) which assigns taxonomy strictly from ribosome-encoding genes including those from the SILVA, RDP, and Greengenes databases (Wilke et al., 2015). We used the "Representative Hit" classification method for organismal annotation, which selects a single, unambiguous annotation for each feature and assigns taxonomy. Default parameters (e-value cutoff = 1e-5, Min. % identity = 60%, Min. alignment length cutoff = 15) were used for both the functional and taxonomic annotations. Low quality sequences (16.4%) were removed prior to assignment of remaining sequences: rRNA genes (0.5%), predicted proteins with known functions (21.7%), predicted proteins with unknown functions (47.9%), and

sequences without any rRNA genes or predicted proteins (13.5%). We obtained usable annotations for approximately 22.2% of the total sequences (0.5% rRNA genes + 21.7% predicted proteins with known functions).

The functional annotations had an average of 1.2 million ± 136 000 observations per sample. The organismal annotations had an average of 304 000 ± 76 000 observations per sample. Annotation tables were subsampled to achieve equal sampling depth across samples. The functional annotation table was rarefied to 1.05 million observations per sample and the organismal table was rarefied to 195 000 observations per sample. All analyses were performed using the R statistical environment (R Development Core Team, 2010) including the vegan package (Oksanen et al., 2015).

We constructed functional group community matrices by selecting only species previously reported in the literature as methanotrophs or methanogens (Supplementary Tables 1 & 2, respectively). Community composition differences were statistically tested using a permutational multivariate ANOVA (PERMANOVA) on Bray Curtis community distances (Bray & Curtis, 1957) and Euclidean distances (ter Braak, 1995) in the case of acetoclastic methanogens because of its treatment of absences. Shannon diversity, species richness, Simpson diversity, and Pielou's evenness were used to assess varying aspects of functional group diversity and were compared across sites using a two-sample two-tailed *t*-test. The abundance of genes encoding methane-cycling functions were compared across samples using a two-sample two-tailed *t*-test. A Shapiro-Wilk Normality test was performed before all *t*-test comparisons to verify normal distribution of the data and log transformation of

counts was used when necessary to achieve normality. Comparison of proportions or ratios across land types was performed by first logit transforming the proportion or ratio (Warton & Hui, 2011), and then testing using a two-sample one-tailed *t*-test because of *a priori* hypotheses regarding the direction of change. For all box plots, the box represents the interquartile range (Q1-Q3), the line represents the median (Q2), the whiskers represent the minimum and maximum non-outlier values below Q1 and above Q3, respectively, and any points outside represent outliers.

Methanotroph genera were characterized under the Competitor-Stress tolerator-Ruderal functional classification (Grime, 1977) proposed for methanotrophs by Ho *et al*. (2013) (Supplementary Table 3). While we are far from having a comprehensive understanding of the life history strategies of all methanotroph taxa (Knief, 2015), traits relevant to life history strategy such as pH and temperature ranges/optima have been shown to be phylogenetically conserved (Krause et al., 2014) in methanotroph lineages, allowing us to putatively assign the taxa detected in our study to life history strategies based on their taxonomic relatedness to the taxa classified by Ho et al (2013). Once taxa were placed into life history groups, the relative abundances of the groups were compared using a two-sample two-tailed *t*-test. Ratios of these groups were first logit-transformed then compared using a two-sample two-tailed *t*-test.

RESULTS & DISCUSSION

Shifts in the proportion of methanotrophs in the methane-cycling community

Methane flux is the net balance between methane production and methane

consumption. Thus changes to the balance of the organisms that mediate these processes, or genes encoding the cellular machinery by which these processes take place, could alter methane flux. The forest site where our study was conducted has been shown to exhibit net negative methane flux (methane consumption), even in the wet season (Fernandes et al., 2002). In contrast, the cattle pasture where we sampled was shown to exhibit positive methane flux (methane emission), even in the dry season (Fernandes et al., 2002). One of our primary findings is that the proportion of methanotroph sequences relative to methanogen sequences is significantly lower in the pasture compared to the forest (Fig. 1a). This was driven by a significant decrease in methanotroph abundance in the pasture and no significant change in the methanogen abundance across land types. Furthermore, we estimated the potential for methanotrophy and methanogenesis from the abundance of genes encoding methane monooxygenases (MMO) and methyl coenzyme M reductase (MCR), respectively. The proportion of MMO genes in the genes coding for these processes shows a trend similar to our taxonomic results (Fig. 1b), driven by a decrease in MMO gene abundance and an increase in MCR gene abundance. A shift in these proportions could impact the amount of methane consumed relative to the amount of methane produced, leading to changes in the net flux of methane from the soil.

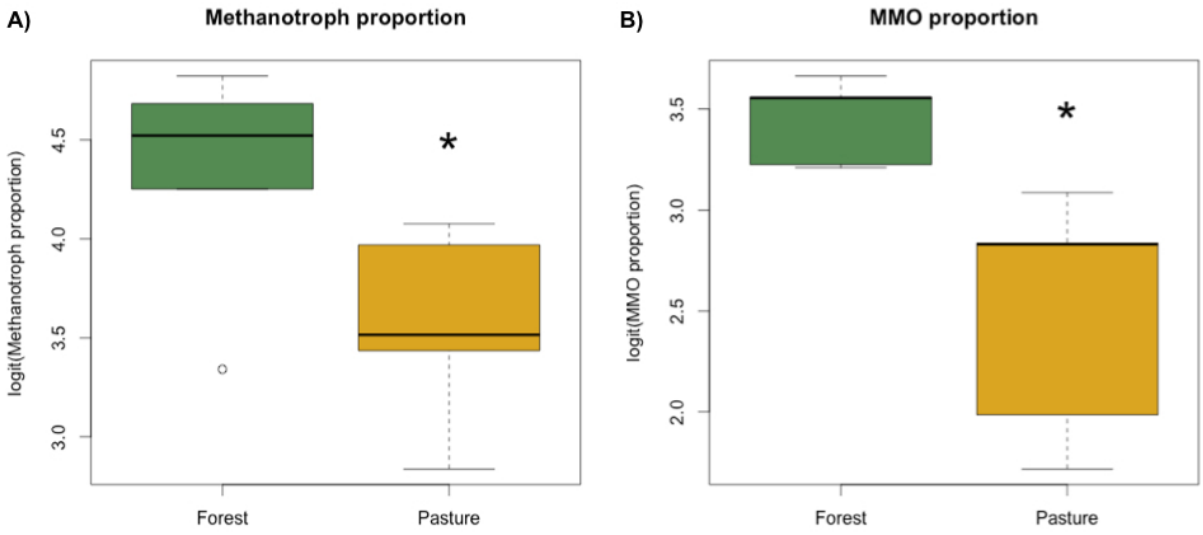


Figure 1: The proportion of methanotrophs differs between forest and pasture. A. The proportion of methanotrophs in the methane-cycling community: methanotrophs/(methanotrophs + methanogens) calculated from rRNA organismal annotations. B. The proportion of methane monooxygenase (MMO) genes within the genes unique to methane-cycling: MMO/(MMO + MCR). MMO is methane mono-oxygenase; a methanotrophy gene marker. MCR is methyl-coenzyme M reductase; a methanogenesis gene marker. Significant differences between forest and pasture are denoted as: * $P < 0.05$.

Shifts in taxonomic groups

Our annotations recovered sequences from *Gammaproteobacteria*, *Alphaproteobacteria*, and *Verrucomicrobia* methanotrophs (Supplementary Table 1), as well as methanogen sequences from the orders *Methanobacteriales, Methanococcales, Methanocellales, Methanomicrobiales, Methanosarcinales*, and

Methanopyrales (Supplementary Table 2).

Methanotrophic taxa. We observed numerous changes to methanotroph community traits that suggest that methanotrophy is altered following conversion of forest to pasture. Methanotroph communities were compositionally different across the two land types (Bray-Curtis R^2=0.41, $P < 0.05$). The forest community exhibited a significantly higher average pairwise dissimilarity than the pasture community (Bray-Curtis, $P < 0.05$), indicating a higher variability in community membership in the forest. The methanotroph community did not differ in richness, Shannon diversity, Simpson diversity, or evenness across land types. However, when the *Proteobacteria* methanotrophs were considered separately, Shannon diversity, Simpson diversity, and evenness were significantly higher in the pasture ($P < 0.05$).

Of the three primary types of methanotrophs at our sites (*Alphaproteobacteria*, *Gammaproteobacteria*, *Verrucomicrobia*), the *Alphaproteobacteria* methanotrophs varied more between the land types (Fig. 2). Neither *Gamma*- nor *Verrucomicrobia* methanotroph communities differed in composition, diversity, abundance or evenness across the land types. The *Alphaproteobacteria* methanotrophs, however, showed a significantly lower relative abundance in pasture, as well as significant compositional differences across forest and pasture soils (Bray-Curtis R^2=0.65, $P < 0.01$). The *Alphaproteobacteria* methanotrophs also exhibited significantly higher Shannon diversity and evenness in pasture, relative to forest. The two most abundant *Alphaproteobacteria* genera (*Methylocella* and *Methylosinus*) were the only genera to significantly change in

abundance between land types, with both showing a decrease in the pasture. The changes observed in *Alphaproteobacteria* methanotrophs could be of concern because numerous *Alphaproteobacteria* taxa have been shown to be capable of consuming atmospheric levels of methane (Holmes et al., 1999; Kolb, 2009), suggesting that these organisms can take up methane even in environments where methane is not produced directly. Due to the short read lengths of our data, we lacked the taxonomic/phylogenetic resolution to identify these high-affinity methanotrophs (i.e. members of upland soil clusters alpha and gamma (reviewed in Kolb 2009)) at the species or strain level. Other studies, including one in the tropics (Knief et al., 2005), have reported lower abundance of *Alphaproteobacteria* methanotrophs in agricultural or grassland soils relative to forest soils and have reported decreases in *Alphaproteobacteria* methanotrophs following deforestation (Singh et al., 2007; Nazaries et al., 2013). For example, in Thailand Knief et al. (2005) observed a lower abundance of *Alphaproteobacteria* methanotrophs (along with a decrease in methane consumption) in agricultural soils relative to tropical forest soils. Thus, the changes we report are in accordance with other reports from agricultural conversion sites and may influence the methane uptake rates of the soils in which they are found.

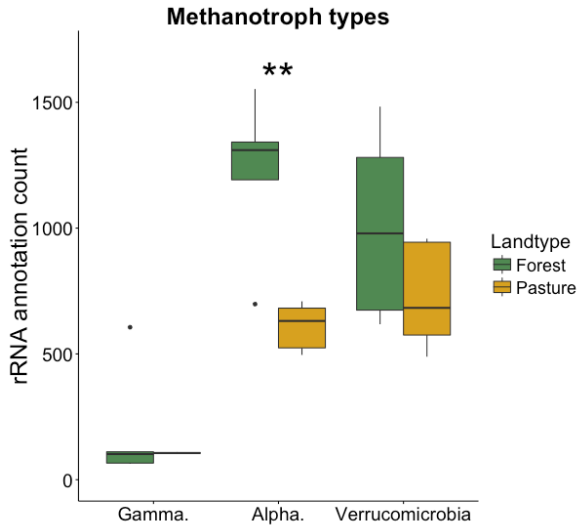


Figure 2: **Differential response of methanotroph types to land use change**. The relative abundance of *Gammaproteobacteria*, *Alphaproteobacteria*, and *Verrucomicrobia* methanotroph rRNA annotations in rainforest and pasture soils. Significant differences between forest and pasture denoted as: ** $P < 0.01$.

Life history strategies of methanotrophs

We grouped methanotroph genera identified in our study according to the Competitor-Stress Tolerator-Ruderal framework proposed by Ho *et al*. (2013, see Supplementary Table 3)(Ho et al., 2013). We observed that the ratio of Stress Tolerators (S) to Ruderals (R, aka Disturbance Specialists) was significantly lower in the pasture ($P < 0.001$, Fig. 3). This change was driven by a significant decline in the Stress Tolerator group in the pasture ($P < 0.01$) and a marginally significant increase in the Ruderal group ($P = 0.05$) in the pasture. Other strategies characterized as Stress Tolerator (C-S and S-R) were also significantly lower in the pasture ($P < 0.01$) compared to the forest. No other groups were significantly

different across the two land types. This could be indicative of a larger trend suggested by other research (Rodrigues et al., 2013; Navarrete et al., 2015a, 2015b) whereby tropical rainforest soils harbor microbial communities that are more stress-tolerant or K-selected (*i.e.* oligotrophic) than those of agricultural soils.

