**University of São Paulo
“Luiz de Queiroz” College of Agriculture
Center for Nuclear Energy in Agriculture**

**Comparative analyses of microbial phylogenetic and functional processes
following long-term land-use change**

Dennis Góss de Souza

Thesis presented to obtain the degree of Doctor in
Science. Area: Applied Ecology

**Piracicaba
2015**

Dennis Góss de Souza
Agronomist

**Comparative analyses of microbial phylogenetic and functional processes following
long-term land-use change**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Prof. Dr. **SIU MUI TSAI**

External Supervisor:
Prof. Ph.D. **JORGE LUIZ MAZZA RODRIGUES**

Thesis presented to obtain the degree of Doctor in
Science. Area: Applied Ecology

Piracicaba
2015

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Souza, Dennis Góss de

Comparative analyses of microbial phylogenetic and functional processes following long-term land-use change / Dennis Góss de Souza. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2015.

106 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz". Centro de Energia Nuclear na Agricultura.

1. Amazônia 2. Diversidade microbiana 3. Funções no solo 4. Mata Atlântica
5. Metagenômica I. Título

CDD 631.46
S731c

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

I dedicate this thesis, to those who did not
have the same opportunities I had.

ACKNOWLEDGMENTS

Thank You God for my life and opportunities along my journey.

Thank you my parents for raising and supporting me for the entire life.

Thanks to Professor Siu Mui Tsai, for sharing her knowledge and wisdom along these years, supporting my colleagues and I, no matter where, when or which situation we have been through.

Thanks to Professor Jorge Rodrigues, for opening his lab's doors to receive me as visiting PhD scholar and support me on the paper's writing. Indeed, I want to thank to LAWR Department, University of California, for the structure and support.

Thanks to Fundação de Amparo à Pesquisa do Estado de São Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico (RPGCC-FAPESP/CNPq # 2008/58114-3, SISBIOTA-CNPq # 563251/2010-7), and Labiomol/CENA, for funding the research project and also my participation in several conferences.

Thanks to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (DS-CAPES PPGIEA/USP, PDSE-CAPES # 99999.003968/2014-06) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (PRONEX-CNPq # 140317/2014-7) for the scholarship grants.

Thanks to all my colleagues of Labiomol - CENA/USP for the good moments we have shared each other. Special thanks to Lucas Mendes, Clovis Borges and Ademir Bigaton for the support and partnership along these years.

Thanks to the technicians José Elias Gomes, Wagner Picinini, Fábio Duarte and Ludmila Campos for all support.

Thanks to the colleagues from Biogeo Lab – CENA/USP, in special Lilian Duarte and Professor Brigitte Feigl.

Thanks to all colleagues from SisBiota-UDESC, for pleasant moments in sampling periods and courses. Special thanks to Professor Dilmar Baretta and Professor José Paulo Souza, for the enormous contribution in the development of this research project.

Thanks to Professor Victoria Ballester and Mara Casarin for the willingness in help me all the time.

To all who somehow contributed to this research project, my gratitude.

*“The illiterate of the 21st century will not be those
who cannot read and write, but those who cannot
learn, unlearn, and relearn”.*

(Alvin Toffler)

CONTENTS

RESUMO	13
ABSTRACT	15
1 INTRODUCTION	17
1.1 Hypothesis	20
1.2 Objectives	21
1.2.1 General.....	21
1.2.2 Specific	21
1.3 Structure of the thesis	21
1.4 Funding.....	21
References	22
2 MICROBIAL ASSEMBLY IN LONG-TERM AGRICULTURE SITES FROM ATLANTIC RAINFOREST BIOME: MAIN PATTERNS AND ECOLOGICAL PROCESSES.....	27
Abstract.....	27
2.1 Introduction	27
2.2 Material and Methods	29
2.2.1 Site description and soil collection.....	29
2.2.2 Soil elemental analyses.....	30
2.2.3 Soil DNA extraction and sequencing of metagenomic libraries	31
2.2.4 Annotation of metagenomic data and analysis	33
2.2.5 Statistical analysis	33
2.3 Results	35
2.3.1 Soil attributes	35
2.3.2 Taxonomic and functional diversities across land uses, geographical locations, and sampling times.....	35
2.3.3 Community structure across land-uses	37
2.3.4 Community dynamics in long-term land-use change	38
2.3.5 Soil functional shifts after long-term land use change	39
2.3.6 Environmental factors associated with taxonomic and functional patterns	41
2.3.7 Correlations among soil microbial taxa, functional categories and environmental factors	42
2.4 Discussion.....	45

2.4.1 A working model for microbial community dynamics in long-term agricultural conversion	48
2.4.2 Linking microbial assembly and community functions	48
2.5 Conclusion	49
References	50
3 LONG-TERM MICROBIAL DYNAMICS AT THE SOYBEAN-RHIZOSPHERE INTERFACE IN AMAZON SOILS	57
Abstract	57
3.1 Introduction	57
3.2 Material and Methods	59
3.2.1 Site description and soil sampling	59
3.2.2 Preparation and conduction of the greenhouse mesocosm experiment	61
3.2.3 Metagenomic DNA extraction and high-throughput sequencing	62
3.2.4 Annotation and data mining of metagenome libraries	63
3.3 Results	64
3.3.1 Sampling site attributes	64
3.3.2 Microbial patterns in the soil-rhizosphere interface along the chronosequence	65
3.3.3 Relationship between soil microbial patterns and environmental factors	71
3.3.4 Three-dimension dynamics of the soil microbial community along the chronosequence	75
3.3.5 Microbial community assembly dynamics along the chronosequence	77
3.4 Discussion	79
3.4.1 Ecological processes governing microbial community assembly	79
3.5 Conclusion	82
References	82
4 GENERAL CONCLUSIONS	87
APPENDICES	89

RESUMO

Análise comparativa de processos filogenéticos e funcionais microbianos seguindo a mudança de uso do solo em longo período

Nos últimos anos, ecologistas microbianos aumentaram drasticamente seus esforços para elucidar a “caixa preta” dos padrões e processos que modulam a diversidade e funcionalidade dos microrganismos do solo, examinando sua diversidade genética (e.g. através de metagenômica) e medindo suas características funcionais. O objetivo dessa tese foi avaliar a interação dos processos ecológicos de dispersão, diversificação, seleção e deriva gênica, sobre (1) as comunidades microbianas do solo, após a conversão da floresta em pastagem ou plantio direto, em longo período e (2) sobre as comunidades microbianas da rizosfera de soja, em sistema de plantio direto, em longo período. O cultivo de pastagens em longo período levou a uma seleção homogeneizante das comunidades microbianas, reduzindo a beta-diversidade, com conseguinte alteração em funções no solo relacionadas ao estresse. O plantio direto em longo período levou a uma menor alteração da diversidade, com manutenção das funções encontradas na floresta. A planta de soja demonstrou poder de seleção homogeneizante, e este aumentou com o tempo. Contudo, as funções selecionadas na rizosfera foram mantidas, indicando resiliência funcional.

Palavras-chave: Amazônia; Diversidade microbiana; Funções no solo; Mata Atlântica; Metagenômica

ABSTRACT

Comparative analyses of microbial phylogenetic and functional processes following long-term land-use change

In the last years, microbial ecologists have dramatically increased their efforts to elucidate the "black box" of patterns and processes that modulate the diversity and functionality of soil microorganisms, examining their genetic diversity (e.g. through metagenomic) and measuring their functional characteristics. The aim of this thesis was to evaluate the interaction of the ecological processes of dispersion, diversification, selection and genetic drift on (1) the soil microbial communities, after conversion of forest to grassland or no-till cropping in long-term and (2) on the microbial communities in the rhizosphere of soybean in long-term no-till system. The cultivation of grassland in long-term led to a homogenizing selection of microbial communities, reducing beta-diversity, with consequent changes in the soil functions related to stress. No-till long-term led to minor changes of diversity, maintaining the functions found in the forest. The soybean plant has shown homogenizing power selection, and this increased with time. However, the functions selected in the rhizosphere were maintained, indicating functional resilience.

Keywords: Amazon; Atlantic rainforest; Metagenomics; Microbial diversity; Soil functions

1 INTRODUCTION

Since the rise of classical microbiology, scientists have applied efforts to unravel the ecological patterns of occurrence, distribution and functions of microorganisms in several ecosystems, including soil. In the last years, the famous Baas-Becking affirmation “Everything is everywhere, but, the environment selects”, has been largely used as central question, in microbial biogeographic studies and also in theoretical niche-based theories (DE WIT; BOUVIER, 2006).

Despite we known the importance of soil microorganisms to maintain the dynamic balance and resilience of ecosystems, little is known about their patterns of assembly and its relationship with the functions resilience due the conversion of natural areas, such as tropical forests, to agriculture. Advances arising from genomics era have enabled microbial ecologists to access the ecological dimension of genetic and functional biodiversity, through genomic sequencing techniques, at scale and depth never seen before (METZKER, 2010).

However, a topic that remains unclear is how to analyze and interpret these patterns of biodiversity generated by millions (or even billions) of taxonomic and functional data, resulting in robust and concise answers about ecological issues, among them: which is/are the effect(s) of conversion of forest to agriculture on microbial ecological patterns? Moreover, how to integrate microbial genetic and functional biodiversity, with the environmental factors that can modulate them? Elucidating these dimensions, through metagenomics community ecology may allow us to gain insights from the microbial “black box” in the agroecosystems, and give answers to Baas-Becking affirmation supporters and opponents.

The genetic and phylogenetic dimensions have been massively analyzed through large scale sequencing of nucleic acids. Recently, metagenomic libraries from soil microbial DNA have been used as template (DANIEL, 2005), in order to evaluate the effects of the conversion of forest into agriculture (MENDES et al., 2014). The functional traits dimension has been assessed by biochemical essays of production/consumption of microbial-mediated greenhouse gases and microbial enzyme activities in soil (MORALES; COSART; HOLBEN, 2010; SAINJU et al., 2012; TABATABAI, 1994; YANG et al., 2006). The environmental dimension is often evaluated by observation or field data collection and analysis, including soil physicochemical, climatic and geographical attributes and their relation with microbial molecular parameters (BELL et al., 2008; DEQUIEDT et al., 2009; DRENOVSKY et al., 2010; LAUBER et al., 2013). A multidimensional approach, linking these dimensions that modulate the distribution and abundance of microorganisms in the ecosystems is obtained by

multivariate pairwise correlations among the parameters evaluated in both dimensions, generating an integrative view.

The rapid development of molecular biology techniques at the end of the twentieth century and their successful application to the study of microbial ecology has changed our view of the assess structure and function of microorganisms. In recent years, advances in the field of molecular microbial ecology, in which are included the NGS techniques (HANDELSMAN; SMALLA, 2003), have revealed a far unknown microbial biodiversity that was not detected previously by classical microbiology.

The advance in sequencing technologies from Sanger to 454 pyrosequencing and Illumina has opened new possibilities in microbial community analysis by making it possible to collect millions of sequences, spanning hundreds of samples. The increase in the number of sequences per run from parallel pyrosequencing technologies such as the Roche 454 GS FLX (5×10^5) to Illumina GAIIx (1×10^8) is of the order of 1,000-fold and greater than the increase in the number of sequences per run from Sanger (1×10^3 through 1×10^4) to 454 (CAPORASO et al., 2011). In addition, the use of barcode strategies allows the analysis of thousands of samples in a single run. With the advance of such technologies the read length has increased, although they are far shorter than the desirable length or the read length obtained from traditional Sanger sequencing (~ 1000 bp) (LUO et al., 2012). The 454 pyrosequencing was the first next-generation sequencing technology available as a commercial product (MARGULIES et al., 2005) and can be considered the cornerstone of the sequencing revolution. The development of the pyrosequencing method allowed an advance of metagenome studies by increasing the number of reads and decreasing costs per sequence, enabling a deep phylogenetic community analysis.

Biogeography and community ecology studies have made possible to investigate the mechanisms leading to microbial diversity generation and maintenance in these ecosystems, such as emergence of new species, extinction, drift, dispersal and ecological interactions that lead to variable or homogenizing selection (DINI-ANDREOTE et al., 2015; LEIBOLD et al., 2004) in several levels of complexity and scale ranges. A framework to investigate microbial patterns is needed, with references to that found to macroorganisms and the establishment of possible exceptions regarding microbial assembly and functional niche occupation. It comprises the knowledge whether microbial assemblies differ among environments and space, besides the effects that modulate this variation. Moreover, a biogeographic multiscale approach can help us to unravel if spatial variation is due to punctual environmental factors,

such as land-use and seasonality (LAUBER et al., 2013) or evolutionary selecting events (MARTINY et al., 2006).

Understanding controls over the distribution of soil microbial communities is a fundamental step toward describing soil ecosystems, understanding their functional capabilities, and predicting their responses to environmental change. However, the complexity of these communities and their interactions with environmental characteristics have made generalizations difficult. Recently, high throughput sequencing technologies have facilitated the investigation of soil bacterial communities at local (ACOSTA-MARTÍNEZ et al., 2008), regional (DEQUIEDT et al., 2009), and global scales (LAUBER et al., 2009).

Microbial groups related to environmental characteristics has been recognized as the most important mechanism controlling soil microbial communities (HANSON et al., 2012) with chemical soil factors identified as a master variable explaining significant portions of the variation in soil microbial diversity and community structure at local (NAVARRETE et al., 2013; ROUSK et al., 2010), regional (JESUS et al., 2009; NAVARRETE et al., 2011; RODRIGUES et al., 2013) and global (CHU et al., 2010; FIERER; JACKSON, 2006) scales. However, while environmental factors have been identified as exerting primary control on soil microbial distribution, on average approximately 50% of the variation in microbial diversity and structure remains unexplained (HANSON et al., 2012). Additionally, very few examinations have been made of how controls on soil microbial communities operate simultaneously at multiple scales to contrast local and regional drivers of microbial diversity and community structure.

The factors that control soil microbial community composition are much debated. It has been suggested that while local scale variation in soil microbial communities can be explained by plant identity, substrate hotspots and soil chemical factors (BEZEMER et al., 2010; ORWIN et al., 2010; THOMSON et al., 2010b), at regional and continental scales, additional factors, such as climate, topography, and soil pH, become more important (ETTEMA; WARDLE, 2002; FIERER et al., 2009; LAUBER et al., 2008). However, pH has been shown to shape soil microbial communities over distances $< 1 \text{ m}^2$ (BAKER et al., 2009), as well as at field and continental scales (LAUBER et al., 2009; ROUSK et al., 2010). Microbial communities have also been shown to be influenced by vegetation type, land-use, soil nutrient status, and soil organic matter quality and quantity at landscape scales (FIERER et al., 2009; JANGID et al., 2008; LAUBER et al., 2008).

The relationship among soil microbial communities and landscape factors, soil factors and plant communities at different spatial scales is relatively lacking, despite their importance

for ecosystem functioning. This lack of understanding of the factors that explain variation in microbial communities at larger spatial scales is surprising given their functional importance in regulating ecosystem processes, such as carbon and nitrogen cycling (BALSER; FIRESTONE, 2005; THOMSON et al., 2010b) and the resistance of nutrient cycles to climate change-related disturbances (DE VRIES et al., 2012). In this sense, just a few studies have simultaneously tested the importance of a range of abiotic factors, including climate and soil properties, and biotic factors, such as vegetation composition, across a wide range of spatial scales. According to (DE VRIES et al., 2012), this represents a major gap in knowledge given the potential for both abiotic and biotic factors to explain variation in soil microbial communities.

In this thesis, we propose the formulation of a framework to evaluate microbial dynamics in agroecosystems, following long-term deforestation. This framework is centered on the conceptual theoretical model proposed by Vellend (2010) for macroecology, in which the processes of diversification, selection, drift and dispersal are found to be involved on community assembly patterns. This model was implemented by Nemergut et al. (2013) and Stegen et al. (2013) for microbial ecology studies. Here, we will synthesize and summarize the prior conceptual and experimental knowledge, in the light of ecological processes, trying to underlie the microbial dynamics in long-term conversion of forests in agriculture. The role of the plant on microbial assembly and ecological processes along time, is also addressed here.

1.1 Hypothesis

The first study hypothesized that microbial community assembly in Atlantic Forest soils respond to land-use change, shifting the abundance of specific taxonomic and functional groups in each management, as well as their correlations with soil environmental factors.

The second study hypothesized that communities from bulk soil and rhizosphere have different mechanisms of assembly and different weights of ecological processes governing these communities along a no-till chronosequence, in Amazon soils.

1.2 Objectives

1.2.1 General

The aim of this thesis was to assess via metagenomics approach, in two different biomes, the role of the processes of selection, drift, dispersal and diversification, in long-term conversion of forest soils in agriculture.

1.2.2 Specific

The aims of the first study were to evaluate (i) microbial patterns alterations imposed by long-term neutral and niche-based model assemblies, and (ii) to gain insights into the relationship between community assembly and ecosystem functions.

The aims of the second study were (i) to evaluate the microbial community dynamics in long-term no-till cropping and the power of selection of soybean along this chronosequence, using agriculture sites arisen by the deforestation of the Amazon Rainforest, and, (ii) to generate a descriptive framework as an outcome of the microbial dynamics along soil fractions and time.

1.3 Structure of the thesis

The present thesis comprises a contextual introduction followed by two studies structured in scientific manuscript format.

1.4 Funding

This thesis was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico (RPGCC-FAPESP/CNPq # 2008/58114-3. D. G. Souza received scholarship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (PRONEX-CNPq # 140317/2014-7) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (DS-CAPES PPGIEA/USP, PDSE-CAPES # 99999.003968/2014-06).

References

- ACOSTA-MARTÍNEZ, V.; DOWD, S.; SUN, Y.; ALLEN, V. Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. **Soil Biology and Biochemistry**, New York, v. 40, p. 2762–2770, 2008.
- BAKER, K.L.; LANGENHEDER, S.; NICOL, G.W.; RICKETTS, D.; KILLHAM, K.; CAMPBELL, C.D.; PROSSER, J.I. Environmental and spatial characterisation of bacterial community composition in soil to inform sampling strategies. **Soil Biology and Biochemistry**, New York, v. 41, p. 2292–2298, 2009.
- BALSER, T.C.; FIRESTONE, M.K. Linking microbial community composition and soil processes in a California annual grassland and mixed-conifer forest. **Biogeochemistry**, New York, v. 73, p. 395–415, 2005.
- BELL, C.; MCINTYRE, N.; COX, S.; TISSUE, D.; ZAK, J. Soil microbial responses to temporal variations of moisture and temperature in a chihuahuan desert grassland. **Microbial Ecology**, New York, v. 56, n. 1, p. 153–167, 2008.
- BEZEMER, T.M.; FOUNTAIN, M.T.; BAREA, J.M.; CHRISTENSEN, S.; DEKKER, S.C.; DUYTS, H.; VAN HAL, R.; HARVEY, J.A.; HEDLUND, K.; MARAUN, M.; MIKOLA, J.; MLADENOV, A.G.; ROBIN, C.; DE RUITER, P.C.; SCHEU, S.; SETÄLÄ, H.M.; SMILAUER, P.; VAN DER PUTTEN, W.H. Divergent composition but similar function of soil food webs of individual plants: plant species and community effects. **Ecology**, Ithaca, v. 91, n. 10, p. 3027–3036, 2010.
- CAPORASO, J.G.; LAUBER, C.L.; WALTERS, W.A.; BERG-LYONS, D.; LOZUPONE, C.A.; TURNBAUGH, P.J.; FIERER, N.; KNIGHT, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 108, p. 4516–4522, 2011.
- CHU, H.; FIERER, N.; LAUBER, C.L.; CAPORASO, J.G.; KNIGHT, R.; GROGAN, P. Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. **Environmental Microbiology**, Oxford, v. 12, n. 11, p. 2998–3006, 2010.
- DANIEL, R. The metagenomics of soil. **Nature Reviews Microbiology**, London, v. 3, n. 6, p. 470–478, 2005.
- DEQUIEDT, S.; THIOULOUSE, J.; JOLIVET, C.; SABY, N.P.A.; LELIEVRE, M.; MARON, P.-A.; MARTIN, M.P.; PRÉVOST-BOURÉ, N.C.; TOUTAIN, B.; ARROUAYS, D.; LEMANCEAU, P.; RANJARD, L. Biogeographical patterns of soil bacterial communities. **Environmental Microbiology Reports**, Malden, v. 1, n. 4, p. 251–255, 2009.
- DE VRIES, F.T.; LIIRI, M.E.; BJØRNLUND, L.; BOWKER, M.A.; CHRISTENSEN, S.; SETÄLÄ, H.M.; BARDGETT, R.D. Land use alters the resistance and resilience of soil food webs to drought. **Nature Climate Change**, London, v. 2, n. 4, p. 276–280, 2012.
- DE WIT, R.; BOUVIER, T. Everything is everywhere, but, the environment selects; what did Baas Becking and Beijerinck really say? **Environmental Microbiology**, Oxford, v. 8, n. 4, p. 755–758, 2006.

DINI-ANDREOTE, F.; STEGEN, J.C.; VAN ELSAS, J.D.; SALLES, J.F. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. **Proceedings of the National Academy of Sciences**, Washington, v. 112, n. 11, p. 1326–1332, 2015.

DRENOVSKY, R.E.; STEENWERTH, K.L.; JACKSON, L.E.; SCOW, K.M. Land use and climatic factors structure regional patterns in soil microbial communities. **Global Ecology and Biogeography**, Malden, v. 19, n. 1, p. 27–39, 2010.

ETTEMA, C.H.; WARDLE, D.A. Spatial soil ecology. **Trends in Ecology & Evolution**, Cambridge, v. 17, n. 4, p. 177–183, 2002.

FIERER, N.; JACKSON, R.B. The diversity and biogeography of soil bacterial communities. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 103, n. 3, p. 626–631, 2006.

FIERER, N.; STRICKLAND, M.S.; LIPTZIN, D.; BRADFORD, M.A.; CLEVELAND, C.C. Global patterns in belowground communities. **Ecology Letters**, Oxford, v. 12, p. 1238–1249, 2009.

HANDELSMAN, J.; SMALLA, K. Techniques: conversations with silent majority. **Current Opinion in Microbiology**, London, v. 6, n. 3, p. 271–273, 2003.

HANSON, C.A.; FUHRMAN, J.A.; HORNER-DEVINE, M.C.; MARTINY, J.B.H. Beyond biogeographic patterns: processes shaping the microbial landscape. **Nature Reviews Microbiology**, London, v. 10, n. 7, p. 497–506, 2012.

JANGID, K.; WILLIAMS, M.A.; FRANZLUEBBERS, A.J.; SANDERLIN, J.S.; REEVES, J.H.; JENKINS, M.B.; ENDALE, D.M.; COLEMAN, D.C.; WHITMAN, W.B. Relative impacts of land-use, management intensity and fertilization upon soil microbial community structure in agricultural systems. **Soil Biology and Biochemistry**, New York, v. 40, p. 2843–2853, 2008.

JESUS, E.C.; MARSH, T.L.; TIEDJE, J.M.; MOREIRA, F.M. Changes in land use alter the structure of bacterial communities in Western Amazon soils. **The ISME Journal**, London, v. 3, n. 9, p. 1004–1011, 2009.

LAUBER, C.L.; HAMADY, M.; KNIGHT, R.; FIERER, N. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. **Applied and Environmental Microbiology**, Washington, v. 75, n. 15, p. 5111–5120, 2009.