By integrating our data into a life history framework we are able to gain a new perspective on the environment as experienced by the microbial community. Much of the differentiation in life history strategies is across the stress gradient. We speculate that because of the stronger fluctuations in temperature, moisture, and oxygen in the pasture soils – likely due to increased exposure and decreased light attenuation- ruderal (disturbance specialist) life history strategies will be favored. By contrast, the forest soil environment, although more heterogeneous over space, may be more stable over time (but overall less rich in nutrients or substrate) and hence favor stress tolerator life history strategies which grow more slowly but tolerate low nutrient conditions. This low substrate environment may also favor strategies exhibiting facultative substrate usage. Several *Alphaproteobacteria* taxa have been shown to be facultative methanotrophs (as reviewed in Ho *et al*. (2013))- one of which (*Methylocella*) significantly decreases in the pasture ($P < 0.01$). This may help to explain why the stress tolerator group is significantly higher in forest. Disturbance specialists would likely persist under adverse conditions such as those of pasture soils but grow rapidly under periodic shifts to optimal conditions. We see other evidence for this strategy in increases in genes related to dormancy and sporulation ($P < 0.01$, Supplementary Fig. 1a) and spore DNA protection ($P < 0.001$, Supplementary Fig. 1b) in pasture soils. These changes in life history traits are likely

to influence broader patterns of ecosystem functioning (*e.g.* C cycling) and may serve as a means to assess ecosystem changes, much like a bioindicator, to inform management practices.

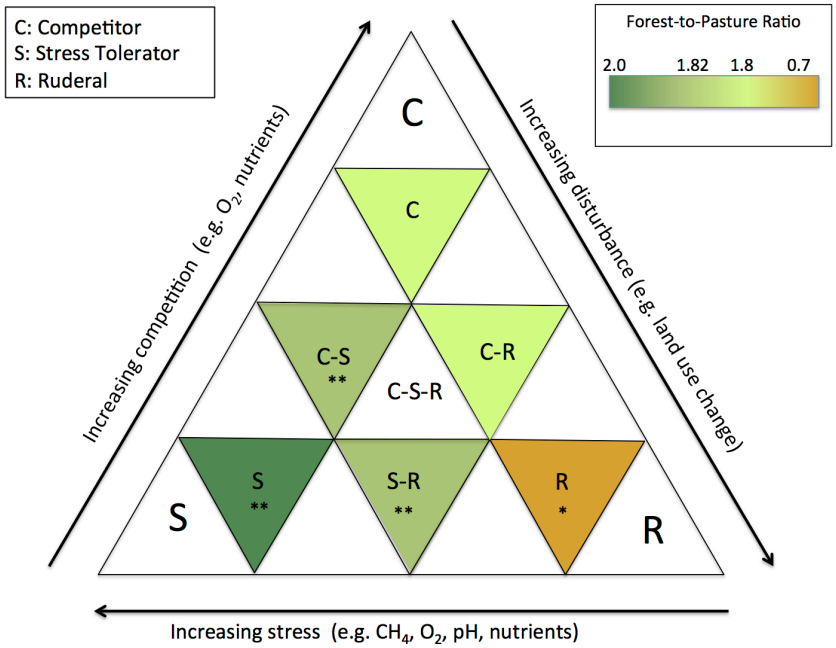


Figure 3: Methanotroph life history strategies vary across rainforest and pasture soils. Life history strategies of methanotroph genera were categorized along three axes (Competitor, Stress tolerator, Ruderal, or combinations thereof) according to the recommendations of Ho *et al.* (2013). Triangles are color-coded by the forest-to-pasture abundance ratio of that strategy group. Significant differences between forest and pasture are denoted as follows: ** $P < 0.01$, * $P = 0.05$.

Methanogenic taxa

The relative abundance of methanogenic microorganisms did not significantly differ across land types. There were, however, compositional differences between forest and pasture soils. The communities were significantly different in composition

across land types (Bray-Curtis R^2=0.61, P<0.001). The average pairwise dissimilarity was higher in the forest soils than the pasture soils (similar to that observed for the methanotroph community), indicating more variation in community membership in the forest than the pasture. We observed no significant differences in diversity or evenness across sites. The proportion of acetoclastic methanogens in the methanogen community was significantly higher in the pasture than the forest (P < 0.01, Supplementary Fig. 2a), driven by a slight (but not significant) increase of acetoclasts and a slight (but not significant) decrease of hydrogenotrophs. When considered separately, the diversity patterns of these two functional groups also varied across forest and pasture. Both groups differed significantly in composition across land types (acetoclast: Euclidean R^2=0.30, P < 0.01, hydrogenotroph: Bray R^2=0.64, P < 0.05). However, changes in diversity across land types were not consistent between the acetoclast and hydrogenotroph communities. The acetoclasts had a significantly higher species richness in the pasture (P < 0.05), while the hydrogenotroph community had a significantly lower species richness (P < 0.05), Simpson diversity (P < 0.01), Shannon diversity (P < 0.05), and evenness (P < 0.05) in the pasture. Six of the eight acetoclast taxa were unique to the pasture. Within the five orders of hydrogenotrophs there was also a differential response to land use change. The second most abundant hydrogenotrophic order in the forest (*Methanopyrales*) was completely absent in pasture, while the most abundant hydrogenotrophic order in the pasture (*Methanocellales*) was more than an order of magnitude less abundant in the forest. This differential response could be the result of variable life history strategies

within this functional group, an interesting avenue for future research.

Changes in gene content

Methanotrophy genes

To investigate the potential for methane oxidation in forest and pasture soils, we analyzed differences in the relative abundance of genes that code for the methane monooxygenase enzyme as well as genetic markers for *Gammaproteobacteria* and *Alphaproteobacteria* methanotrophs. The methane monooxygenase enzyme is unique to methanotrophs and is the only currently known way that aerobic methanotrophs can utilize methane-derived carbon (Anthony, 1991; Hanson & Hanson, 1996). This enzyme has two forms: soluble methane monooxygenase (sMMO) and particulate methane monooxygenase (pMMO)(Hanson & Hanson, 1996). The particulate form has a higher substrate-specificity and cells containing pMMO tend to have a higher methane affinity and thus may play a role in the consumption of atmospheric methane (Hanson & Hanson, 1996). Similar to another study at this site (Paula *et al*. 2014), we observed that the relative abundance of genes encoding enzymes in the pMMO pathway was significantly lower in pasture soils ($P < 0.01$, Fig. 4a). In contrast, genes for sMMO did not differ in abundance between forest and pasture (Fig. 4b). We, however, were not able to distinguish between pMMO and the higher-affinity isoenzyme pMMO2 (Tchawa Yimga et al., 2003) due to limitations of the database used for annotations and our short read lengths.

There are two biochemical pathways for assimilation of carbon from

methane by methanotrophs: the ribulose-monophosphate pathway (used by Type I methanotrophs) and the serine pathway (used by *Alphaproteobacteria* methanotrophs)(Hanson & Hanson, 1996). Despite observing changes to the relative abundance and composition of *Alphaproteobacteria* methanotrophs, we did not detect changes in the abundance of genes involved in the ribulose-monophosphate or serine pathways. We attribute these results to the broader involvement of these pathways in one-carbon metabolism (methylotrophy), resulting in these pathways being shared by more groups than just methanotrophs (Anthony, 1991).

Methanogenesis genes

The overall abundance of all genes related to methanogenesis did not significantly differ between land types (Fig. 4c). However, the relative abundance of genes encoding the common marker enzyme, methyl coenzyme M reductase, was significantly higher in the pasture than the forest ($P < 0.01$, Fig. 4d). Genes involved in methanogenesis from methylated compounds (an alternative pathway by which methane can be produced) were also significantly more abundant in the pasture ($P < 0.05$, Supplementary Fig. 2b). This pathway has not been the focus of many landscape-level studies, and it is unclear to what extent these changes might influence soil methane emissions. Finally, we observed that genes encoding two enzymes involved in the production of acetate (formyltetrahydrofolate synthetase and the beta subunit of the acetyl-CoA synthase) were significantly more abundant in the pasture ($P=0.05$, $P<0.05$, respectively). This indicates that the potential to synthesize acetate (the substrate for acetoclastic methanogenesis) is higher in

pasture and could be related to the increase in the proportion acetoclastic methanogens we observed in the pasture soils.

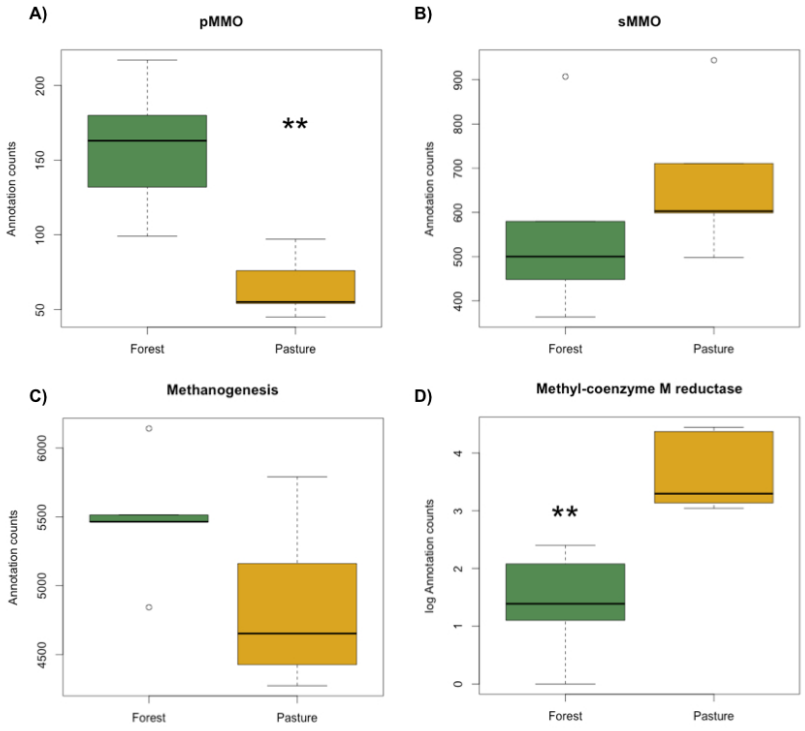


Figure 4: Functional genes related to methane cycling respond differently to land use change: A) The relative abundance of genes encoding particulate methane monooxygenase (pMMO). B) The relative abundance of genes encoding soluble methane monooxygenase (sMMO). C) The relative abundance of genes involved in the methanogenesis pathway. D) The relative abundance of genes encoding the methyl-coenzyme M reductase enzyme. Significant differences between forest and pasture are denoted as follows: ** $P < 0.01$, * $P < 0.05$.

Previous research has shown that forest soils are methane sinks while agricultural (including pasture) soils are methane sources (Maxfield et al., 2008; Nazaries et al., 2013) . This is true of tropical soils, including soils in the Amazon

basin (Steudler et al., 1996; Verchot et al., 2000; Carmo et al., 2012), and specifically those at the site at which our study was conducted (Fernandes et al., 2002). The causes of this difference remain unknown. Researchers have proposed that increased water filled pore space and decreased O_2 diffusion in pasture and agricultural soils are responsible for this trend through the reduction in oxygen and a subsequent increase in methanogenesis, which is a strictly anaerobic process (Verchot et al., 2000). However, other studies have correlated increases in methane fluxes with changes in the methanotroph community (*e.g.* community structure, diversity, and activity (Seghers et al., 2003; Maxfield et al., 2008; Bodelier et al., 2012, 2013)). Evidence from our study suggests that changes to the methane-cycling community could be playing a role in the observed shift from methane sink to source (Fernandes et al., 2002) at our site. Although we observed some differences in the methanogen communities of forest and pasture (most notably a change in the relative abundance of acetoclasts), there were few large-scale shifts to the methanogen community (*e.g.* the abundance of methanogens did not change across sites, nor did the abundance of total genes involved in methanogenesis pathways). In general we observed much larger and more varied differences between forest and pasture for the methanotroph community. Among the most striking differences we observed were a decrease in the abundance of methanotroph taxa, a decline in the proportion of methanotrophs in methane-cycling community, a decrease in pMMO abundance, a decrease in the relative abundance of *Alphaproteobacteria* methanotrophs, and a decrease in the methanotroph Stress Tolerator-to-Ruderal ratio. Hence the shift to methane

emission at our site may be due, at least in part, to altered methane consumption rates caused by changes to methanotroph community traits.