LAUBER, C.L.; STRICKLAND, M.S.; BRADFORD, M.A.; FIERER, N. The influence of soil properties on the structure of bacterial and fungal communities across land-use types. **Soil Biology and Biochemistry**, New York, v. 40, p. 2407–2415, 2008.

LAUBER, C.L.; RAMIREZ, K.S.; AANDERUD, Z.; LENNON, J.; FIERER, N. Temporal variability in soil microbial communities across land-use types. **The ISME Journal**, London, v. 7, n. 8, p. 1641–1650, 2013.

LEIBOLD, M.A.; HOLYOAK, M.; MOUQUET, N.; AMARASEKARE, P.; CHASE, J.M.; HOOPES, M.F.; HOLT, R.D.; SHURIN, J.B.; LAW, R.; TILMAN, D.; LOREAU, M.; GONZALEZ, A. The metacommunity concept: a framework for multi-scale community ecology. **Ecology Letters**, Oxford, v. 7, n. 7, p. 601–613, 2004.

LUO, C.; TSEMENTZI, D.; KYRPIDES, N.; READ, T.; KONSTANTINIDIS, K.T. Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. **PloS One**, New York, v. 7, n. 2, p. 1–12, 2012.

MARGULIES, M.; EGHOLM, M.; ALTMAN, W.E.; ATTIYA, S.; BADER, J.S.; BEMBEN, L.A.; BERKA, J.; BRAVERMAN, M.S.; CHEN, Y.-J.; CHEN, Z.; DEWELL, S.B.; DU, L.; FIERRO, J.M.; GOMES, X.V.; GODWIN, B.C.; HE, W.; HELGESEN, S.; HO, C.H.; IRZYK, G.P.; JANDO, S.C.; ALENQUER, M.L.I.; JARVIE, T.P.; JIRAGE, K.B.; KIM, J.-B.; KNIGHT, J.R.; LANZA, J.R.; LEAMON, J.H.; LEFKOWITZ, S.M.; LEI, M.; LI, J.; LOHMAN, K.L.; LU, H.; MAKHIJANI, V.B.; MCDADE, K.E.; MCKENNA, M.P.; MYERS, E.W.; NICKERSON, E.; NOBILE, J.R.; PLANT, R.; PUC, B.P.; RONAN, M.T.; ROTH, G.T.; SARKIS, G.J.; SIMONS, J.F.; SIMPSON, J.W.; SRINIVASAN, M.; TARTARO, K.R.; TOMASZ, A.; VOGT, K.A.; VOLKMER, G.A.; WANG, S.H.; WANG, Y.; WEINER, M.P.; YU, P.; BEGLEY, R.F.; ROTHBERG, J.M. Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, London, v. 437, n. 7057, p. 376–380, 2005.

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. Microbial biogeography: putting microorganisms on the map. **Nature Reviews Microbiology**, London, v. 4, n. 2, p. 102–112, 2006.

MENDES, L.W.; KURAMAE, E.E.; NAVARRETE, A.A.; VAN VEEN, J.A.; TSAI, S.M. Taxonomical and functional microbial community selection in soybean rhizosphere. **The ISME Journal**, v. 8, p. 1577–1587, 2014.

METZKER, M.L. Sequencing technologies - the next generation. **Nature Reviews Genetics**, London, v. 11, n. 1, p. 31–46, 2010.

MORALES, S.E.; COSART, T.; HOLBEN, W.E. Bacterial gene abundances as indicators of greenhouse gas emission in soils. **The ISME Journal**, London v. 4, n. 6, p. 799–808, 2010.

NAVARRETE, A.A.; KURAMAE, E.E.; DE HOLLANDER, M.; PIJL, A.S.; VAN VEEN, J.A.; TSAI, S.M. Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. **FEMS Microbiology Ecology**, Amsterdam, v. 83, n. 3, p. 607–621, 2013.

NAVARRETE, A.A.; TAKETANI, R.G.; MENDES, W.; CANNAVAN, F.D.S.; MOREIRA, F.M.S.; TSAI, S.M. Land-use systems affect archaeal community structure and functional diversity in western amazon soils. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 35, n. 1, p. 1527–1540, 2011.

NEMERGUT, D.R.; SCHMIDT, S.K.; FUKAMI, T.; O'NEILL, S.P.; BILINSKI, T.M.; STANISH, L.F.; KNELMAN, J.E.; DARCY, J.L.; LYNCH, R.C.; WICKEY, P.; FERRENBURG, S. Patterns and processes of microbial community assembly. **Microbiology and Molecular Biology Reviews**, Washington, v. 77, n. 3, p. 342–356, 2013.

ORWIN, K.H.; BUCKLAND, S.M.; JOHNSON, D.; TURNER, B.L.; SMART, S.; OAKLEY, S.; BARDGETT, R.D. Linkages of plant traits to soil properties and the functioning of temperate grassland. **Journal of Ecology**, London, v. 98, p. 1074–1083, 2010.

RODRIGUES, J.L.M.; PELLIZARI, V.H.; MUELLER, R.; BAEK, K.; JESUS, E.D.C.; PAULA, F.S.; MIRZA, B.; HAMAOU, G.S.; TSAI, S.M.; FEIGL, B.; TIEDJE, J.M.; BOHANNAN, B.J.M.; NÜSSLEIN, K. Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 110, n. 3, p. 988–993, 2013.

ROUSK, J.; BAATH, E.; BROOKES, P.C.; LAUBER, C.L.; LOZUPONE, C.; CAPORASO, J.G.; KNIGHT, R.; FIERER, N. Soil bacterial and fungal communities across a pH gradient in an arable soil. **The ISME Journal**, London, v. 4, p. 1340–1351, 2010.

SAINJU, U.M.; STEVENS, W.B.; CAESAR-TONTHAT, T.; LIEBIG, M.A. Soil greenhouse gas emissions affected by irrigation, tillage, crop rotation, and nitrogen fertilization. **Journal of Environmental Quality**, Madison, v. 41, n. 6, p. 1774–1786, 2012.

STEGEN, J.C.; LIN, X.; FREDRICKSON, J.K.; CHEN, X.; KENNEDY, D.W.; MURRAY, C.J.; ROCKHOLD, M.L.; KONOPKA, A. Quantifying community assembly processes and identifying features that impose them. **The ISME Journal**, London, v. 7, n. 11, p. 2069–2079, 2013.

TABATABAI, M.A. Soil enzymes. In: WEAVER, R.W.; ANGLE, J.S.; BOTTOMLEY, P.S. (Ed.). **Methods of soil analysis**. 2nd ed. Madison: SSSA Book Series, 1994. pt. 2: p. 775–833.

THOMSON, B.C.; OSTLE, N.; MCNAMARA, N.; BAILEY, M.J.; WHITELEY, A.S.; GRIFFITHS, R.I. Vegetation affects the relative abundances of dominant soil bacterial taxa and soil respiration rates in an upland grassland soil. **Microbial Ecology**, New York, v. 59, p. 335–343, 2010.

VELLEND, M. Conceptual synthesis in community ecology. **The Quarterly Review of Biology**, Chicago, v. 85, n. 2, p. 183–206, 2010.

YANG, Z.; LIU, S.-Q.; ZHENG, D.; FENG, S. Effects of cadmium, zinc and lead on soil enzyme activities. **Journal of Environmental Sciences**, Cambridge, v. 18, n. 6, p. 1135–1141, 2006.

2 MICROBIAL ASSEMBLY IN LONG-TERM AGRICULTURE SITES FROM ATLANTIC RAINFOREST BIOME: MAIN PATTERNS AND ECOLOGICAL PROCESSES

Abstract

Microbial ecologists have emphasized the importance of community patterns in the study of assembly mechanisms. We reason that this approach has hampered our understanding of microbial diversity and ability to test novel hypotheses in community ecology. Here, we evaluated the relative importance of the ecological processes of diversification, selection, drift and dispersal on community assembly in soils under different land uses. We used a high throughput metagenomics approach to investigate microorganism diversity and composition and functional alterations associated with forest-to-agricultural conversion. Although no significant alterations in alpha diversity were observed among different land uses, the beta diversity in grassland was lower than those observed in forest and no-till soils. At the functional level, grassland microbial communities showed significant changes for 12 categories in comparison to forests, while long-term conversion of forest to no-till system resulted in modifications of four categories. Grassland communities showed assembly that predominantly followed the niche model, i.e. high selection with moderate dispersion, leading to biotic homogenization. Both no-till and forest soil communities were found to have assembly that predominantly followed a neutral model, i.e. low rates of dispersal and weak selection, resulting in maintenance of higher beta diversity relative to grasslands. The co-occurrence of taxa, functions, and environmental parameters, visualized in networks, revealed a non-modular structure for grassland communities, due the high influence of environmental variables. Taken together, our results indicate that the patterns of assembly and their governing processes are dependent on the land use employed after deforestation, with consequences to taxa turnover and functional maintenance.

Keywords: Assembly; Community ecology; Neutral and niche model; Specialization

2.1 Introduction

Ecologists have long established the mechanisms ruling community assembly, but quantifying the importance of the different processes for microorganisms remains a fundamental research challenge. Most ecological models based on selection were developed for interactions occurring among a few and/or discrete species (LEIBOLD et al., 2004; MACARTHUR; WILSON, 1963; WHITTAKER, 1962), but microbial communities are known to be more complex and species-rich than those model/simpler communities (FIERER; LENNON, 2011; TRINGE et al., 2005). One of the challenges in working with communities containing a large number of species is that assembly configurations are nearly limitless. Thus, testing community ecology theories to microorganisms can only be achieved by

developing a framework that takes into consideration the particularities of complex communities.

The current theory proposes two contrasting models of assembly: (i) a niche-based (deterministic) model that considers abiotic and biotic factors as primary drivers of taxa differences (CARROLL; CARDINALE; NISBET, 2011) and (ii) a neutral (stochastic) model that is focused on probabilistic and random events to explain the structuring of communities (HUBBELL, 2001). Although only a few microbiological studies have examined the above models directly, a consensus has been built that considers microbial community assembly as the interplay between stochastic and deterministic processes (DINI-ANDREOTE et al., 2015; FERRENBURG et al., 2013; MENDES et al., 2014; OFITERU et al., 2010; STEGEN et al., 2012). Indeed, experimental support has been provided for the importance of niche-based and stochasticity factors on the assembly of soil microbial communities (BARBERÁN et al., 2012; OFITERU et al., 2010; WANG et al., 2013). It remains, however, of limited understanding how these processes take place in soils under different land uses and the consequences of forest-to-agriculture conversion to community assembly.

Change in land use due to agriculture has been identified as the most important driver of biodiversity losses (CHAPIN et al., 2000; SALA et al., 2000). This is particularly severe in tropical systems with rising societal demands for food, fiber, and energy resources (FOLEY et al., 2005, 2011; THOMSON et al., 2010a) and an increasing number of studies have documented biodiversity losses for animals (CARDINALE et al., 2012; MCCONKEY; FARRILL, 2015) and plants (FEELEY; SILMAN, 2009; PEREIRA et al., 2010; WARDLE et al., 2011). Recently, several studies have shown that microbial diversity losses also occur in tropical forests converted to grassland (MENDES et al., 2015; MUELLER et al., 2014; NAVARRETE et al., 2011; RODRIGUES et al., 2013) or no-till systems (CARBONETTO et al., 2014; FIGUEROLA et al., 2012; LIENHARD et al., 2013), but the consequences of land use change to the ecological processes ruling assembly remain unknown.

In this study, the relative contributions of the ecological processes of selection, drift, dispersal, and diversification to community assembly—previously reviewed by Vellend (2010) and Nemergut et al. (2013)—were examined in agricultural soils after long-term conversion of the Brazilian Atlantic Rainforest, a biome with the highest number of endemic species among all terrestrial biomes (MYERS et al., 2000; RIBEIRO et al., 2009). We tested the hypothesis that ecological processes differ across long-term land uses. In a corollary hypothesis, we tested whether variation in microbial function is dependent on assembly. By combining the conceptual framework in community ecology proposed by Vellend (2010) with

implementations conceived by Stegen et al. (2013) and high throughput metagenomics, we aimed: (i) to assess the processes of selection, drift, dispersal, and diversification in forest, grassland, and no-till soils, (ii) to evaluate microbial functional alterations imposed by neutral and niche-based model assemblies, and (iii) to gain insights into the relationship between community assembly and ecosystem functions.

2.2 Material and Methods

2.2.1 Site description and soil collection

Soil samples were collected within the Brazilian Atlantic Rainforest Biome composed of remnants of the original forest cover, long-term grassland and no-till cropping sites, with a minimum of 10 years of conversion (Figure 2.1). Forest sites comprised a natural transition between mixed ombrophilous forest and semi-deciduous forest (Supplementary Information). Grassland sites were characterized as perennial with predominance of *Axonopus affinis* (Poaceae). No-till cropping systems were characterized by successive cultivation of wheat in the winter, followed by maize and soybean in the summer. The site selection criteria were based in previous exploratory visits and history of land use and management acquired through interviews with farmers and regional experts.

Figure 2.1 - Soil sampling sites location on the Atlantic Rainforest Biome, Santa Catarina State, Brazil. (a) Nine samples from each site were collected, in two sampling seasons, at 0-10 cm profile. (b) Samples were collected in a square geogrid scheme, representing an area of 3,600 m² of sampling in each sampling site. Samples were used for both metagenomics and environmental analyses. Altitude, latitude and longitude were geo-referenced at the central point of each geogrid site

Samples were collected in July 2011 and January 2012, comprising winter and summer, respectively, in the municipalities of Chapecó (27°2'S; 52°41'W) and Xanxerê (26°50'S; 52°27'W), State of Santa Catarina, Brazil (Figure 2.1) Non-deformed soil samples from the 0-10 cm profile were taken with sterile PVC tubes (5 cm diameter X 10 cm depth), yielding at least 500 g of soil each. A total of 36 soil samples were collected (2 municipalities X 2 sampling periods X 3 sites X 3 samples per site). Each sample was composed by 3 subsamples in a geogrid system, equidistantly every 30 meters from each other (Figure 2.1). The litter layer was removed prior sampling collection. Samples for microbial analysis were kept on dry ice and transported to Cell and Molecular Biology Laboratory, Center for Nuclear Energy in Agriculture (Piracicaba, Brazil), while samples for physical and chemical analyses were maintained at 4°C and transported to the Soil Physicochemistry and Fertility Analysis Laboratory, Santa Catarina State University (Lages, Brazil).

The six study sites are located in the Brazilian Atlantic Rainforest Biome, with humid subtropical climate without drought period (Cfa), according to Köppen-Geiger classification. Annual precipitation average is about 1,400 mm. Rhodic Nitisol is the predominant soil order of the sampling sites. The main plants in the forest canopy were *Apuleia leiocarpa*, *Balfourodendron riedelianum*, *Cabralea glaberrima*, *Cedrela fissilis*, *Cordia trichotoma*, *Diatenopterix sorbifolia*, *Enterolobium contortisiliquum*, *Lonchocarpus leucanthus*, *Parapiptadenia rigida*, *Patagonula americana* and *Peltophorum dubium*. Forest areas were deforested via timber slash-and-burn in the late 1980s. Then, these areas were converted into two distinct land-uses: (I) Grassland - livestock in extensive grazing, with predominance of *Brachiaria* spp. (Poaceae) in the landscape, and (II) No-till cropping - soybean and sporadically maize in the summer, and wheat or ryegrass/oat in the winter. Since the soils from the region are predominantly acids (pH ~ 4), after conversion, no-till cropping sites have often received liming to increase the pH close to 6, every 4th or 5th year, fertilizers to increase nutrient availability, every year, and pesticides to control weeds, plagues and pathogens, every year. Grassland sites have received these same management practices, occasionally.

2.2.2 Soil elemental analyses

Soil chemical and physical properties were determined for each sample based on 500 g of soil. A total of 35 soil parameters were analyzed. Most of them (29) were analyzed at the Soil Physicochemistry and Fertility Analysis Laboratory, Santa Catarina State University,

Lages, Brazil, performed at the Laboratory of Soil Analysis at “Luiz de Queiroz” College of Agriculture (ESALQ/USP, Piracicaba, Brazil), following routine methodology. Soil pH was measured in a 1:2.5 soil/water suspension. Exchangeable Al, Ca, and Mg were extracted with KCl 1 M. Calcium and Mg were determined by atomic absorption spectrometry and Al by acid-base titration. Phosphorus and K were extracted by ion-exchange resin. Potential acidity (H+Al) was estimated by an equation based on the pH determined in SMP buffer solution (pH SMP). Available micronutrients (Fe, Mn, Zn, and Cu) were extracted by Mehlich-1 and determined by atomic absorption spectrometry. Boron was extracted with hot water and determined by spectrophotometry with azomethine-H at 420 nm. Some of the results allowed the calculation of other parameters such as exchangeable bases (SB), the sum of Ca, Mg, and K; cation exchange capacity (CEC), the sum of Ca, Mg, K, Al, and H; base saturation (V%), the percentage relation between SB and CEC; and Al saturation (m%), the percentage relation between exchangeable Al and CEC. Total C, H, N and S were extracted and determined by combustion catalytic oxidation method in a TOC Analyzer. Soil texture was determined using Bouyoucos densimeter, after shaking the soil vigorously with NaOH 1 M as dispersant. Penetration resistance was measured by a penetrometer. Total porosity, macro and microporosity were measured. The gravimetric moisture was obtained in percentage, through the difference between the weight of the sample at the moment of sampling and its dry weight, after 48 hours in an incubator at 105 Celsius degrees. Soil density was measured by Kopecky's ring method. Biopores were calculated by measuring the pores continuity. Total porosity was calculated through saturation method. Microporosity was obtained by tension table method. Macroporosity was calculated by difference, deducting the microporosity from the total porosity.

2.2.3 Soil DNA extraction and sequencing of metagenomic libraries

Total DNA was extracted from soil samples (0.25g) using the PowerLyzer PowerSoil DNA Isolation Kit (Mo Bio Laboratories, Carlsbad, USA), following instructions of the manufacturer. Soil DNA was visualized in 1% Tris-saline buffer (TBS) with sodium boric acid (BRODY; KERN, 2004) agarose gel electrophoresis and quantified with the Qubit fluorometer (Thermo Fischer Scientific, Waltham, USA).

Metagenomic libraries were constructed with the Nextera DNA Sample Preparation Kit (Illumina Inc., San Diego, USA), as specified by the manufacturer. Barcoding preparation reaction was performed with 50ng of soil DNA in 20 µl volume in a 96 well-plate with

addition of 25 μ l of Tagment DNA buffer (TD) and 5 μ l of TD1 enzyme. Reactions were carried out in an Applied Biosystems GeneAmp PCR System 9700 (Thermo Fischer Scientific, Waltham, USA) under the following settings: 55°C for 5 min and enzyme inactivation at 10°C for 5 min. Products were purified with the GFX PCR DNA and Gel Band Purification Kit (GE Healthcare, Little Chalfont, UK) and resuspended in 25 μ l of water.

The insertion of barcoding indexes was performed with the Nextera Index Kit (Illumina Inc., San Diego, USA). A combination of 4 x 6 indexes was used, totaling 24 samples per sequencing run. For the each PCR reaction, 20 μ l of purified DNA, 5 μ l of index 1 + 5 μ l of index 2, 15 μ l of Nextera PCR Master Mix (NPM) and 5 μ l of PCR Primer Mix were mixed to a final volume of 50 μ l. The PCR reaction was performed in the Applied Biosystems GeneAmp PCR System 9700 (Thermo Fischer Scientific, Waltham, USA) and included a pre-extension at 72°C for 3 min, pre-denaturation at 98°C for 30 s, followed by 5 cycles of denaturation at 98°C for 10 s, annealing at 63°C for 30 s, extension at 72°C for 3 min, and a final inactivation at 10°C for 3 min. Products were purified with Agencourt AMPure XP reagents (Beckman Coulter, Brea, USA), on a magnetic rack, as manufacturer's instructions. DNA was dried at 37°C for 3 min in a Biological Oxygen Demand (BOD) incubator. Sample were resuspended in 32.5 μ l of RSB Buffer (supplied by Nextera Index Kit), incubated for 2 min and transferred to a new 96-well plate. DNA samples were quantified with the Qubit fluorometer (Thermo Fischer Scientific, Waltham, USA) and equimolar concentrations (80 ng) of metagenomic libraries pooled together for sequencing.

Prior to sequencing, pooled samples were quantified with the KAPA SYBR FAST Universal qPCR Kit (Kappa Biosystems, Woburn, USA). A standard curve with six known DNA concentrations, ranging from 0.0002 pM to 20 pM, was used with three replicates for each point. Each PCR reaction contained 4 μ l of Master Mix, 1 μ l of each primer provided by the manufacturer, 2 μ l of sterile water and 2 μ l of purified DNA diluted 100, 1000 and 2000 times. PCR was performed in the StepOne Plus thermocycler (Thermo Fischer Scientific, Waltham, USA) and consisted of a pre-denaturation at 95°C for 5 min, followed by 35 cycles of: denaturation at 95°C for 30 s and annealing at 60°C for 45 s.

The final step of sequencing was performed with MiSeq Reagent Kit V2 (Illumina Inc., San Diego, USA). Ten μ l of NaOH (0.2 M) were added to 10 μ l of pooled metagenomic libraries (2nM). The solution was mixed briefly in a vortex and incubated at room temperature for 5 min. followed by addition of 980 μ l of HT1 buffer. A total of 600 μ l of the mixture (17 pM of DNA) was loaded in the sequencing cartridge and inserted into the MiSeq Desktop Sequencer (Illumina Inc., San Diego, USA) at the Center for Nuclear Energy in Agriculture,

University of Sao Paulo, Brazil. The sequencing consisted of a paired-end reads (2 x 250bp) 500-cycle/39-hour run.

2.2.4 Annotation of metagenomic data and analysis

Unassembled fastq DNA sequences were annotated with the Metagenomics Rapid Annotation (MG-RAST) pipeline version 3.3.9 (MEYER et al., 2008). Taxonomic and functional profiles were generated using the normalized abundance of sequences matches to the M5nr (WILKE et al., 2012) and SEED (OVERBEEK et al., 2013) databases, respectively. BlastX searches with a minimum alignment length of 50 bp and an *E*-value cut-off of 1×10^{-5} was used for generating data matrices. Tables of frequency for hits associated with each individual taxon (taxonomy) or subsystem (function) for different metagenomes were normalized to remove bias in read lengths and sequencing efforts. Sequences have been deposited in MG-RAST as project ID = 5426.

Shannon's alpha-diversity and Whittaker's global beta diversity were calculated based on taxonomic relative abundance matrix at genus level and functional relative abundance matrix at subsystem Level 2, using the PAST software, v. 3.0 (HAMMER; HARPER; RYAN, 2001). Nonmetric multidimensional scaling (NMDS) plots were used to visualize the structure among samples, using the above described taxonomic and functional abundance matrices. NMDS rotations of PCA were performed on Canoco software v. 5.0 (LEPŠ; ŠMILAUER, 2005).