Our work illustrates the importance of using environmental metagenomics to address questions regarding microbial functional ecology. Rather than perform shallow sequencing on samples from a number of different pastures and forests, we chose to deeply sequence replicate samples from a limited number of sites, in order to detect low-abundance functional genes. Although this approach restricted our scale of inference to the sites we sampled, rather than forests and pastures across the tropics, the results we report could not have been obtained without deep metagenomic sequencing. The depth at which we performed sequencing allowed us to investigate changes to rare genetic traits that are difficult to assess without PCR amplification (*i.e.* those involved in methanotrophy or acetate production). Our gene annotations yielded a wide breadth of genes; 13 418 different genes, 245 of which have the potential to be directly involved in the cycling of methane. We were able to detect methanotrophic taxa (*e.g. Methylocella*, which is not known to have pMMO (Dedysh et al., 2000)) that would not have been detected using PCR-based, culture independent methods (i.e. amplification of *pmoA* gene regions). The use of metagenomics also allowed us to simultaneously survey taxa and genes across multiple functional groups. Finally, our work is an example of the power of combining trait inference from metagenomics with life history theory to generate novel hypotheses regarding the functional responses of microorganisms to environmental change (Barberán et al., 2012).

Although a variety of studies have investigated the impact of land use change

on methane-cycling organisms, to our knowledge this is the first to do so in the Amazon Basin using metagenomics. The majority of past studies have focused on temperate regions where rates of agricultural conversion have been relatively low over the last 50 years. Tropical regions, in contrast, are currently facing a faster rate of land use change than any other region (Dirzo & Raven, 2003; Foley et al., 2005; Laurance et al., 2014), and the Amazon Basin is facing the highest rate of all tropical regions. Several studies have reported increases in methane flux from soil in the Amazon as a result of agricultural development (Steudler et al., 1996; Verchot et al., 2000; Fernandes et al., 2002; Carmo et al., 2012). Our study shows that there are numerous alterations to methane-cycling community traits, many of which have been linked to variation in methane flux in other studies (Seghers et al., 2003; Maxfield et al., 2008; Bodelier et al., 2012, 2013). Thus we suggest that alterations to soil microbial communities could be one of the driving factors behind the shift from methane sink to source following land use change in the Amazon (Fernandes et al., 2002).

CHAPTER V

CONCLUSION

Synthesis

Studying the spatial distribution of organisms has played a central role in the development of many fundamental principles in ecology and evolution. Incorporating microbial life forms into a spatial framework that is consistent with macro-organisms, however, has been a challenge both conceptually and quantitatively (Martiny et al., 2006; Hanson et al., 2012). Ever since it was shown that microbial communities tend to exhibit much weaker spatial patterns than macro-organismal communities, ecologists have been puzzled by this disparity. Are microorganisms fundamentally different from macro-organisms, or is this disparity somehow a product of how we conceptualize or measure microbial communities? This notion not only brings into question the universality of some of the most foundational principles in ecology, but it also has the potential to impact the strategies we use to conserve communities and the functions they perform. Indeed the idea that microbial communities change over space means that alterations to the environment could disrupt their spatial ecology and potentially drive the extinction of microbial lineages. Thus, our basic understanding of microbial spatial and functional ecology must be expanded in order to minimize our impacts on these communities.

Incorporating spatial and functional ecology can help us better understand microbial communities and their responses to environmental change. Predicting community responses to change requires a fundamental understanding of the

factors structuring microbial communities. In Chapter II, I demonstrated that bacteria and trees showed very different spatial patterns, bringing into question whether microbial and macro-organismal communities are shaped by different spatial processes. I showed that this difference in spatial patterning could not be explained by differences in species definition, spatial scale, or community activity. Rather, it appears that these differences are likely an effect of undersampling. Both microbial and tree community members followed a frequency-abundance relationship whereby abundant members were more widespread and low abundance members were more restricted in range. Hence by under-sampling a community (a problem which is particularly pronounced in highly complex microbial communities), we tend to miss many of the taxa that are low-abundance and restricted in range, driving the underestimation of spatial diversity. What this suggests is that microorganisms and macro-organisms may not be so fundamentally different in their spatial scaling, and since conservation strategies tend to incorporate perspectives on the spatial distributions of taxa (Diamond, 1975), this finding suggests we may not need an alternate framework to incorporate microbes into our conservation strategies.

Biotic homogenization includes a loss of local diversity, a loss of spatial diversity, and a loss of endemic taxa (Olden & Poff, 2003; Olden et al., 2004; Smart et al., 2006). These trends have been documented in microbial communities in regions undergoing agricultural conversion in the Amazon Basin (Rodrigues et al., 2013). It, however, has remained unclear whether land use change in other regions of the tropics would drive similar responses. In Chapter III, I show that the same

indicators of biotic homogenization can also be detected in Central Africa. Most related work has compared an intact ecosystem (such as a rainforest) to an established agricultural system (such as a pasture or plantation). I was able to gain an additional perspective by incorporating a recently burned site into my design. In nearly all cases the burned site appeared more impacted, showing for example, larger losses in species richness, spatial diversity, and landscape diversity. The plantation site, in contrast, seemed much less impacted, being for the most part indistinguishable from the forest site in diversity. This additional perspective shows that the act of ecosystem conversion (*i.e.* the slash-and-burn process), as opposed to the act of planting and managing land, may be the driving force behind community change in this system. Furthermore, I asked whether the active fraction of the microbial community (as opposed to the total community that includes active and inactive taxa) is responding to land use change in a similar manner. Although it seems logical to distinguish the active taxa in a community if one is interested in ecosystem function, to date there has been little effort to do so in the context of environmental change, and especially not in the tropics. In almost all cases the active fraction followed the same trends as the total community, however this fraction tended to exhibit a more pronounced response (*e.g.* larger losses of species richness and larger changes to the rate of spatial turnover). Other studies have illustrated that the active community is more responsive to seasonal changes (Barnard et al., 2013) and that it is shaped to a larger extent by the environment (Zhang et al., 2014) relative to the total community. My work contributes to this

body of work by suggesting that the active community may also be more responsive to environmental change.

Community responses to environmental change are often accompanied by changes to ecosystem function. Land use change in the Amazon Basin, for example, has been linked to numerous changes to microbially-mediated functions such as nitrogen- and methane-cycling (Fernandes et al., 2002; Neill et al., 2005). Connecting measurements of these processes to changes in community structure is a considerable challenge in communities characterized by high levels of diversity. Many processes also vary in their levels of phylogenetic conservatism (Martiny et al., 2013), making it even more difficult to predict processes from observations of individual taxa. One approach researchers use to circumvent this problem is to survey the gene content of a community (*i.e.* the metagenome) as opposed to surveying the individual taxa (Fierer et al., 2014). In this way the presence and abundance of the genes encoding a given process may be a more suitable proxy for that process than the taxa capable of performing that process. In Chapter IV, I applied this approach to Amazonian soils to investigate shifts in the methane-cycling community following conversion to cattle pasture. Soils at these sites have previously been reported to shift from a methane sink (showing net consumption of methane) to a methane source (showing net methane emission) following conversion to agriculture and it has remained unclear whether this shift is driven by physico-chemical alterations to the soil (*e.g.* changes in soil compaction or O_2 content) or whether this is driven by changes to the community of microorganisms that perform these processes (Fernandes et al., 2002). Methane flux is governed by

two microbial processes: methane production and methane consumption. My work illustrates that there are numerous alterations to the methane-cycling community and that the methane consumption process is more impacted by ecosystem conversion than the methane production process. This includes a decrease in the genes encoding the enzymes responsible for methane consumption, a pronounced change in the composition of methane-consuming bacteria, and a shift in the life history strategies employed by these taxa. This is a strong indication that this functional shift could be driven by changes to the community of microorganisms performing these processes. Moreover, it illustrates that this function, which is a balance between two processes, could be largely changing because of a perturbation to one process (methane consumption). This work highlights the perspective gained from implementing a metagenomic approach in studies of environmental change, and it contributes a new conceptual framework of community structure-function relationships in a changing environment.

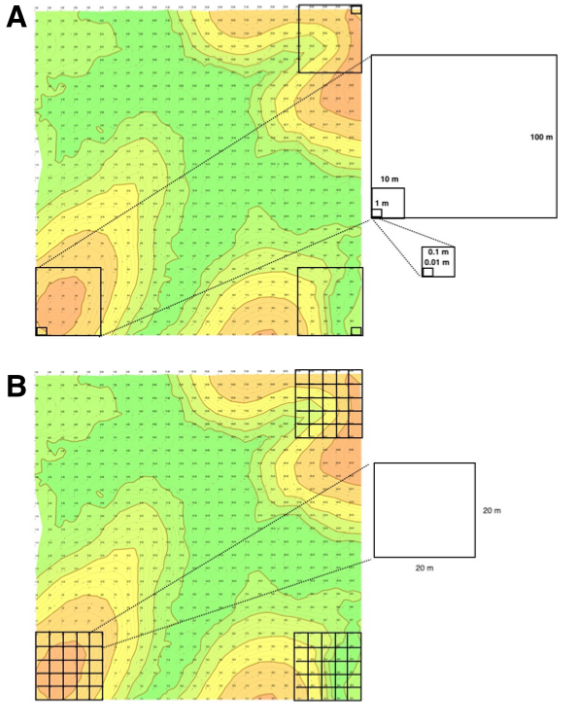
Moving microbial ecology from descriptions to predictions

Microbial communities shape the chemical composition of the planet. In this way, the functions these communities perform can play a decisive role in the fate of our planet. Moving from descriptions of community structure to predictions of community function is an important step forward if we are to better estimate the impacts of our actions on the environment. Understanding the relationship between community attributes and ecosystem function has been a central goal of ecology, and this task will only become more important in a changing environment.

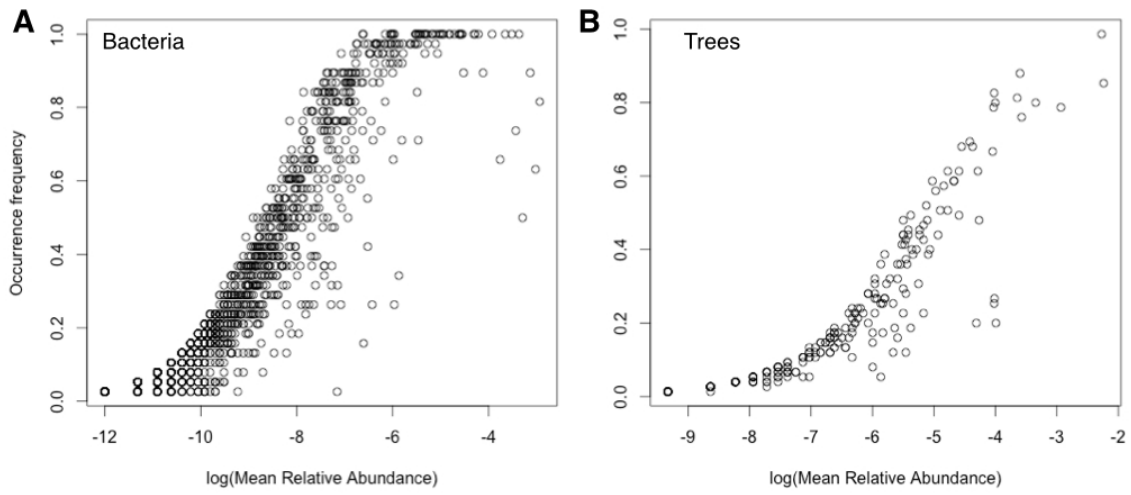
Microbial communities are characteristically highly complex and have many attributes that set them apart from the rest of life (*e.g.* dormancy and horizontal gene transfer). These nuances make understanding the relationship between variation in community attributes and variation in the functions these communities mediate particularly challenging, but this provides an excellent opportunity to expand the generality of ecological principles. In my work I have begun to ask how microbial communities scale with space and how microbial activity can help us better understand the impacts of environmental change. I have also asked how microbial gene content (as opposed to taxonomic composition) can provide an alternative perspective on a system undergoing changes in ecosystem function. My goal in this work has been to explore the many attributes of microbial communities in the hope of mitigating future changes to these communities and better predicting their functions.

APPENDIX A

SUPPLEMENTARY INFORMATION FOR CHAPTER II



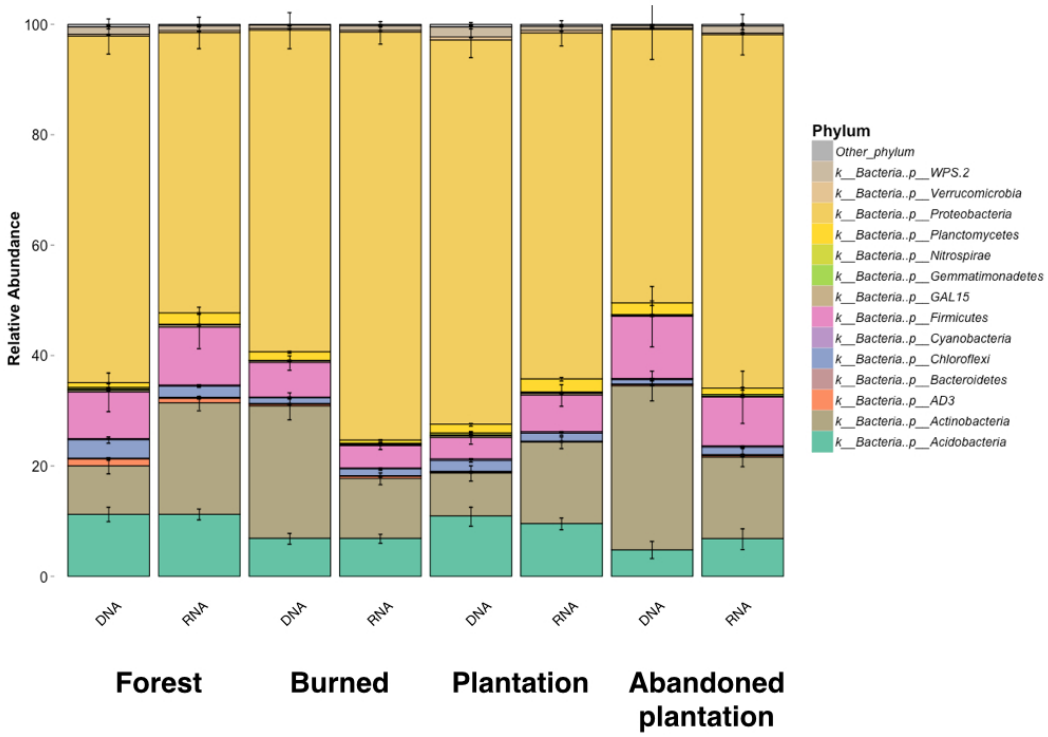
Supplementary Figure 1: Sampling scheme for A) soil microbial communities using a spatially explicit nested design, and B) tree communities taken on 20m x 20m quadrats on the Rabi CTFS plot, Gabon.



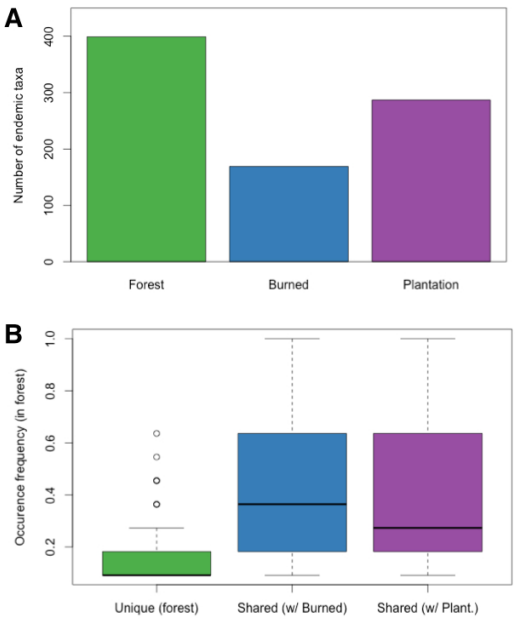
Supplementary Figure 2: The frequency-abundance relationship for A) Bacteria and B) Trees in the Rabi plot, Gabon.

APPENDIX B

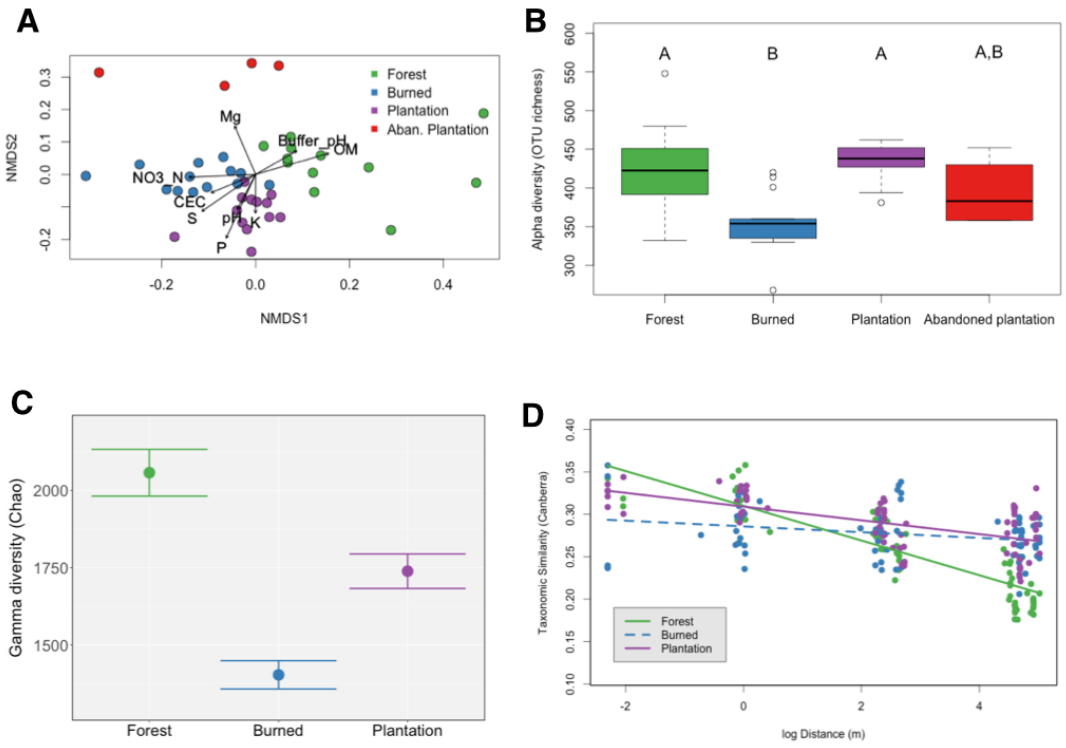
SUPPLEMENTARY INFORMATION FOR CHAPTER III



Supplementary Figure 1: Phylum-level changes to the composition of the DNA- and RNA-inferred communities across the slash-and-burn chronosequence.

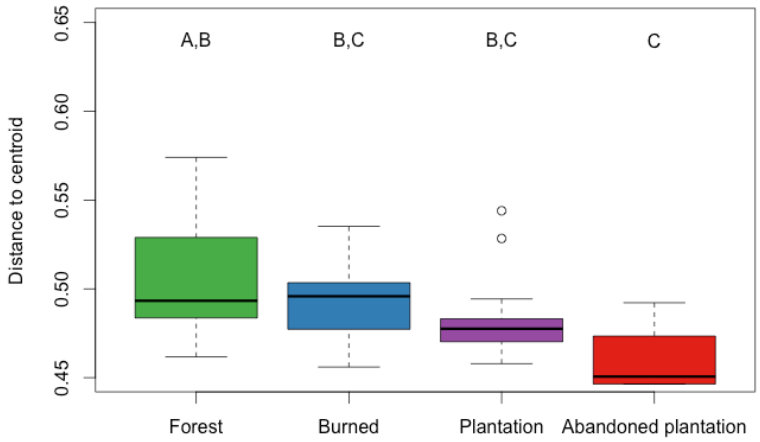


Supplementary Figure 2: The abundance and spatial distribution of endemic taxa is impacted by land use change. A) The total number of endemic taxa (in the DNA-inferred community) across sites. B) Forest endemics tend to be more spatially restricted than the taxa that are shared between the forest and the impacted sites.

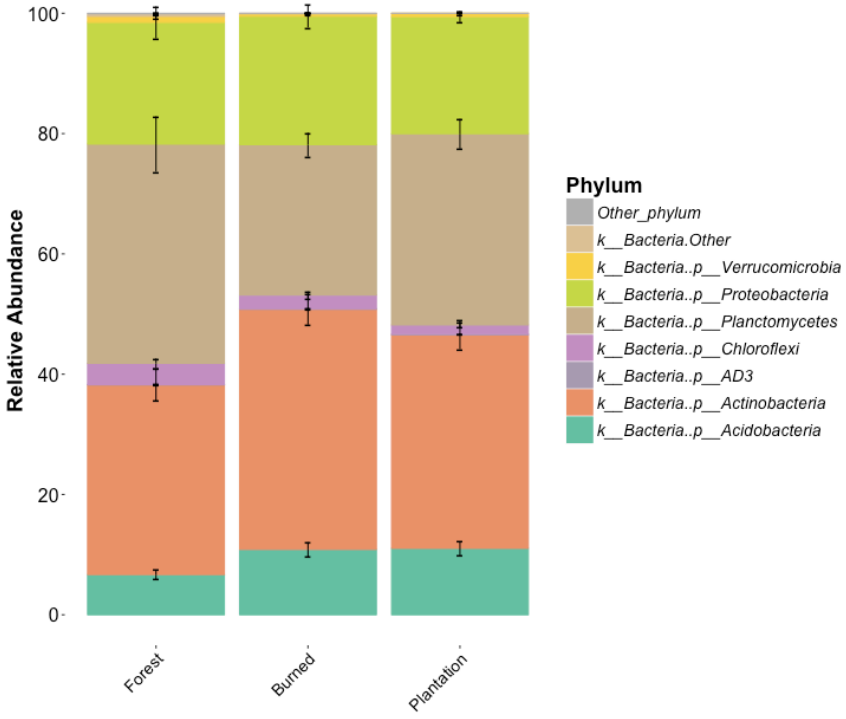


Supplementary Figure 3: Changes to the active (RNA-inferred) community qualitatively follow the same trend as the DNA-inferred community, but are slightly more pronounced. A) Composition of the active community clusters by land use, non-metric multidimensional scaling (NMDS) of Canberra dissimilarities showing significant ($P < 0.05$) environmental vectors. OM: organic matter, K: potassium percent saturation, P: weak-bray phosphorus, pH: soil pH, S: sulphate-sulfur, CEC: cation exchange capacity, NO_3-N: nitrate nitrogen, and Mg: magnesium percent saturation B) within-sample OTU richness of the active community across conversion sites, significant differences assessed using Tukeys HSD on one-way ANOVA. C) Gamma (landscape-level) diversity of the active community from Chao1 estimator showing standard error bars. D) The rate of distance-decay of active

community similarity (1- Canberra dissimilarity) significantly differs ($P < 0.05$) by land use. Dashed line indicates a non-significant ($P > 0.05$) trend.



Supplementary Figure 4: Beta-diversity (average distance to centroid) tends to decrease across the slash-and-burn chronosequence. Based on Canberra distances. Significant differences assessed using Tukey's HSD of one-way ANOVA results.



Supplementary Figure 5: Phylum-level composition of the differentially active taxa differs across land use. Differentially active taxa are defined as those taxa with a significant ($P_{adj} < 0.05$) enrichment in the RNA relative to the DNA.