2.2.5 Statistical analysis

All statistical analyses were performed using the following design: three land uses, two seasons and two true replicates (municipalities). Data were presented together when the effect of season and/or geography was negligible. Abundance data from taxa (M5nr database) and functional subsystems (SEED database) of each soil metagenome were downloaded from MG-RAST and used to generate taxonomic and functional matrices as input for the Statistical Analysis of Metagenomic Profiles (STAMP) software, v. 3.0 (PARKS; BEIKO, 2010). The *q*-values were calculated using two-sided Welch's *t*-test (WELCH, 1934), while confidence intervals were calculated using the Newcombe–Wilson method (NEWCOMBE, 1998) with corrections being performed using the Benjamini–Hochberg false discovery rate (BENJAMINI; HOCHBERG, 1995).

Distance-based redundancy analysis (db-RDA) of Bray-Curtis dissimilarity matrices, with forward selection was performed, in order to identify the main environmental drivers of microbial taxonomic assembly and functional structure. In addition, a variation partitioning of redundancy analysis (p-RDA), generated by principal coordinates analysis of neighbor matrices (PCNM) with forward selection was performed to evaluate the possible single and joint effects of environmental variables, space and time on the variation of taxonomic and functional profiles of microbial communities in each land use. Latitude and longitude were used as constraining spatial coordinates. Both analyses were performed using the Canoco software, v.5.0.

An all-possible rank pairwise Spearman's correlation with Benjamini-Hochberg false discovery rate correction was used to determine correlations within taxonomic groups and functions, their interactions, and their connectivity with environmental factors. The analysis was performed on the Vegan package from the R software, v. 3.1.2 (R DEVELOPMENT CORE TEAM, 2008). Correlation patterns across land uses, we visualized with Gephi software (BASTIAN; HEYMANN; JACOMY, 2009), version 0.8.2 using only high Spearman's pairwise correlations, with cut-off at $-0.7 \leq R \leq 0.7$, with P -value ≤ 0.01 for both taxonomy, function and environmental variables. Nodes in each network represent taxa, functions and environmental factors. Edges represent high and significant positive or negative correlations between nodes.

We calculated rank abundance distributions and immigration rates, based on taxonomic genus level matrices, for each sample, from each land use system. The species rank abundance for each metagenomic sample were fitted to the neutral zero-sum model (ZSM) on the TeTame software (JABOT; ETIENNE; CHAVE, 2008), version 1.9. The neutral broken stick model (null model) and three niche-based models (log-normal, pre-emption, and Zipf) were calculated using the function 'radfit' from Vegan package on R software, version 3.1.2. Models were compared based on the Akaike Information Criterion (AIC) — the lower the value, the better the microbial community fitted to a specific model. The dispersal rate was calculated on the TeTame software, through Etienne's formula.

2.3 Results

2.3.1 Soil attributes

We performed a Tukey's Honest Significant Difference (HSD) test for 35 soil physicochemical and 3 geographical parameters measured for forest, grassland and no-till samples taken during winter and summer (See appendix, supplementary Table ST 2.1). The total N, total C, SOM, potential acidity (H+Al), aluminum (Al^{3+}), and Al^{3+} saturation, total porosity, macroporosity, and bioporosity decreased after long term conversion, while pH, PO_4^- , Ca^{+2} , Mg^{+2} , base saturation, soil density, microporosity, and penetration resistance increased ($P < 0.001$). The concentration of K^+ and N-NH_4^+ only increased in the non-till system. When comparisons were performed between sampling periods, N-NH_4^+ , N-NO_3^- , total S concentrations and the C:N ratio were higher for all land uses during the summer ($P < 0.001$).

2.3.2 Taxonomic and functional diversities across land uses, geographical locations, and sampling times

A total of 27 million of sequences were obtained for 36 soil samples using a shotgun metagenomic approach. The taxonomic and functional diversities across land uses were measured with both Shannon's alpha diversity and the Whittaker's global beta diversity indices for two municipalities in winter and summer (Figure 2.2). There were no significant changes in alpha diversity with forest-to-agriculture conversion. Although not significant, alpha diversities were higher when sampling was performed in the summer, in both municipalities for taxonomic data (Figures 2.2a and 2.2b). No differences were found for functional data (Figures 2.2c and 2.2d).

Figure 2.2 - Soil microbial diversity at taxonomic genus level (a and b) and functional subsystems level 2 (c and d), based on shotgun metagenomics of three land uses: forest, grassland and no-till. Two municipalities: Chapecó and Xanxerê. Bars represent Shannon's alpha-diversity and dispersion points represent Whittaker's Global beta-diversity for two sampling seasons: winter (black) and summer (gray). Error bars show standard deviation of triplicate samples

The taxonomic Whittaker's global beta diversity values were lower for grasslands in both municipalities, with no changes between the forest and no-till system (Figures 2.2a and 2.2b). The same patterns were observed for functional beta diversity (Figures 2.2c and 2.2d). There was a shift in Whittaker's global beta diversity between locations and sampling times, but when an analysis for explanatory variables with stepwise forward selection was performed (Supplementary Tables ST2 and ST3), both were found significant only for grassland.

2.3.3 Community structure across land uses

Since grassland land-use showed a lower Whittaker's global beta diversity index in comparison to forest and no-till for both locations, we asked whether community structure displayed similar patterns among different land uses. Non-metric multidimensional scaling plots (NMDS) for taxonomy at genus level and functional level 2 were generated using a Bray-Curtis dissimilarity matrix (Figure 2.3).

Figure 2.3 - Non-metric multidimensional scaling 3D plot generated from a Bray-Curtis dissimilarity matrix of metagenomic profiles based on (a) taxonomic genus level and (b) functional subsystems level 2. Colored circles as follows: forest (green), grassland (red) and no-till (brown)

At taxonomic level, grassland soil samples were grouped together and separated from forest and no-till samples (Figure 2.3a). Resulting coordinates were used to calculate the mean distance between neighbor samples (MDN) through nearest neighbor algorithm and the mean density of coordinates (MDC) through convex hull area estimation. Forest microbial communities showed significant overdispersion on taxa assembly (MDN = 0.048; MDC = 250; $P < 0.001^{***}$), while non-till (MDN = 0.038; MDC = 408; $P = 0.14^{NS}$) and grassland (MDN = 0.027; MDC = 943; $P = 0.06^{NS}$) communities were found to be clustered within each land use (Figure 2a). The functional structure (Figure 2.3b) became more homogeneous after long-term conversion of forest (MDN = 0.046; MDC = 366; $P = 0.005^{**}$) in grassland (MDN = 0.032; MDC = 919; $P = 0.002^{**}$) and, more disperse in no-till (MDN = 0.061; MDC = 228; $P = 0.003^{**}$), different from what was found in taxonomy, being that all land-use presented statistically significant overdispersion on functional categories structure.

2.3.4 Community dynamics in long-term land use change

Sequence reads comprised of 32 phyla with 28 belonging to Bacteria and 4 to Archaea based on a non-redundant assignment of the M5nr database. We calculated the normalized taxa relative abundance based on a 1,000 resampling permutations of the shotgun metagenomes (Figure 2.4). Overall, the most abundant phylum was Proteobacteria with an average of 46.6% of the sequences, followed by the phylum Actinobacteria representing 20.9% of the total. The Acidobacteria and Firmicutes were represented by a total of 9.5 and 4.1%, respectively. Sequences matching the phyla Verrucomicrobia, Planctomycetes, Chloroflexi, Cyanobacteria, and Bacteroidetes ranged from 2 to 3.3%. The same patterns were observed when comparing summer and winter sampling periods (Figure 2.4a).

Figure 2.4 - Relative abundance of soil bacterial and archaeal communities at phylum level based on shotgun metagenomics. Only significantly altered taxa are shown: forest (green bars and circles), grassland (red bars and circles), and no-till (brown bars and circles). Error bars show standard deviation. Panels: (a) Percentages of sequence reads are shown for two sampling seasons. Differences in mean proportions are shown: (b) comparison between forest and grassland and (c) comparison between forest and no-till. A Welch's t-test with corrected q-values calculated using the Benjamini-Hochberg false discovery rate was performed ($q\text{-value} \leq 0.05$)

When a Welch's t-test corrected with the Benjamini-Hochberg false discovery rate (FDR) was used, long-term conversion of forest to grassland and forest to no-till showed significant changes for 11 and 8 phyla, respectively (Figures 2.4b and 2.4c, q-values are shown in the panels). The relative abundances of the Planctomycetes, Verrucomicrobia, Fibrobacteres, Synergistetes and Euryarchaeota increased with conversion to pasture (Figure 2.4b). The largest proportional decrease in response to conversion was by the phylum Proteobacteria, followed by Gemmatimonadetes, Nitrospirae, and Thaumarchaeota. The long-term no-till conversion resulted in the increase (combined total of 3%) in relative abundances of several phyla, Bacteroidetes, Chloroflexi, Gemmatimonadetes, Nitrospirae, Planctomycetes, Dictyoglomi, and Poribacteria, with a corresponding decrease (2.7%) for the Acidobacteria (Figure 2.4c).

2.3.5 Soil functional shifts after long-term land use change

In order to verify whether the taxonomic differences observed with forest conversion were reflected at functional level, metagenomic libraries were classified according to the SEED subsystems database. Out of 48 possible categories, soil metagenomes were assigned to 28 different functional subsystems. After the Welch's t-test with FDR correction, a number of 13 categories were significantly altered by forest conversion with no seasonal effect. (Figure 2.5, corrected q-values are shown in lower panels).

Figure 2.5 - Relative abundance of soil functional categories at subsystems level 1 based on shotgun metagenomics. Only significantly altered categories are shown: forest (green bars and circles), grassland (red bars and circles), and no-till (brown bars and circles). Error bars show standard deviation. Panels: (a) Percentages of sequence reads are shown for two sampling seasons. Differences in mean proportions are shown: (b) comparison between forest and grassland and (c) comparison between forest and no-till. A Welch's t-test with corrected q-values calculated using the Benjamini-Hochberg false discovery rate was performed (q-value < 0.05)

Long-term grasslands resulted in changes in relative abundance for a total of 12 categories with increases in functions associated with carbohydrates, respiration, virulence, sulfur metabolism, regulation and cell signaling, motility and chemotaxis and iron acquisition and metabolism, while amino acids and derivatives, fatty acids, lipids and isoprenoids, cofactors and vitamins, metabolism of aromatic compounds and photosynthesis decreased (Figure 2.5b). When forests were converted to long-term no-till systems, only four categories were significantly changed, with functions associated to amino acids and nucleosides and nucleotides increasing and those related to photosynthesis and virulence decreasing (Figure 2.5c).

2.3.6 Environmental factors associated with taxonomic and functional patterns

We performed an Euclidean distance-based redundancy analysis (db-RDA) for 39 explanatory variables (35 soil physico-chemical parameters, 3 geographical coordinates, and time of sampling) to identify factors modulating soil microbial structure at taxonomic and functional levels. Overall, when all 39 factors were computed in our model, the total variation in taxonomic and functional structures were 40.2 and 4.8%, respectively. However, when a stepwise forward selection was performed, the subset of predictors ($P_{\text{adjusted}} < 0.05$) with the highest scores contributed to 36.8% of the explanation in taxonomic structure (Figure 2.6a, notice that 31.7% is represented in the first two axes of the db-RDA, while 5.1% is explained in the remaining axes). The P content was the main factor associated with differences in community composition, explaining 20.9% of the total variation (pseudo-F = 9.0; $P_{\text{adjusted}} = 0.01$), followed by $\text{Ca}^{2+}/\text{Mg}^{2+}$ ratio and N-NO_3^- content explaining 8.1% (pseudo-F = 3.8; $P_{\text{adjusted}} = 0.025$) and 7.8% (pseudo-F = 3.9; $P_{\text{adjusted}} = 0.031$), respectively.