Changes to soil chemistry across land types

The chemical profiles of the soil were significantly different among land types (PERMANOVA on Gower Distance $R^2 = 0.61$, $P < 0.001$). The largest change to soil organic matter (OM) was a 2-fold decrease in average from the forest to burned site and decreasing further in the plantation (Supp. Table 1). Na levels as well as cation exchange capacity (CEC) of the soil increase from the forest to burned site then decrease in the plantation and abandoned plantation. Nitrate-nitrogen (NO_3-N) and sulfate-sulfur (S) follow the same trend, increasing at first in the burned sites

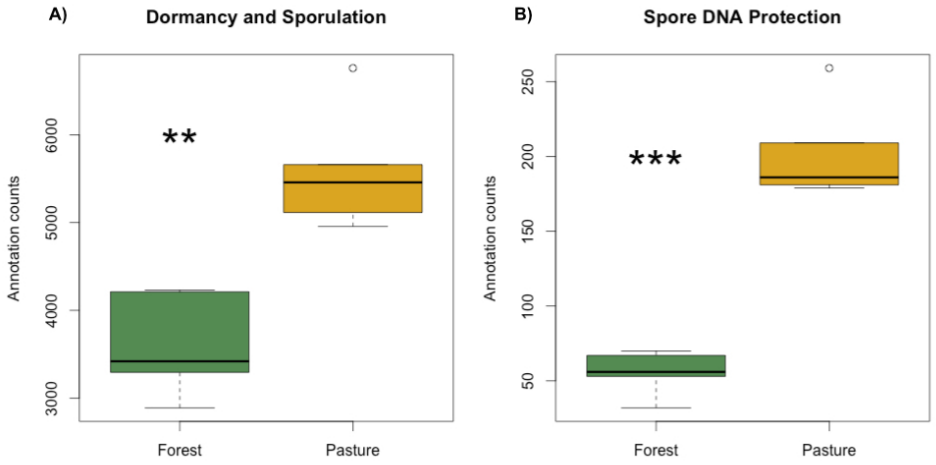
then decreasing. Weak Bray phosphorus (P1) incrementally increase from the forest to burned sites then the burned to plantation sites, but decrease in the abandoned plantation. Finally, the pH of the soil increases from forest to plantation.

Supplemental Table 1: Chemical profiles of the four land types.

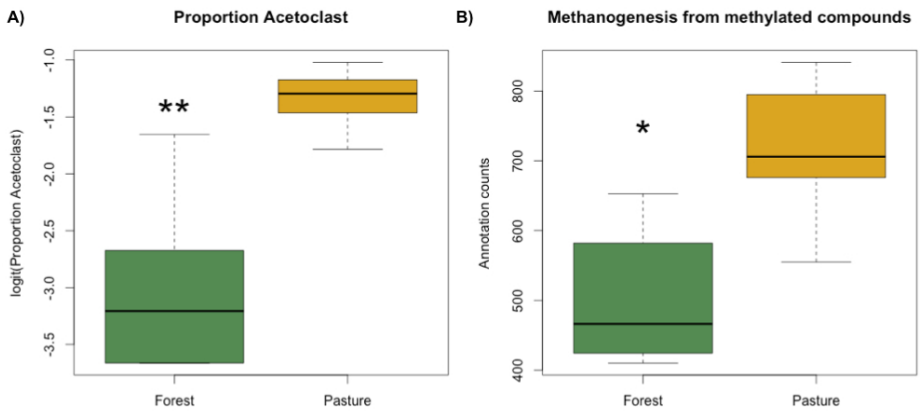
	OM (%)	P (ppm)	pH	Na (ppm)	Buffer pH	CEC (meq/100g)	K_PCT (%)	Mg_PCT (%)	Ca_PCT (%)	NO_3-N (ppm)
orest	6.95±11.73	26.08±12	3.9±0.19	10.18±4.24	6.82±0.12	1.99±1.41	3.86±0.92	7.78±1.1	10.91±1.88	2.62±0.43
urned	3.39±0.82	83.08±20.97	3.98±0.21	25.89±31.57	6.61±0.09	4.78±1.29	3.48±0.6	8.81±1.16	10.37±1.61	14.34±3.73
antation	2.45±0.49	157.08±24.75	4.12±0.18	8.63±3	6.72±0.06	2.76±0.83	4.56±0.88	7.94±1.03	11.05±1.21	2.82±1.03
b. Plantation	2.48±0.84	18.25±8.66	3.94±0.23	6.53±0.46	6.8±0.08	2.13±0.62	2.2±0.8	11.23±2.45	10.13±2.39	3.78±0.85

APPENDIX C

SUPPLEMENTARY MATERIALS FOR CHAPTER IV



Supplementary Figure 1: Land use change impacts the relative abundance of genes related to dormancy and spore protection. A) The relative abundance of dormancy and sporulation genes is higher in pasture than forest soils. B) The relative abundance of genes related to spore DNA protection is higher in pasture relative to forest soils. Significant differences between forest and pasture are denoted as follows: ** $P < 0.01$, *** $P < 0.001$.



Supplementary Figure 2: Land use change impacts methanogen community composition and alternative pathway abundance. A) The proportion of acetoclastic methanogens (relative to hydrogenotrophs) is higher in pasture than forest soils. B) The relative abundance of genes related to methanogenesis from methylated compounds is higher in pasture relative to forest soils. Significant differences between forest and pasture are denoted as follows: ** $P < 0.01$, * $P < 0.05$.

Supplementary Table 1: Methanotroph Classifications

Type	Taxonomic classification
Gammaproteobacteria	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylobacter; Methylobacter sp. BB5.1; Methylobacter sp. BB5.1
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylobacter; Methylobacter sp. HG-1; Methylobacter sp. HG-1
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylobacter; uncultured Methylobacter sp.; uncultured Methylobacter sp.
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylohalobius; Methylohalobius sp. IT-

	9; Methylohalobius sp. IT-9
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylomicrobium; Methylomicrobium album; Methylomicrobium album
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylomicrobium; Methylomicrobium buryatense; Methylomicrobium buryatense
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylomicrobium; Methylomicrobium pelagicum; Methylomicrobium pelagicum
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylomonas; Methylomonas sp. LC 1; Methylomonas sp. LC 1
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylothermus; Methylothermus subterraneus; Methylothermus subterraneus
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylocaldum; Methylocaldum gracile; Methylocaldum gracile
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylocaldum; Methylocaldum sp. 05J-I-7; Methylocaldum sp. 05J-I-7
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylocaldum; Methylocaldum sp. E10a; Methylocaldum sp. E10a
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylocaldum; Methylocaldum szegediense; Methylocaldum szegediense
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylococcus; Methylococcus capsulatus; Methylococcus capsulatus
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylococcus; Methylococcus capsulatus; Methylococcus capsulatus str. Bath
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylothermus; Methylothermus thermalis; Methylothermus thermalis
Alphaproteobacteria	Proteobacteria; Alphaproteobacteria; Rhizobiales; Beijerinckiaceae; Methylocapsa; Methylocapsa acidiphila; Methylocapsa acidiphila
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Beijerinckiaceae; Methylocapsa; Methylocapsa aurea; Methylocapsa sp. KYG
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Beijerinckiaceae; Methylocella; Methylocella palustris; Methylocella palustris
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Beijerinckiaceae; Methylocella; Methylocella silvestris; Methylocella silvestris

Proteobacteria; Alphaproteobacteria; Rhizobiales; Beijerinckiaceae; Methylocella; Methylocella silvestris; Methylocella silvestris BL2
Proteobacteria; Alphaproteobacteria; Rhizobiales; Beijerinckiaceae; Methylocella; Methylocella tundrae; Methylocella tundrae
Proteobacteria; Alphaproteobacteria; Rhizobiales; Beijerinckiaceae; Methylocella; uncultured Methylocella sp.; uncultured Methylocella sp.
Proteobacteria; Alphaproteobacteria; Rhizobiales; Beijerinckiaceae; Methylovirgula; Methylovirgula ligni; Beijerinckiaceae bacterium BW872
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis echinoides; Methylocystis echinoides
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis heyeri; Methylocystis heyeri
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis hirsuta; Methylocystis hirsuta
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis minimus; Methylocystis minimus
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis parvus; Methylocystis parvus
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis rosea; Methylocystis rosea
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. 10; Methylocystis sp. 10
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. 2-19; Methylocystis sp. 2-19
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. 39; Methylocystis sp. 39
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. 5FB1; Methylocystis sp. 5FB1
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. 5FB2; Methylocystis sp. 5FB2
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. 62/12; Methylocystis sp. 62/12
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. Ch22;

Methylocystis sp. Ch22
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. F10V2a; Methylocystis sp. F10V2a
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. H2s; Methylocystis sp. H2s
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. H9a; Methylocystis sp. H9a
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. IMET 10484; Methylocystis sp. IMET 10484
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. KS3; Methylocystis sp. KS3
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. LW5; Methylocystis sp. LW5
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. M; Methylocystis sp. M
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. m1511; Methylocystis sp. m1511
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. m231; Methylocystis sp. m231
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. Pi5/4; Methylocystis sp. Pi5/4
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; Methylocystis sp. WI14; Methylocystis sp. WI14
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylocystis; uncultured Methylocystis sp.; uncultured Methylocystis sp.
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylopila; Methylopila capsulata; Methylopila capsulata
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylosinus; Methylosinus acidophilus; Methylosinus acidophilus
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylosinus; Methylosinus sp. D15a; Methylosinus sp. D15a
Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylosinus; Methylosinus sp. LW2; Methylosinus sp. LW2

	Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylosinus; Methylosinus sp. LW3; Methylosinus sp. LW3
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylosinus; Methylosinus sp. PW1; Methylosinus sp. PW1
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylosinus; Methylosinus sporium; Methylosinus sporium
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylosinus; Methylosinus trichosporium; Methylosinus trichosporium
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylosinus; Methylosinus trichosporium; Methylosinus trichosporium OB3b
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; Pleomorphomonas; Pleomorphomonas oryzae; Pleomorphomonas oryzae
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; unclassified (derived from Methylocystaceae); methanotroph E10; methanotroph E10
	Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylocystaceae; unclassified (derived from Methylocystaceae); uncultured Methylocystaceae bacterium; uncultured Methylocystaceae bacterium
Verrucomicrobia	Verrucomicrobia; unclassified (derived from Verrucomicrobia); Methylacidiphilales; Methylacidiphilaceae; Methylacidiphilum; Methylacidiphilum fumariolicum; Methylacidiphilum fumariolicum
	Verrucomicrobia; unclassified (derived from Verrucomicrobia); Methylacidiphilales; Methylacidiphilaceae; Methylacidiphilum; Methylacidiphilum infernorum; Methylacidiphilum infernorum V4
	Verrucomicrobia; unclassified (derived from Verrucomicrobia); Methylacidiphilales; Methylacidiphilaceae; Methylacidiphilum; Methylacidiphilum kamchatkense; Methylacidiphilum kamchatkense
Unknown	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; unclassified (derived from Methylococcaceae); methanotroph FL-DIKO; methanotroph FL-DIKO
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; unclassified (derived from Methylococcaceae); Methylococcaceae bacterium ET-HIRO; Methylococcaceae bacterium ET-HIRO
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; unclassified (derived from Methylococcaceae); Methylococcaceae bacterium ET-SHO; Methylococcaceae bacterium ET-SHO
	Proteobacteria; Gammaproteobacteria; Methylococcales; Methylococcaceae; unclassified (derived from Methylococcaceae); Methylococcaceae bacterium T2-1; Methylococcaceae bacterium T2-1
Creno-thrix	Proteobacteria; Gammaproteobacteria; Methylococcales; Crenotrichaceae; Crenothrix; Crenothrix polyspora; Crenothrix polyspora

Supplementary Table 2: Methanogen Classifications

Type	Taxonomic classification
Hydrogenotroph	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobacterium; Methanobacterium bryantii; Methanobacterium bryantii
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobacterium; Methanobacterium formicicum; Methanobacterium formicicum
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobacterium; Methanobacterium kanagiense; Methanobacterium kanagiense
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobacterium; Methanobacterium oryzae; Methanobacterium oryzae
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobacterium; Methanobacterium palustre; Methanobacterium palustre
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobacterium; Methanobacterium sp. MB4; Methanobacterium sp. MB4
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobrevibacter; Methanobrevibacter arboriphilus; Methanobrevibacter arboriphilus
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobrevibacter; Methanobrevibacter ruminantium; Methanobrevibacter ruminantium
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobrevibacter; Methanobrevibacter smithii; Methanobrevibacter smithii
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobrevibacter; Methanobrevibacter smithii; Methanobrevibacter smithii ATCC 35061
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanobrevibacter; Methanobrevibacter smithii; Methanobrevibacter smithii DSM 2374
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanosphaera; Methanosphaera stadtmanae; Methanosphaera stadtmanae DSM 3091
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanothermobacter; Methanothermobacter thermautotrophicus; Methanothermobacter thermautotrophicus
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanothermobacter; Methanothermobacter thermautotrophicus; Methanothermobacter thermautotrophicus str. Delta H
	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales; Methanobacteriaceae; Methanothermobacter;

	Methanothermobacter wolfeii; Methanothermobacter wolfeii
	Archaea; Euryarchaeota; Methanococci; Methanococcales; Methanocaldococcaceae; Methanocaldococcus; Methanocaldococcus vulcanius; Methanocaldococcus vulcanius M7
	Archaea; Euryarchaeota; Methanococci; Methanococcales; Methanococcaceae; Methanococcus; Methanococcus aeolicus; Methanococcus aeolicus
	Archaea; Euryarchaeota; Methanococci; Methanococcales; Methanococcaceae; Methanococcus; Methanococcus aeolicus; Methanococcus aeolicus Nankai-3
	Archaea; Euryarchaeota; Methanococci; Methanococcales; Methanococcaceae; Methanococcus; Methanococcus maripaludis; Methanococcus maripaludis C6
	Archaea; Euryarchaeota; Methanococci; Methanococcales; Methanococcaceae; Methanococcus; Methanococcus voltae; Methanococcus voltae A3
	Archaea; Euryarchaeota; Methanomicrobia; Methanocellales; Methanocellaceae; Methanocella; Methanocella paludicola; Methanocella paludicola SANAE
	Archaea; Euryarchaeota; Methanomicrobia; Methanomicrobiales; Methanocorpusculaceae; Methanocorpusculum; Methanocorpusculum bavaricum; Methanocorpusculum bavaricum
	Archaea; Euryarchaeota; Methanomicrobia; Methanomicrobiales; Methanomicrobiaceae; Methanoculleus; Methanoculleus chikugoensis; Methanoculleus chikugoensis
	Archaea; Euryarchaeota; Methanomicrobia; Methanomicrobiales; Methanomicrobiaceae; Methanoculleus; Methanoculleus thermophilus; Methanoculleus thermophilus
	Archaea; Euryarchaeota; Methanomicrobia; Methanomicrobiales; Methanomicrobiaceae; Methanofollis; Methanofollis liminatans; Methanofollis liminatans
	Archaea; Euryarchaeota; Methanomicrobia; Methanomicrobiales; Methanospirillaceae; Methanospirillum; Methanospirillum hungatei; Methanospirillum hungatei
	Archaea; Euryarchaeota; Methanomicrobia; Methanomicrobiales; Methanospirillaceae; Methanospirillum; Methanospirillum hungatei; Methanospirillum hungatei JF-1
	Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae; Methanopyrus; Methanopyrus kandleri; Methanopyrus kandleri AV19
Acetoclast	Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methanosaetaceae; Methanosaeta; Methanosaeta concilii; Methanosaeta concilii
	Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methanosaetaceae; Methanosaeta; Methanosaeta harundinacea; Methanosaeta harundinacea
	Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methanosaetaceae; Methanosaeta; Methanosaeta thermophila; Methanosaeta thermophila PT

Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methanosarcinaceae; Methanococcoides; Methanococcoides burtonii; Methanococcoides burtonii DSM 6242
Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methanosarcinaceae; Methanolobus; Methanolobus zinderi; Methanolobus zinderi
Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methanosarcinaceae; Methanosarcina; Methanosarcina acetivorans; Methanosarcina acetivorans C2A
Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methanosarcinaceae; Methanosarcina; Methanosarcina barkeri; Methanosarcina barkeri str. Fusaro
Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methanosarcinaceae; Methanosarcina; Methanosarcina mazei; Methanosarcina mazei Go1

Supplementary Table 3: Competitor-Stress Tolerator-Ruderal Methanotroph Classifications

Genus	Classification
Methylobacter	C, C-R
Methylocaldum	C, C-R, R
Methylococcus	C, C-R, R
Methylomicrobium	C, C-R
Methylomonas	C, C-R
Methylocapsa	S
Methylocella	S
Methylocystis	C-S, S, S-R
Methylosinus	C-S, S, S-R

REFERENCES CITED

Alele P.O., Sheil D., Surget-Groba Y., Lingling S., & Cannon C.H. (2014) How does conversion of natural tropical rainforest ecosystems affect soil bacterial and fungal communities in the nile river watershed of Uganda? *PLoS One*, **9**, e104818.

Anderson M.J., Crist T.O., Chase J.M., Vellend M., Inouye B.D., Freestone A.L., Sanders N.J., Cornell H. V, Comita L.S., Davies K.F., Harrison S.P., Kraft N.J.B., Stegen J.C., & Swenson N.G. (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecol. Lett.*, **14**, 19–28.

Anthony C. (1991) Assimilation of Carbon by Methylotrophs. *Biology of Methylotrophs* (ed. by I. Goldberg and J.S. Rokem), pp. 79–109. Butterworth-Heinemann, Jerusalem.

Barberán A., Fernández-Guerra A., Bohannan B.J.M., & Casamayor E.O. (2012) Exploration of community traits as ecological markers in microbial metagenomes. *Mol. Ecol.*, **21**, 1909–17.

Barnard R.L., Osborne C.A., & Firestone M.K. (2013) Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. *ISME J.*, **7**, 2229–41.

Bell T. (2010) Experimental tests of the bacterial distance-decay relationship. *ISME J.*, **4**, 1357–65.

Berg G. & Smalla K. (2009) Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol. Ecol.*, **68**, 1–13.

Bissett A., Richardson A.E., Baker G., Wakelin S., & Thrall P.H. (2010) Life history determines biogeographical patterns of soil bacterial communities over multiple spatial scales. *Mol. Ecol.*, **19**, 4315–4327.

Blagodatskaya E. & Kuzyakov Y. (2013) Active microorganisms in soil : Critical review of estimation criteria and approaches. *Soil Biol. Biochem.*, **67**, 192–211.

Blazewicz S.J., Barnard R.L., Daly R. a, & Firestone M.K. (2013) Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J.*, **7**, 2061–8.

Boddey R.M., Macedo R., Tarré R.M., Ferreira E., de Oliveira O.C., de P. Rezende C., Cantarutti R.B., Pereira J.M., Alves B.J.R., & Urquiaga S. (2004) Nitrogen cycling in Brachiaria pastures: the key to understanding the process of pasture decline. *Agric. Ecosyst. Environ.*, **103**, 389–403.

Bodelier P.L.E., Bär-Gilissen M.-J., Meima-Franke M., & Hordijk K. (2012) Structural and functional response of methane-consuming microbial communities to different flooding regimes in riparian soils. *Ecol. Evol.*, **2**, 106–27.

Bodelier P.L.E., Meima-Franke M., Hordijk C. a, Steenbergh A.K., Hefting M.M., Bodrossy L., von Bergen M., & Seifert J. (2013) Microbial minorities modulate methane consumption through niche partitioning. *ISME J.*, **7**, 2214–28.

Bossio D.A., Girvan M.S., Verchot L., Bullimore J., Borelli T., Albrecht A., Scow K.M., Ball a S., Pretty J.N., & Osborn a M. (2005) Soil microbial community response to land use change in an agricultural landscape of western Kenya. *Microb. Ecol.*, **49**, 50–62.

Ter Braak C.J.F. (1995) Ordination. *Data Analysis in Community and Landscape Ecology* (ed. by R.H.G. Jongman, C.J.F. ter Braak, and O.F.R. van Tongeren), pp. 152–153. Cambridge University Press, Cambridge.

Bray J. & Curtis J. (1957) An ordination of upland forest communities of southern Wisconsin. *Ecol. Monogr.*, **27**, 325–349.

Butler R.A. & Laurance W.F. (2008) New strategies for conserving tropical forests. *Trends Ecol. Evol.*, **23**, 469–72.

Cam E., Nichols J.D., Hines J.E., Sauer J.R., Alpizar-Jara R., & Flather C.H. (2002) Disentangling sampling and ecological explanations underlying species-area relationships. *Ecology*, **83**, 1118–1130.

Caporaso J.G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.D., Costello E.K., Fierer N., Gonzalez Peña A., 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 Intensity normalization improves color calling in SOLiD sequencing. *Nat. Methods*, **7**, 335–336.

Cardinale B.J., Duffy J.E., Gonzalez A., Hooper D.U., Perrings C., Venail P., Narwani A., Mace G.M., Tilman D., & Wardle D.A. (2012) Biodiversity loss and its impact on humanity. *Nature*, **486**, 59–67.

Carmo J.B. Do, de Sousa Neto E.R., Duarte-Neto P.J., Ometto J.P.H.B., & Martinelli L.A. (2012) Conversion of the coastal Atlantic forest to pasture: Consequences for the nitrogen cycle and soil greenhouse gas emissions. *Agric. Ecosyst. Environ.*, **148**, 37–43.

Chao A., Colwell R.K., Lin C.W., & Gotelli N.J. (2009) Sufficient sampling for asymptotic minimum species richness estimators. *Ecology*, **90**, 1125–1133.

Chapman H.D. (1965) Cation-exchange capacity. *Methods of soil analysis* pp. 891–901. American Society of Agronomy, Soil Science Society of America,

Collomb J.-G., Mikissa J.-B., Minnemeyer S., Mundunga S., Nzao H.N., Madouma J., Mapaga J. de D., Mikolo C., Rabenkogo N., Akagah S., Bayani-Ngoye E., & Moufama A. (2000) A first look at logging in Gabon. .

Condit R., Pitman N., Leigh E.G., Chave J., Terborgh J., Foster R.B., Núñez P., Aguilar S., Valencia R., Villa G., Muller-Landau H.C., Losos E., & Hubbell S.P. (2002) Beta-diversity in tropical forest trees. *Science*, **295**, 666–9.

Conrad R. (1999) Contribution of hydrogen to methane production and control of hydrogen concentrations in methanogenic soils and sediments. *FEMS Microbiol. Ecol.*, **28**, 193–202.

Dean W.E. (1974) Determination of carbonate and organic matter in calcareous sediments and sedimentary rocks by loss on ignition: comparison with other methods. *J. Sediment. Res.*, **44**, .

Dedysh S.N., Liesack W., Khmelenina V.N., Suzina N.E., Trotsenko Y.A., Semrau J.D., Bares A.M., Panikov N.S., & Tiedje J.M. (2000) Methylocella palustris gen. nov., sp. nov., a new methane-oxidizing acidophilic bacterium from peat bogs, representing a novel subtype of serine-pathway methanotrophs. *Int. J. Syst. Evol. Microbiol.*, **50**, 955–969.

Delègue M.-A., Fuhr M., Schwartz D., Mariotti A., & Nasi R. (2001) Recent Origin of a Large Part of the Forest Cover in the Gabon Coastal Area Based on Stable Carbon Isotope Data. *Oecologia*, **129**, 106–113.

Diamond J.M. (1975) The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biol. Conserv.*, **7**, 129–146.

Dias-Filho M.B., Davidson E.A., & Reis de Carvalho C.. (2001) Linking biogeochemical cycles to cattle pasture management and sustainability in the Amazon Basin. *The Biogeochemistry of the Amazon Basin* (ed. by M.E. McClain, R.L. Victoria, and J.E. Richey), pp. 84–115. Oxford University Press, New York.

Dimitrakopoulos P.G. & Schmid B. (2004) Biodiversity effects increase linearly with biotope space. *Ecol. Lett.*, **7**, 574–583.

Dirzo R. & Raven P.H. (2003) Global State of Biodiversity and Loss. *Annu. Rev. Environ. Resour.*, **28**, 137–167.

Drenovsky R.E., Vo D., Graham K.J., & Scow K.M. (2004) Soil water content and organic carbon availability are major determinants of soil microbial community composition. *Microb. Ecol.*, **48**, 424–30.

Edgar R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460–1.