Figure 2.6 - Distance-based redundancy analysis (db-RDA) of soil microbial communities at (a) genus level and (b) functional categories at level 2. Plots were generated using Euclidean distance matrices with 1,000 Monte-Carlo permutations. Vectors represent environmental variables, which were forward-selected ($P_{\text{adjusted}} < 0.05$). Samples were colored as follows: Forest (green circles), grassland (red circles) and no-till (brown circles)

When the same analysis was performed for functional structure, the forward-selected variables ($P_{\text{adjusted}} < 0.05$) explained 28.7% of the total variation. Phosphorus was again the major factor contributing to 14.3% (pseudo-F = 5.7; $P_{\text{adjusted}} = 0.014$) of the explained variation, followed by penetration resistance (8.3%; pseudo-F = 3.5; $P_{\text{adjusted}} = 0.014$), C/N ratio (6.6%; pseudo-F = 3.0; $P_{\text{adjusted}} = 0.03$), and longitude (5.8%; pseudo-F = 2.8, $P_{\text{adjusted}} = 0.01$) (Fig. 2.6b, notice that only 28.7% is represented in the first two axes).

2.3.7 Correlations among soil microbial taxa, functional categories and environmental factors

We calculated pairwise Spearman's correlation coefficients to establish co-occurrence patterns between all taxonomic, functional, and environmental data within each land use (Table 2.2 and Supplementary Table ST 2.4).

Table 2.2 - Number of correlations as inferred by Spearman's coefficient calculated among microbial taxonomy at phylum level, functions at SEED subsystem level 1, and environmental variables

Land-use \ Pairwise (n)*	Taxonomic	Functional	Taxonomic X functional	Taxonomic X environmental	Functional X environmental	Total
Forest						
Pairwise correlations ¹	561	378	944	1,348	1,091	4,322
Significant correlations ²	29 (10)	6 (0)	18 (3)	1 (0)	11 (2)	65 (15)
Significant positive correlations	25 (9)	2 (0)	9 (3)	0 (0)	7 (2)	43 (14)
Significant negative correlations	4 (1)	4 (0)	9 (0)	1 (0)	4 (0)	22 (1)
Grassland						
Pairwise correlations	560	376	948	1,352	1,103	4,339
Significant correlations	138 (65)	28 (9)	49 (17)	394 (109)	28 (0)	637 (200)
Significant positive correlations	128 (63)	12 (5)	26 (9)	199 (54)	16 (0)	381 (131)
Significant negative correlations	10 (2)	16 (4)	23 (8)	195 (55)	12 (0)	256 (69)
No-till						
Pairwise correlations	560	375	947	1,343	1,107	4,332
Significant correlations	39 (7)	19 (4)	39 (11)	38 (6)	49 (9)	184 (37)
Significant positive correlations	35 (3)	8 (1)	17 (6)	24 (3)	21 (6)	105 (19)
Significant negative correlations	4 (4)	11 (3)	22 (5)	14 (3)	28 (3)	79 (18)

¹ Number of possible pairwise correlations

² Number of strong (Spearman's $\rho \geq |0.7|$) and significant ($P \leq 0.01$) correlations. Values shown in parenthesis indicate a significance of $P < 0.001$.

Grassland samples had the highest number of significant correlations (637), followed by no-till (184) and forest (65). All land uses showed a higher number of total positive correlations in comparison to the negative ones. Correlations were consistently higher for grassland at taxa-taxa, function-function, taxa-function and taxa-environment combinations in comparison to no-till and forest, with the latter showing the lowest numbers of correlations for all combinations.

In order to visualize the connectivity patterns among the three dataset, we generated networks for strong (Spearman's $\rho \geq |0.7|$) and highly significant ($P < 0.01$) pairwise correlations (Figure 2.7). The forest network had a larger diameter (7) and higher modularity (0.65) and average path length (2.65) in comparison to converted systems, but lower number of nodes (55), edges (65), average clustering (0.10), and average degree (2.40). The complexity of the network in grassland was higher than that measured for the no-till system, evidenced by a higher number of nodes, edges, positive and negative edges, average degree and average weighted degree (Table 2.3). The grassland network structure was found to be non-modular (modularity = 0.11).

Figure 2.7 - Networks based on Spearman's pairwise correlation matrices at phylum level (blue nodes), functional level 1 (red nodes), and environmental variables (green nodes). Bold edges represent strong and positive correlations ($\rho \geq 0.7$), while regular edges represent strong and negative correlations ($\rho \leq -0.7$). Only significant edges are shown ($P < 0.01$). Size of each node is proportional to its number of edges. Labels as follows: Aq, Aquificae; Ba, Bacteroidetes; Ch, Chlorobi; Cl, Chlamydiae; Cy, Cyanobacteria; De, Deferribacteres; DT, Deinococcus-Thermus; Di, Dictyoglomi; El, Elusimicrobia; Eu, Euryarchaeota; Fi, Firmicutes; Le, Lentisphaerae; Ni, Nitrospirae; Pl, Planctomycetes; Po, Poribacteria; Pr, Proteobacteria; Sp, Spirochaetes; Te, Tenericutes; Th, Thaumarchaeota; CDCC, cell division and cell cycle; DNA met., DNA metabolism; MAC, metabolism of aromatic compounds; RNA met., RNA metabolism; SM, sulfur metabolism

Four phyla, Firmicutes, Verrucomicrobia, Elusimicrobia, and Thaumarchaeota, had a high number of correlations among themselves as well as with other factors in the forest network. The Proteobacteria, although not directly connected with the above phyla, presented high number of correlations with functional categories and environmental variables. After long-term land use change, co-occurring phyla were altered. The taxa with more connections in the no-till network were Deinococcus-Thermus, Euryarchaeota, Nitrospirae,

Planctomycetes and Proteobacteria, but none of them were predominant. The grassland network had a high number of correlations for *Deinococcus-Thermus*, Cyanobacteria, Aquificae, Tenericutes and Deferribacteres, which were strongly connected among them. Functional categories with more correlations in the forest were metabolism of aromatic compounds, cell division and cell cycle, DNA metabolism and sulfur metabolism. Long-term forest conversion showed a high number of connections for functions related to metabolism of aromatic compounds and RNA metabolism in grassland and sulfur metabolism and metabolism of aromatic compounds in the no-till system.

Table 2.3 - Topological properties of networks. Each network represents the number correlations inside taxonomy, functions, their interactions, as well as their correlations with environmental variables

Network properties	Land-use		
	Forest	Grassland	No-till
Number of nodes ^a	55	91	85
Number of edges ^b	65 (15)	637 (200)	184 (37)
Number of positive edges ^c	43 (14)	381 (131)	105 (19)
Number of negative edges ^d	22 (1)	256 (69)	79 (18)
Network diameter ^e	7	5	5
Modularity ^f	0.65	0.11	0.55
Average path length ^g	2.65	1.72	2.28
Average clustering coefficient ^h	0.10	0.20	0.15
Average degree ⁱ	2.40	14.00	4.32
Average weighted degree ^j	0.61	2.37	0.49

^a Number of features with at least one strong correlation (Spearman's $\rho \geq |0.7|$ and statistically high significant ($P \leq 0.01$). The values in parentheses show higher significance correlations ($P < 0.001$).

^b Number of connections obtained by pairwise correlations between nodes.

^c Spearman's pairwise positive correlations ($\rho \geq 0.7$) and high significant ($P < 0.01$).

^d Spearman's pairwise negative correlations ($\rho \leq -0.7$) and high significant ($P < 0.01$).

^e The longest distance between nodes in the network, measured in number of edges.

^f The capability of the nodes to form a modular structure, that is, a structure with high density of nodes and clustered topology. Values > 0.4 suggest that the network has a modular structure.

^g Average network distance between all pair of nodes or the average length of all edges in the network.

^h How nodes are embedded in their neighborhood and the degree to which they tend to cluster together.

ⁱ The average number of connections per node in the network, that is, the node connectivity.

^j The average number of connections per node in the network, corrected according to the proportion of nodes and edges in the network.

2.4 Discussion

Overall, we found that community alpha diversity did not vary among land uses in space and time, while the Beta diversity decreased in grassland. Our results agree with previous reports that tropical forest-to-agriculture conversion maintains (LEE-CRUZ et al., 2013) or increases (JESUS et al., 2009) alpha diversity, but leads to significant decreases in beta diversity in grassland systems (MENDES et al., 2015; RODRIGUES et al., 2013).

Many studies have described the consequences of land use change to microbial dynamics (DRENOVSKY et al., 2010; FIERER et al., 2012; KURAMAE et al., 2012; LAUBER et al., 2013), but the processes governing the alterations are often neglected (KNELMAN; NEMERGUT, 2014). Owing to observed differences in diversity, abundance and composition for our long-term converted study systems, we hypothesized that grassland and no-till microbial communities were assembled differently.

We calculated the corrected Akaike Information Criterion ($AIC_{corrected}$) for five different rank abundance models: two of them fitted for neutral- (HUBBELL, 2001; VOLKOV et al., 2003) and three for niche-based assembly (OFITERU et al., 2010; PHOLCHAN et al., 2013). There were two distinct patterns of assembly in our study: most of the forest and no-till samples showed a lognormal distribution, while grassland samples fitted the zero-sum multinomial model (Table 2.4). These results suggest that forest and no-till microbial communities have assemblages based on stochastic processes and, thus, are less prone to be influenced by selective pressures imposed by the surrounding environment. Differently, grassland communities have a niche-based assembly, a pattern that reflects the importance of environmental factors in regulating abundance and composition of the microbial community (DINI-ANDREOTE et al., 2015). We found support in our results that abiotic drivers (i.e. physical and chemical variables) have a stronger influence on grassland microbial communities than those observed in no-till and forest: first, grassland taxa showed a high number of significant correlations ($n = 394$, $P < 0.01$) associated with abiotic parameters (Table 2.1); second, the grassland network was found to be non-modular with a disproportional number of environmental variables surrounding phyla in a single cluster (Figure 2.7). This characteristic of non-modularity combined with a short average path length could lead to a prompt response of microbial communities to environmental perturbations (FAUST; RAES, 2012).

Table 2.4 - Samples fitting to theoretical ecological models of assembly, based on Akaike Information Criterion (AIC) for rank abundance distribution at genus level, based on M5nr database. The tendency to migration was calculated through dispersal rate by the Etienne's formula

Land-use	Location	Sampling	Akaike Information Criterion ^a					<i>Dispersal</i> ^b
			Neutral		Niche-based			
			<i>Null</i>	<i>ZSM</i>	<i>Lognormal Preemption</i>	<i>Zipf</i>		
Forest	Chapecó	Winter	1402.1	1252.0	1238.1	1347.9	1311.3	0.11
Forest	Chapecó	Summer	1274.3	1172.8	1163.2	1232.8	1217.1	0.07
Forest	Xanxerê	Winter	1401.6	1244.9	1240.2	1347.4	1308.5	0.08
Forest	Xanxerê	Summer	1781.0	1405.0	1435.1	1671.8	1568.6	0.06
Grassland	Chapecó	Winter	2172.7	1643.6	1699.7	1958.4	1881.9	0.18
Grassland	Chapecó	Summer	1713.4	1523.9	1518.0	1640.5	1600.1	0.64
Grassland	Xanxerê	Winter	1973.8	1608.9	1613.1	1845.6	1721.1	0.11
Grassland	Xanxerê	Summer	2231.3	1671.7	1711.6	2029.3	1887.8	0.58
No-till	Chapecó	Winter	1411.7	1241.3	1230.8	1344.3	1311.0	0.08
No-till	Chapecó	Summer	1287.8	1192.1	1174.0	1237.2	1252.8	0.05
No-till	Xanxerê	Winter	1371.3	1212.2	1207.0	1311.3	1289.3	0.08
No-till	Xanxerê	Summer	1665.6	1368.6	1396.2	1554.3	1543.2	0.05

^a Poisson Rank abundance distribution models, based on corrected AIC values, using maximum likelihood estimation. The lowest AIC value for each sample represented the best fitted model for general community's assembly (in bold). AIC corrected values were calculated by the Radfit function on Vegan R Package, with exception of zero-sum model distribution, which was calculated using TeTame Software. AIC calculated by the general equation $AIC = -2\log\text{-likelihood} + 2 \times \text{npar}$.