Edgar R.C. (2013) UPARSE : highly accurate OTU sequences from microbial amplicon reads. *Nature*, **10**, 996–1000.

Fadrosh D.W., Ma B., Gajer P., Sengamalay N., Ott S., Brotman R.M., & Ravel J. (2014) An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*, **2**, 1–7.

Falkowski P.G., Fenchel T., & Delong E.F. (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science (80-.).*, **320**, 1034–1039.

Feigl B., Cerri C., Piccolo M., Noronha, Norberto Augusti K., & Melillo, Jerry Eschenbrenner, Vincent Melo L. (2006) Biological survey of a low-productivity pasture in Rondonia state. *Outlook Agric.*, **35**, 199–208.

Fernandes S.A.P., Bernoux M., Cerri C.C., Feigl B.J., & Piccolo M.C. (2002) Seasonal variation of soil chemical properties and CO2 and CH4 fluxes in unfertilized and P-fertilized pastures in an Ultisol of the Brazilian Amazon. *Geoderma*, **107**, 227–241.

Fierer N., Barberán A., & Laughlin D.C. (2014) Seeing the forest for the genes: using metagenomics to infer the aggregated traits of microbial communities. *Front. Microbiol.*, **5**, 1–6.

Fierer N., Bradford M.A., & Jackson R.B. (2007) Toward and ecological classification of soil bacteria. *Ecology*, **88**, 1354–1364.

Fierer N. & Jackson R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 626–31.

Finlay B.J. (2002) Global dispersal of free-living microbial eukaryote species. *Science (80-.).*, **296**, 1061–3.

Finlay B.J., Esteban G.F., & Fenchel T. (2004) Protist Diversity is Different? *Protist*, **155**, 15–22.

Foley J. a, Defries R., Asner G.P., Barford C., Bonan G., Carpenter S.R., Chapin F.S., Coe M.T., Daily G.C., Gibbs H.K., Helkowski J.H., Holloway T., Howard E. a, Kucharik C.J., Monfreda C., Patz J. a, Prentice I.C., Ramankutty N., & Snyder P.K. (2005) Global consequences of land use. *Science*, **309**, 570–4.

Fox R.L., Olson R.A., & Rhoades H.F. (1964) Evaluating the sulfur status of soils by plant and soil tests. *Soil Sci. Soc. Am. J.*, **28**, 243–246.

Franklin R.B. & Mills A.L. (2003) Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. *FEMS Microbiol. Ecol.*, **44**, 335–346.

Freitag T.E. & Prosser J.I. (2009) Correlation of methane production and functional gene transcriptional activity in a peat soil. *Appl. Environ. Microbiol.*, **75**, 6679–87.

Freitag T.E., Toet S., Ineson P., & Prosser J.I. (2010) Links between methane flux and transcriptional activities of methanogens and methane oxidizers in a blanket peat bog. *FEMS Microbiol. Ecol.*, no–no.

Fuhrman J.A. & Campbell L. (1998) Marine ecology: microbial microdiversity. *Nature*, **393**, 410–411.

Goslee S.C. & Urban D.L. (2007) The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.*, **22**, 1–19.

Gower J.C. (1971) A general coefficient of similarity and some of its properties. *Biometrics*, **27**, 857–871.

Green J. & Bohannan B.J.M. (2007) Biodiversity scaling relationships: are microorganisms fundamentally different? *Scaling Biodiversity* (ed. by D. Storch, P.A. Marquet, and J.H. Brown), pp. 129–149. Cambridge University Press, Cambridge.

Green J.L., Bohannan B.J.M., & Whitaker R.J. (2008) Microbial biogeography: from taxonomy to traits. *Science*, **320**, 1039–43.

Green J.L., Holmes A.J., Westoby M., Oliver I., Briscoe D., Dangerfield M., Gillings M., & Beattie A.J. (2004) Spatial scaling of microbial eukaryote diversity. *Nature*, **432**, 747–50.

Grime J.P. (1977) Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *Am. Nat.*, **111**, 1169–1194.

Hamaoui G.S., Rodrigues J.L.M., Bohannan B.J.M., Tiedje J.M., & N??sslein K. (2016) Land-use change drives abundance and community structure alterations of thaumarchaeal ammonia oxidizers in tropical rainforest soils in Rondonia, Brazil. *Appl. Soil Ecol.*, **107**, 48–56.

Hanson C.A., Fuhrman J.A., Horner-Devine M.C., & Martiny J.B.H. (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.*, **10**, 497–506.

Hanson R.S. & Hanson T.E. (1996) Methanotrophic bacteria. *Microbiol. Rev.*, **60**, 439–71.

Hedderich R. & Whitman W.B. (2013) Physiology and Biochemistry of the Methane-Producing Archaea. *The Prokaryotes* pp. 635–662.

Van der Heijden M.G. a, Bardgett R.D., & van Straalen N.M. (2008) The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol. Lett.*, **11**, 296–310.

Herpin U., Cerri C.C., Conceic M., Markert B., Enzweiler J., Friese K., & Breulmann G. (2002) Biogeochemical dynamics following land use change from forest to pasture in a humid tropical area (Rondonia, Brazil): a multi-element approach by means of XRF-spectroscopy. *Sci. Total Environ.*, **286**, 97–109.

Hillebrand H., Watermann F., Karez R., & Berninger U.G. (2001) Differences in species richness patterns between unicellular and multicellular organisms. *Oecologia*, **126**, 114–124.

Ho A., Kerckhof F.-M., Luke C., Reim A., Krause S., Boon N., & Bodelier P.L.E. (2013) Conceptualizing functional traits and ecological characteristics of methane-oxidizing bacteria as life strategies. *Environ. Microbiol. Rep.*, **5**, 335–45.

Holmes A.J., Roslev P., McDonald I.R., Iversen N., Henriksen K., & Murrell J.C. (1999) Characterization of methanotrophic bacterial populations in soils showing atmospheric methane uptake. *Appl. Environ. Microbiol.*, **65**, 3312–3318.

Hooper D.U., Adair E.C., Cardinale B.J., Byrnes J.E.K., Hungate B. a, Matulich K.L., Gonzalez A., Duffy J.E., Gamfeldt L., & O'Connor M.I. (2012) A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature*, **486**, 105–8.

Horner-Devine M.C., Lage M., Hughes J.B., & Bohannan B.J.M. (2004) A taxa-area relationship for bacteria. *Nature*, **432**, 750–3.

Houghton R.A. (1994) The worldwide extent of land-use change. *Bioscience*, **44**, 305–313.

Jesus E., Marsh T.L., Tiedje J.M., & de S Moreira F.M. (2009) Changes in land use alter the structure of bacterial communities in Western Amazon soils. *ISME J.*, **3**, 1004–11.

Jost L., Chao A., & Chazdon R.L. (2011) Compositional similarity and Beta diversity. *Biological Diversity: frontiers in measurement and assessment* (ed. by A.E. Magurran and B. McGill), pp. 66–84. Oxford University Press, Oxford.

Jurasinski G. & Retzer V. (2012) simba: A collection of functions for similarity analysis of vegetation data. .

Kamprath E.J. & Watson M.E. (1980) Conventional soil and tissue tests for assessing the phosphorus status of soils. *role phosphorus Agric.*, 433–469.

Knief C. (2015) Diversity and habitat preferences of cultivated and uncultivated aerobic methanotrophic bacteria evaluated based on pmoA as molecular marker. *Front. Microbiol.*, **6**, .

Knief C., Vanitchung S., Harvey N.W., Dunfield P.F., Chidthaisong A., & Conrad R. (2005) Diversity of Methanotrophic Bacteria in Tropical Upland Soils under Different Land Uses. *Appl. Environ. Microbiol.*, **71**, 3826–3831.

Kolb S. (2009) The quest for atmospheric methane oxidizers in forest soils. *Environ. Microbiol. Rep.*, **1**, 336–346.

Kozich J.J., Westcott S.L., Baxter N.T., Highlander S.K., & Schloss P.D. (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.*, **79**, 5112–20.

Krause S., Le Roux X., Niklaus P. a, Van Bodegom P.M., Lennon J.T., Bertilsson S., Grossart H.-P., Philippot L., & Bodelier P.L.E. (2014) Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Front. Microbiol.*, **5**, 251.

Lauber C.L., Ramirez K.S., Aanderud Z., Lennon J., & Fierer N. (2013) Temporal variability in soil microbial communities across land-use types. *ISME J.*, **7**, 1641–50.

Laurance W.F., Alonso A., Lee M., & Campbell P. (2006) Challenges for forest conservation in Gabon, Central Africa. *Futures*, **38**, 454–470.

Laurance W.F., Sayer J., & Cassman K.G. (2014) Agricultural expansion and its impacts on tropical nature. *Trends Ecol. Evol.*, **29**, 107–116.

Lee M.E., Alonso A., Dallmeier F., Campbell P., & Pauwels O.S.G. (2006) The Gamba Complex of Protected Areas : An Illustration of Gabon's Biodiversity. *Bull. Biol. Soc. Washingt.*, 229–242.

Lennon J.T. & Jones S.E. (2011) Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.*, **9**, 119–30.

Livingstone D.A. (2001) A geological perspective on the conservation of African forests. *African Rain Forest Ecology and Conservation* (ed. by W. Weber, L.J.T.

White, A. Vedder, and L. Naughton-Treves), pp. 50–55. Yale University Press, New Haven and London.

Locey K.J. (2010) Synthesizing traditional biogeography with microbial ecology: The importance of dormancy. *J. Biogeogr.*, **37**, 1835–1841.

Louca S., Wegener Parfrey L., & Doebli M. (2016) Decoupling function and taxonomy in the global ocean microbiome. *Science (80-.).*, **353**, 1272–1277.

Love M.I., Huber W., & Anders S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.

Maechler M., Rousseeuw P., Struyf A., Hubert M., & Hornik K. (2016) cluster: Cluster analysis basics and extensions. .

Martiny A.C., Treseder K., & Pusch G. (2013) Phylogenetic conservatism of functional traits in microorganisms. *ISME J.*, **7**, 830–8.

Martiny J.B.H., Bohannan B.J.M., Brown J.H., Colwell R.K., Fuhrman J. a, Green J.L., Horner-Devine M.C., Kane M., Krumins J.A., Kuske C.R., Morin P.J., Naeem S., Ovreås L., Reysenbach A.-L., Smith V.H., & Staley J.T. (2006) Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.*, **4**, 102–112.

Martiny J.B.H., Eisen J.A., Penn K., Allison S.D., & Horner-devine M.C. (2011) Drivers of bacterial β -diversity depend on spatial scale. *Proc. Natl. Acad. Sci.*, **108**, 7850–7854.

Maxfield P.J., Hornibrook E.R.C., & Evershed R.P. (2008) Acute impact of agriculture on high-affinity methanotrophic bacterial populations. *Environ. Microbiol.*, **10**, 1917–1924.

McCalley C.K., Woodcroft B.J., Hodgkins S.B., Wehr R. a., Kim E.-H., Mondav R., Crill P.M., Chanton J.P., Rich V.I., Tyson G.W., & Saleska S.R. (2014) Methane dynamics regulated by microbial community response to permafrost thaw. *Nature*, **514**, 478–481.

Memiaghe H.R., Lutz J. a., Korte L., Alonso A., & Kenfack D. (2016) Ecological Importance of Small-Diameter Trees to the Structure, Diversity and Biomass of a Tropical Evergreen Forest at Rabi, Gabon. *PLoS One*, **11**, e0154988.

Meyer F., Paarmann D., D'Souza M., Olson R., Glass E.M., Kubal M., Paczian T., Rodriguez a, Stevens R., Wilke a, Wilkening J., & Edwards R. a (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, **9**, 386.

Mirza B.S., Potisap C., Nüsslein K., Bohannan B.J.M., & Rodriguesa J.L.M. (2014) Response of free-living nitrogen-fixing microorganisms to land use change in the amazon rainforest. *Appl. Environ. Microbiol.*, **80**, 281–288.

De Moraes J.F.L., Volkoff B., Cerri C.C., & Bernoux M. (1996) Soil properties under Amazon forest and changes due to pasture installation in Rondônia, Brazil. *Geoderma*, **70**, 63–81.