^b Dispersal rate were calculated by Etienne's formula, using TeTame Software. Values of dispersal are between 0 and 1, where 0 means no tendency to migration and 1 means total tendency to migration in a certain community. Dispersal values ≥ 0.3 were highlighted (in bold).

We investigated the relative importance of the ecological processes modulating microbial assembly. While diversification and dispersal apparently acted on the composition of microbial communities, selection and drift clearly altered the relative abundance of their members. This is consistent with previous observations (NEMERGUT et al., 2013; STEGEN et al., 2013). Furthermore, we determined that the long-term forest to grassland conversion reduced diversification. This was true not only for the beta diversity, a measure of shared taxa between samples in space and/or time, but also for structure and distance to the nearest neighbor, both measures of community similarity. Together, these results indicate that grassland communities have a more homogeneous composition. When the rate of dispersal

was calculated for each community, forest and no-till samples had average values of 0.08 and 0.07, respectively, suggesting their microbial populations had a very low tendency to migrate. Conversely, grassland samples showed consistently higher dispersal rates (average $m = 0.38$). A combination of decreased diversification and increased dispersal, as observed in grassland, often leads to biotic homogenization (OLDEN, 2006). The weight of selection was calculated using a variation partitioning of redundancy analysis (pRDA) that decomposed environment, geography + time, and their interaction.

Figure 2.8 - Variation partitioning of redundancy analysis (pRDA) with forward selection, using Principal Coordinates of Neighbor Matrices (PCNM). Data show the adjusted coefficient of multiple determination (R^2) from simple effects of environmental variables (black), geography + time (gray), and their interactions (dashed gray). Latitude and longitude were used as constraining spatial coordinates. Panels: (a) Taxonomy at genus level based on M5nr database and (b) Function based on the SEED subsystems level 2 database. Symbols mean: * = $P_{\text{adjusted}} < 0.05$, ** = $P_{\text{adjusted}} < 0.01$, *** = $P_{\text{adjusted}} < 0.001$

The variation explained by selection in grassland was 83.8% ($P < 0.001$), followed by forest and no-till with 40.6% ($P < 0.05$) and 20.9%, respectively (Figure 2.8a), implying that selection is an important driver of community assembly in grassland. Accordingly, forest-to-grassland conversion has been shown to alter microbial abundance as a result of changes in C and N concentrations (CENCIANI et al., 2009; NÜSSLEIN; TIEDJE, 1999; RODRIGUES et al., 2013). Owing to the difficulty of directly measuring the contribution of drift for non-discrete populations in large communities, the process was estimated as a component of the pRDA residue — non-explained variation + stochastic forces (STEGEN et al., 2013; VELLEND, 2010). When selection is weak (low R^2) and residual variation is high, such as

observed in our study for no-till and forest, drift is likely to play an important role in community assembly (KNELMAN; NEMERGUT, 2014; NEMERGUT et al., 2013). These results reflected in the importance of selection to soil functions. Selection showed no effect on functional composition in forest and no-till, with moderate effect on grassland (Figure 2.8b).

2.4.1 A working model for microbial community dynamics in long-term agricultural conversion

Together, the above results suggest that land use is more important driver for microbial assembly than the conversion itself. While the no-till system, in our study, is a successive rotation of soybean and/or maize in the summer, followed by wheat in the winter; the grassland is dominated by the perennial species *Axonopous affinis*. Two complementary factors are postulated to influence soil microbial assembly under a grass monoculture: First, the restriction in range of C and energy sources in grassland subjects microbial communities to a strong process of selection (WAWRIK et al., 2005). Second, the grass root system increases the soil connectivity by removing potential barriers to migration (e.g. aggregate-to-aggregate continuity) and leads to higher microbial dispersal in space and time (MURPHY; FOSTER, 2014). As a consequence, the overall microbial diversity is impacted. However, the no-till cropping with rotation increases carbon storage by maintaining residues from previous cultivations and supplying microbial communities with more heterogeneous sources. Classical ecological theory states that species coexistence increases as a result of niche partitioning due to resource availability (CARROLL; CARDINALE; NISBET, 2011) and empirical support has been provided for plant and animal communities (FINKE; SNYDER, 2008; MORI et al., 2013). We found that resource heterogeneity in no-till increased niche specialization and reduced the importance of the process of selection on assembly.

2.4.2 Linking microbial assembly and community functions

Niche and neutral processes have been studied in terrestrial ecosystems (DINI-ANDREOTE et al., 2015; FERRENBURG et al., 2013; STEGEN et al., 2012), but the relationship between assembly and function has yet to be investigated (KNELMAN; NEMERGUT, 2014). Given that our results point towards distinctive assemblies according to land use, we argue as follows: (1) if community assembly occurs mostly by neutral processes, environmental conditions should exert a stronger effect on microbial functions; and (2) if

community assembly is mostly determined by niche-based processes, the environment should be less likely to modulate functions. While the grassland communities were found to be neutral, forest and no-till communities had a predominantly niche-based assembly. We attribute this observation in no-till as result of functional niche partitioning, in which species can coexist by differential resource utilization. When resources are more heterogeneous, lower rates of dispersal and weak selection lead to decrease in competition, thus functions are maintained through niche specialization (BLÜTHGEN; KLEIN, 2011). A second implication of niche partitioning in no-till is the buffering effect against species losses (BIHN; GEBAUER; BRANDL, 2010; MOUILLOT et al., 2013), as observed by similar values of global beta diversity when compared to those found in forest. In grassland, root exudates are mostly composed of easily degradable and homogenous C sources that might intensify the process of selection, thus leading to an increase in functional categories related to carbohydrates. As a response to environmental pressure, microbial communities increase functional categories associated with stress responses (regulation and cell signaling, respiration, virulence, disease and defense) and motility and chemotaxis (Figure 2.5b), the latter associated with the moderate tendency to migration.

2.5 Conclusion

Forest conversion to agricultural systems has been generally regarded as detrimental to biodiversity. We have shown, by analyzing a large range of taxa, functional and environmental data, that soil microbial community assembly is not dependent on conversion, but on the land use after conversion, and influenced by the net biodiversity outcome. Furthermore, the relative weight of divergent ecological processes, depending on land use, determined a niche- or neutral-based model of assembly, with implications to function. We showed that long-term grassland led to biotic homogenization and functional stress, through environmental selection. In the other hand, long-term no-till cropping has shown no differences, when compared to forest. Future investigations stand to gain valuable information from assessing ecological processes combined effect in different ecosystems and across multiple scales of space and time. Understanding microbial community assembly as well as characterizing fundamental processes that guide dynamic responses in community organization have the potential to provide important insights into ecosystem modeling of biodiversity and microbial conservation.

References

- BARBERÁN, A.; BATES, S.T.; CASAMAYOR, E.O.; FIERER, N. Using network analysis to explore co-occurrence patterns in soil microbial communities. **The ISME Journal**, London, v. 6, n. 2, p. 343–351, 2012.
- BASTIAN, M.; HEYMAN, S.; JACOMY, M. Gephi: an open source software for exploring and manipulating networks. In: INTERNATIONAL AAAI CONFERENCE ON WEBLOGS AND SOCIAL MEDIA, 2009, San Jose. **Proceedings...** Disponível em: <<http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>>. Acesso em: 10 nov. 2014.
- BENJAMINI, Y.; HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society**, London, v. 57, n. 1, p. 289–300, 1995.
- BIHN, J.H.; GEBAUER, G.; BRANDL, R. Loss of functional diversity of ant assemblages in secondary tropical forests. **Ecology**, Ithaca, v. 91, n. 3, p. 782–792, 2010.
- BLÜTHGEN, N.; KLEIN, A.M. Functional complementarity and specialisation: The role of biodiversity in plant-pollinator interactions. **Basic and Applied Ecology**, Amsterdam, v. 12, n. 4, p. 282–291, 2011.
- BRODY, J.R.; KERN, S.E. Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis. **BioTechniques**, New York, v. 36, n. 2, p. 214–216, 2004.
- CARBONETTO, B.; RASCOVAN, N.; ÁLVAREZ, R.; MENTABERRY, A.; VÁZQUEZ, M.P. Structure, composition and metagenomic profile of soil microbiomes associated to agricultural land use and tillage systems in Argentine Pampas. **PLoS One**, New York, v. 9, n. 6, p. 1–11, 2014.
- 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.; KINZIG, A.P.; DAILY, G.C.; LOREAU, M.; GRACE, J.B.; LARIGAUDERIE, A.; SRIVASTAVA, D.S.; NAEEM, S. Biodiversity loss and its impact on humanity. **Nature**, London, v. 489, n. 7415, p. 326–326, 2012.
- CARROLL, I.T.; CARDINALE, B.J.; NISBET, R.M. Niche and fitness differences relate the maintenance of diversity to ecosystem function. **Ecology**, Ithaca, v. 92, n. 5, p. 1157–1165, 2011.
- CENCIANI, K.; LAMBAIS, M.R.; CERRI, C.C.; AZEVEDO, L.C.B.; FEIGL, B.J. Bacteria diversity and microbial biomass in forest, pasture and fallow soils in the southwestern Amazon basin. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 33, n. 1, p. 907–916, 2009.
- CHAPIN, F.S.; ZAVALA, E.S.; EVINER, V.T.; NAYLOR, R.L.; VITOUSEK, P.M.; REYNOLDS, H.L.; HOOPER, D.U.; LAVOREL, S.; SALA, O.E.; HOBIE, S.E.; MACK, M.C.; DÍAZ, S. Consequences of changing biodiversity. **Nature**, London, v. 405, n. 6783, p. 234–242, 2000.

DINI-ANDREOTE, F.; STEGEN, J.C.; VAN ELSAS, J.D.; SALLES, J.F. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. **Proceedings of the National Academy of Sciences**, Washington, v. 112, n. 11, p. 1326–1332, 2015.

DRENOVSKY, R.E.; STEENWERTH, K.L.; JACKSON, L.E.; SCOW, K.M. Land use and climatic factors structure regional patterns in soil microbial communities. **Global Ecology and Biogeography**, Malden, v. 19, n. 1, p. 27–39, 2010.

FAUST, K.; RAES, J. Microbial interactions: from networks to models. **Nature Reviews Microbiology**, London, v. 10, n. 8, p. 538–550, 2012.

FEELEY, K.J.; SILMAN, M.R. Extinction risks of Amazonian plant species. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 106, n. 30, p. 12382–12387, 2009.

FERRENBURG, S.; O'NEILL, S.P.; KNELMAN, J.E.; TODD, B.; DUGGAN, S.; BRADLEY, D.; ROBINSON, T.; SCHMIDT, S.K.; TOWNSEND, A.R.; WILLIAMS, M.W.; CLEVELAND, C.C.; MELBOURNE, B.A.; JIANG, L.; NEMERGUT, D.R. Changes in assembly processes in soil bacterial communities following a wildfire disturbance. **The ISME Journal**, London, v. 7, n. 6, p. 1102–1111, 2013.

FIERER, N.; LENNON, J.T. The generation and maintenance of diversity in microbial communities. **American Journal of Botany**, Saint Louis, v. 98, n. 3, p. 439–448, 2011.

FIERER, N.; LAUBER, C.L.; RAMIREZ, K.S.; ZANEVELD, J.; BRADFORD, M.A.; KNIGHT, R. Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. **The ISME Journal**, London, v. 6, n. 5, p. 1007–1017, 2012.