Morlon H., Chuyong G., Condit R., Hubbell S., Kenfack D., Thomas D., Valencia R., & Green J.L. (2008) A general framework for the distance-decay of similarity in ecological communities. *Ecol. Lett.*, **11**, 904–917.

Mueller R.C., Paula F.S., Mirza B.S., Rodrigues J.L., Nüsslein K., & Bohannan B.J. (2014) Links between plant and fungal communities across a deforestation chronosequence in the Amazon rainforest. *ISME J.*, **8**, 1548–50.

Mueller R.C., Rodrigues J.L., Nüsslein K., & Bohannan B.J. (2016) Land use change in the Amazon rainforest favors generalist fungi. *Funct. Ecol.*, .

Myhre G., Shindell D., Bréon F.M., Collins W., Fuglestvedt J., Huang J., Koch D., Lamarque J.-F., Lee D., Mendoza B., Nakajima T., Robock A., Stephens G., Takemura T., & Zhang H. (2013) 2013: Anthropogenic and Natural Radiative Forcing. 659–740.

Nannipieri P., Ascher J., Ceccherini M., Landi L., Pietramellara G., & Renella G. (2003) Microbial diversity and soil functions. *Eur. J. Soil Sci.*, **54**, 655–670.

Naughton-Treves L. & Weber W. (2001) Human dimensions of the African rain forest. *African Rain Forest Ecology and Conservation* (ed. by W. Weber, L.J.T. White, A. Vedder, and L. Naughton-Treves), pp. 30–49. Yale University Press, New Haven and London.

Navarrete A. a, Tsai S.M., Mendes L.W., Faust K., de Hollander M., Cassman N. a, van Veen J. a, & Kuramae E.E. (2015a) Soil microbiome responses to the short-term effects of Amazonian deforestation. *Mol. Ecol.*, 2433–2448.

Navarrete A. a, Venturini A.M., Meyer K.M., Klein A.M., James M., Bohannan B.J., Nüsslein K., Tsai S.M., & Rodrigues J.L. (2015b) Differential response of Acidobacteria subgroups to forest-to-pasture conversion and their biogeographic patterns in the western Brazilian Amazon. *Front. Microbiol.*, **6**, 1–10.

Navarrete A., Tsai S.M., Mendes L.W., Faust K., de Hollander M., Cassman N. a, van Veen J. a, & Kuramae E.E. (2015c) Soil microbiome responses to the short-term effects of Amazonian deforestation. *Mol. Ecol.*, .

Nazaries L., Pan Y., Bodrossy L., Baggs E.M., Millard P., Murrell J.C., & Singh B.K. (2013) Evidence of microbial regulation of biogeochemical cycles from a study on methane flux and land use change. *Appl. Environ. Microbiol.*, **79**, 4031–40.

Neill C., Melillo J.M., Steudler P. a., Cerri C.C., De Moraes J.F.L., Piccolo M.C., & Brito M. (1997a) Soil carbon and nitrogen stocks following forest clearing for pasture in the southwestern Brazilian Amazon. *Ecol. Appl.*, **7**, 1216–1225.

Neill C., Piccolo M.C., Cerri C.C., Steudler P. a., Melillo J.M., & Brito M. (1997b) Net nitrogen mineralization and net nitrification rates in soils following deforestation for pasture across the southwestern Brazilian Amazon Basin landscape. *Oecologia*, **110**, 243–252.

Neill C., Steudler P.A., Garcia-Montiel D.C., Melillo J.M., Feigl B.J., Piccolo M.C., & Cerri C.C. (2005) Rates and controls of nitrous oxide and nitric oxide emissions following conversion of forest to pasture in Rondonia. *Nutr. Cycl. Agroecosystems*, **71**, 1–15.

Nekola J.C. & White P.S. (1999) The distance decay of similarity in biogeography and ecology. *J. Biogeogr.*, **26**, 867–878.

Ofek M., Voronov-Goldman M., Hadar Y., & Minz D. (2014) Host signature effect on plant root-associated microbiomes revealed through analyses of resident vs . active communities. *Environ. Microbiol.*, **16**, 2157–2167.

Oksanen J., Blanchet F.G., Roeland K., Legendre P., Minchin P., O'Hara R.B., Simpson G., Solymos P., Stevens M.H.H., & Wagner H. (2015) vegan: Community ecology package. .

Olden J.D., Leroy Poff N., Douglas M.R., Douglas M.E., & Fausch K.D. (2004) Ecological and evolutionary consequences of biotic homogenization. *Trends Ecol. Evol.*, **19**, 18–24.

Olden J.D. & Poff N.L. (2003) Toward a mechanistic understanding and prediction of biotic homogenization. *Am. Nat.*, **162**, 442–460.

Olsen S.R. (1954) Estimation of available phosphorus in soils by extraction with sodium bicarbonate. *United States Dep. Agric. Circ.*, .

Östman Ö., Drakare S., Kritzberg E.S., Langenheder S., Logue J.B., & Lindström E.S. (2010) Regional invariance among microbial communities. *Ecol. Lett.*, **13**, 118–127.

Overbeek R., Olson R., Pusch G.D., Olsen G.J., Davis J.J., Disz T., Edwards R. a., Gerdes S., Parrello B., Shukla M., Vonstein V., Wattam a. R., Xia F., & Stevens R. (2014)

The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.*, **42**, D206–D214.

Pace N.R. (1997) A Molecular View of Microbial Diversity and the Biosphere. *Science (80-.).*, **276**, 734–740.

Papke R.T., Ramsing N.B., Bateson M.M., & Ward D.M. (2003) Geographical isolation in hot spring cyanobacteria. *Environ. Microbiol.*, **5**, 650–659.

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. *Mol. Ecol.*, **23**, 2988–99.

Preston F.W. (1960) Time and Space and the Variation of Species. *Ecology*, **41**, 612–627.

R Development Core Team (2010) R: A language and environment for statistical computing. .

Rodrigues J.L.M., Pellizari V.H., Mueller R., Baek K., Jesus E.D.C., Paula F.S., Mirza B., Hamaoui G.S., Tsai S.M., Feigl B., Tiedje J.M., Bohannan B.J.M., & Nüsslein K. (2013) Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 988–93.

Romanowicz K.J., Freedman Z.B., Upchurch R. a., Argiroff W. a., & Zak D.R. (2016) Active microorganisms in forest soils differ from the total community yet are shaped by the same environmental factors: the influence of pH and soil moisture. *FEMS Microbiol. Ecol.*, fiw149.

Rosenzweig M.L. (1995) *Species diversity in space and time.* Cambridge University Press, Cambridge.

Schimel J.P. & Gulledge J. (1998) Microbial community structure and global trace gases. *Soil Biota Glob. Chang.*, **4**, 745–758.

Seghers D., Top E.M., Reheul D., Bulcke R., Boeckx P., Verstraete W., & Siciliano S.D. (2003) Long-term effects of mineral versus organic fertilizers on activity and structure of the methanotrophic community in agricultural soils. *Environ. Microbiol.*, **5**, 867–877.

Simberloff D. & Abele L.G. (1982) Refuge design and island biogeographic theory: effects of fragmentation. *Am. Nat.*, **120**, 41–50.

Singh B.K., Tate K.R., Kolipaka G., Hedley C.B., Macdonald C. a., Millard P., & Murrell J.C. (2007) Effect of afforestation and reforestation of pastures on the activity and population dynamics of methanotrophic bacteria. *Appl. Environ. Microbiol.*, **73**, 5153–5161.

Sloan W.T., Lunn M., Woodcock S., Head I.M., Nee S., & Curtis T.P. (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ. Microbiol.*, **8**, 732–740.

Smart S.M., Thompson K., Marrs R.H., Le Duc M.G., Maskell L.C., & Firbank L.G. (2006) Biotic homogenization and changes in species diversity across human-modified ecosystems. *Proc. Biol. Sci.*, **273**, 2659–65.

Soininen J., McDonald R., & Hillebrand H. (2007) The distance decay of similarity in ecological communities. *Ecography (Cop.).*, **30**, 3–12.

Soule M.E. & Simberloff D. (1986) What do genetics and ecology tell us about the design of nature reserves? *Biol. Conserv.*, **35**, 19–40.

Steudler A., Melillo M., Feigl J., Piccolo C., & Cerri C. (1996) Consequence of forest-to-pasture conversion on CH4 fluxes in the Brazilian Amazon Basin. *J. Geophys. Res.*, **101**, 547–554.

Storch D. & Šizling A.L. (2008) The concept of taxon invariance in ecology: Do diversity patterns vary with changes in taxonomic resolution? *Folia Geobot.*, **43**, 329–344.

Sul W.J., Asuming-brempong S., Wang Q., Tourlousse D.M., Penton C.R., Deng Y., Rodrigues J.L.M., Adiku S.G.K., Jones J.W., Zhou J., Cole J.R., & Tiedje J.M. (2013) Tropical agricultural land management influences on soil microbial communities through its effect on soil organic carbon. *Soil Biol. Biochem.*, **65**, 33–38.

Tchawa Yimga M., Dunfield P.F., Ricke P., & Liesack W. (2003) Wide Distribution of a Novel pmoA-Like Gene Copy among Type II Methanotrophs, and its Expression in Methylocystis Strain SC2. *Appl. Environ. Microbiol.*, **69**, 5593–5602.

Thompson K., Hodgson J.G., & Gaston K.J. (1998) Abundance–range size relationships in the herbaceous flora of central England. *J. Ecol.*, **86**, 439–448.

Tiedje J.M. (1995) Approaches to the comprehensive evaluation of prokaryote diversity of a habitat. *Microb. Divers. Ecosyst. Funct.*, 73–97.

Tuomisto H., Ruokolainen K., & Yli-halla M. (2003) Dispersal , Environment , and Floristic Variation of Western Amazonian Forests. *Science (80-.).*, **299**, 241–245.

Verchot L. V., Davidson E. a., Cattânio J.H., & Ackerman I.L. (2000) Land-Use Change and Biogeochemical Controls of Methane Fluxes in Soils of Eastern Amazonia. *Ecosystems*, **3**, 41–56.

Wardle D.A., Bardgett R.D., Klironomos J.N., Setälä H., Van Der Putten W.H., & Wall D.H. (2004) Ecological linkages between aboveground and belowground biota. *Science (80-.).*, **304**, 1629–1633.

Warton D.I. & Hui F.K. (2011) The arcsine is asinine: the analysis of proportions in ecology. *Ecology*, **92**, 3–10.

Whitaker R.J., Grogan D.W., & Taylor J.W. (2003) Geographic Barriers Isolate Endemic Populations of Hyperthermophilic Archaea. *Science (80-.).*, **301**, 976–978.

White L.J.T. (2001) The African Rain Forest: Climate and vegetation. *African Rain Forest Ecology and Conservation* pp. 3–29. Yale University Press, New Haven and London.

Wickham H. (2009) *ggplot2: elegant graphics for data analysis.* Springer Science & Business Media,

Wilke A., Glass E., Bischof J., Braithwaite D., Gerlach W., Harrison T., Keegan K., Paczian T., Trimble W., & Meyer F. (2015) MG-RAST Manual for version 3.6, revision 2. 20.

Wilke A., Harrison T., Wilkening J., Field D., Glass E.M., Kyrpides N., Mavrommatis K., & Meyer F. (2012) The M5nr: a novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. *BMC Bioinformatics*, **13**, 1–5.

Wilkie D.S. & Laporte N. (2001) Forest area and deforestation in Central Africa: Current knowledge and future directions. *African Rain Forest Ecology and Conservation* pp. 119–139. Yale University Press, New Haven and London.

Woodcock S., Curtis T.P., Head I.M., Lunn M., & Sloan W.T. (2006) Taxa-area relationships for microbes: The unsampled and the unseen. *Ecol. Lett.*, **9**, 805–812.

Zhang Y., Zhao Z., Dai M., Jiao N., & Herndl G.J. (2014) Drivers shaping the diversity and biogeography of total and active bacterial communities in the South China Sea. *Mol. Ecol.*, **23**, 2260–74.

Zhou J., Kang S., Schadt C.W., & Garten C.T. (2008) Spatial scaling of functional gene diversity across various microbial taxa. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 7768–7773.

Zinger L., Boetius A., & Ramette A. (2014) Bacterial taxa-area and distance-decay relationships in marine environments. *Mol. Ecol.*, .