FIGUEROLA, E.L.M.; GUERRERO, L.D.; ROSA, S.M.; SIMONETTI, L.; DUVAL, M.E.; GALANTINI, J.A.; BEDANO, J.C.; WALL, L.G.; ERIJMAN, L. Bacterial indicator of agricultural management for soil under no-till crop production. **PloS One**, New York, v. 7, n. 11, p. 1–12, 2012.

FINKE, D.L.; SNYDER, W.E. Niche partitioning increases resource exploitation by diverse communities. **Science**, Washington, v. 321, p. 1488–1490, 2008.

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. Global consequences of land use. **Science**, Washington, v. 309, n. 5734, p. 570–574, 2005.

FOLEY, J.A.; RAMANKUTTY, N.; BRAUMAN, K.A.; CASSIDY, E.S.; GERBER, J.S.; JOHNSTON, M.; MUELLER, N.D.; O'CONNELL, C.; RAY, D.K.; WEST, P.C.; BALZER, C.; BENNETT, E.M.; CARPENTER, S.R.; HILL, J.; MONFREDA, C.; POLASKY, S.; ROCKSTRÖM, J.; SHEEHAN, J.; SIEBERT, S.; TILMAN, D.; ZAKS, D.P.M. Solutions for a cultivated planet. **Nature**, London, v. 478, n. 7369, p. 337–342, 2011.

HAMMER, Ø.; HARPER, D.A.T.; RYAN, P.D. PAST: Paleontological statistics software package for education and data analysis. **Paleontologia Electronica**, College Station, v. 4, n. 1, p. 1–9, 2001.

HUBBELL, S.P. **The unified neutral theory of biodiversity and biogeography**. Princeton: Princeton University Press, 2001. 448 p.

JABOT, F.; ETIENNE, R.S.; CHAVE, J. Reconciling neutral community models and environmental filtering: theory and an empirical test. **Oikos**, Copenhagen, v. 117, p. 1308–1320, 2008.

JESUS, E.C.; MARSH, T.L.; TIEDJE, J.M.; MOREIRA, F.M. Changes in land use alter the structure of bacterial communities in Western Amazon soils. **The ISME Journal**, London, v. 3, n. 9, p. 1004–1011, 2009.

KNELMAN, J.E.; NEMERGUT, D.R. Changes in community assembly may shift the relationship between biodiversity and ecosystem function. **Frontiers in Microbiology**, Lausanne, v. 5, p. 1–4, 2014.

KURAMAE, E.E.; YERGEAU, E.; WONG, L.C.; PIJL, A.S.; VAN VEEN, J.A.; KOWALCHUK, G.A. Soil characteristics more strongly influence soil bacterial communities than land-use type. **FEMS Microbiology Ecology**, Amsterdam, v. 79, n. 1, p. 12–24, 2012.

LAUBER, C.L.; RAMIREZ, K.S.; AANDERUD, Z.; LENNON, J.; FIERER, N. Temporal variability in soil microbial communities across land-use types. **The ISME Journal**, London, v. 7, n. 8, p. 1641–1650, 2013.

LEE-CRUZ, L.; EDWARDS, D.P.; TRIPATHI, B.M.; ADAMS, J.M. Impact of logging and forest conversion to oil palm plantations on soil bacterial communities in borneo. **Applied and Environmental Microbiology**, Baltimore, v. 79, n. 23, p. 7290–7297, 2013.

LEIBOLD, M.A.; HOLYOAK, M.; MOUQUET, N.; AMARASEKARE, P.; CHASE, J.M.; HOOPES, M.F.; HOLT, R.D.; SHURIN, J.B.; LAW, R.; TILMAN, D.; LOREAU, M.; GONZALEZ, A. The metacommunity concept: a framework for multi-scale community ecology. **Ecology Letters**, Oxford, v. 7, n. 7, p. 601–613, 2004.

LEPŠ, J.; ŠMILAUER, P. Multivariate Analysis of Ecological Data using CANOCO. **Bulletin of the Ecological Society of America**, Ithaca, v. 86, p. 6–6, 2005.

LIENHARD, P.; TIVET, F.; CHABANNE, A.; DEQUIEDT, S.; LELIÈVRE, M.; SAYPHOUMMIE, S.; LEUDPHANANE, B.; PRÉVOST-BOURÉ, N.C.; SÉGUY, L.; MARON, P.A.; RANJARD, L. No-till and cover crops shift soil microbial abundance and diversity in Laos tropical grasslands. **Agronomy for Sustainable Development**, Dijon, v. 33, n. 2, p. 375–384, 2013.

MACARTHUR, R.H.; WILSON, E.O. An equilibrium theory of insular zoogeography. **Evolution**, New York, v. 17, n. 4, p. 373–387, 1963.

MCCONKEY, K.R.; FARRILL, G.O. Cryptic function loss in animal populations. **Trends in Ecology & Evolution**, Cambridge, v. 30, n. 4, p. 182–189, 2015.

MENDES, L.W.; KURAMAE, E.E.; NAVARRETE, A.A.; VAN VEEN, J.A.; TSAI, S.M. Taxonomical and functional microbial community selection in soybean rhizosphere. **The ISME Journal**, London, v. 8, p. 1577–1587, 2014.

MENDES, L.W.; TSAI, S.M.; NAVARRETE, A.A.; DE HOLLANDER, M.; VAN VEEN, J.A.; KURAMAE, E.E. Soil-borne microbiome: linking diversity to function. **Microbial Ecology**, New York, v. 70, p. 255–265, 2015.

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. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. **BMC Bioinformatics**, London, v. 9, n. 386, p. 1–8, 2008.

MORI, A.S.; SHIONO, T.; KOIDE, D.; KITAGAWA, R.; OTA, A.T.; MIZUMACHI, E. Community assembly processes shape an altitudinal gradient of forest biodiversity. **Global Ecology and Biogeography**, Malden, v. 22, n. 7, p. 878–888, 2013.

MOUILLOT, D.; BELLWOOD, D.R.; BARALOTO, C.; CHAVE, J.; GALZIN, R.; HARMELIN-VIVIEN, M.; KULBICKI, M.; LAVERGNE, S.; LAVOREL, S.; MOUQUET, N.; PAINE, C.E.T.; RENAUD, J.; THUILLER, W. Rare species support vulnerable functions in high-diversity ecosystems. **PLoS Biology**, New York, v. 11, n. 5, p. 1–11, 2013.

MUELLER, R.C.; PAULA, F.S.; MIRZA, B.S.; RODRIGUES, J.L.M.; NÜSSLEIN, K.; BOHANNAN, B.J.M. Links between plant and fungal communities across a deforestation chronosequence in the Amazon rainforest. **The ISME Journal**, London, v. 8, p. 1548–1550, 2014.

MURPHY, C.A.; FOSTER, B.L. Soil properties and spatial processes influence bacterial metacommunities within a grassland restoration experiment. **Restoration Ecology**, Malden, v. 22, n. 5, p. 685–691, 2014.

MYERS, N.; MITTERMEIER, R.A.; MITTERMEIER, C.G.; FONSECA, G.A.B.; KENT, J. Biodiversity hotspots for conservation priorities. **Nature**, London, v. 403, n. 6772, p. 853–858, 2000.

NAVARRETE, A.A.; TAKETANI, R.G.; MENDES, L.W.; CANNAVAN, F.D.S.; MOREIRA, F.M.S.; TSAI, S.M. Land-use systems affect archaeal community structure and functional diversity in western amazon soils. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 35, n. 1, p. 1527–1540, 2011.

NEMERGUT, D.R.; SCHMIDT, S.K.; FUKAMI, T.; O’NEILL, S.P.; BILINSKI, T.M.; STANISH, L.F.; KNELMAN, J.E.; DARCY, J.L.; LYNCH, R.C.; WICKEY, P.; FERRENBURG, S. Patterns and processes of microbial community assembly. **Microbiology and Molecular Biology Reviews**, Washington, v. 77, n. 3, p. 342–356, 2013.

NEWCOMBE, R.G. Improved confidence intervals for the difference between binomial proportions based on paired data. **Statistics in Medicine**, Chichester, v. 17, p. 2635–2650, 1998.

NÜSSLEIN, K.; TIEDJE, J.M. Soil bacterial community shift correlated with change from forest to pasture vegetation in a tropical soil. **Applied and Environmental Microbiology**, Baltimore, v. 65, n. 8, p. 3622–3626, 1999.

OFITERU, I.D.; LUNN, M.; CURTIS, T.P.; WELLS, G.F.; CRIDDLE, C.S.; FRANCIS, C.A.; SLOAN, W.T. Combined niche and neutral effects in a microbial wastewater treatment community. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 107, n. 35, p. 15345–15350, 2010.

OLDEN, J.D. Biotic homogenization: a new research agenda for conservation biogeography. **Journal of Biogeography**, Malden, v. 33, p. 2027–2039, 2006.

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. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). **Nucleic Acids Research**, Oxford, v. 42, p. 1–9, 2013.

PARKS, D.H.; BEIKO, R.G. Identifying biologically relevant differences between metagenomic communities. **Bioinformatics**, Oxford, v. 26, n. 6, p. 715–721, 2010.

PEREIRA, H.M.; LEADLEY, P.W.; PROENÇA, V.; ALKEMADE, R.; SCHARLEMANN, J.P.W.; FERNANDEZ-MANJARRÉS, J.F.; ARAÚJO, M.B.; BALVANERA, P.; BIGGS, R.; CHEUNG, W.W.L.; CHINI, L.; COOPER, H.D.; GILMAN, E.L.; GUÉNETTE, S.; HURTT, G.C.; HUNTINGTON, H.P.; MACE, G.M.; OBERDORFF, T.; REVENGA, C.; RODRIGUES, P.; SCHOLLES, R.J.; SUMAILA, U.R.; WALPOLE, M. Scenarios for global biodiversity in the 21st century. **Science**, Washington, v. 330, n. 6010, p. 1496–1501, 2010.

PHOLCHAN, M.K.; BAPTISTA, J.C.; DAVENPORT, R.J.; SLOAN, W.T.; CURTIS, T.P. Microbial community assembly, theory and rare functions. **Frontiers in Microbiology**, Lausanne, v. 4, p. 1–9, 2013.

RIBEIRO, M.C.; METZGER, J.P.; MARTENSEN, A.C.; PONZONI, F.J.; HIROTA, M.M. The Brazilian Atlantic Forest: how much is left, and how is the remaining forest distributed? Implications for conservation. **Biological Conservation**, Cambridge v. 142, n. 6, p. 1141–1153, 2009.

RODRIGUES, J.L.M.; PELLIZARI, V.H.; MUELLER, R.; BAEK, K.; JESUS, E.D.C.; PAULA, F.S.; MIRZA, B.; HAMAOU, G.S.; TSAI, S.M.; FEIGL, B.; TIEDJE, J.M.; BOHANNAN, B.J.M.; NÜSSLEIN, K. Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 110, n. 3, p. 988–993, 2013.

SALA, O.E.; CHAPIN, F.S.; ARMESTO, J.J.; BERLOW, E.; BLOOMFIELD, J.; DIRZO, R.; HUBER-SANWALD, E.; HUENNEKE, L.F.; JACKSON, R.B.; KINZIG, A.; LEEMANS, R.; LODGE, D.M.; MOONEY, H.A.; OESTERHELD, M.; POFF, N.L.; SYKES, M.T.; WALKER, B.H.; WALKER, M.; WALL, D.H. Global biodiversity scenarios for the year 2100. **Science**, Washington, v. 287, n. 5459, p. 1770–1774, 2000.

STEGEN, J.C.; LIN, X.; KONOPKA, A.E.; FREDRICKSON, J.K. Stochastic and deterministic assembly processes in subsurface microbial communities. **The ISME Journal**, London, v. 6, n. 9, p. 1653–1664, 2012.

STEGEN, J.C.; LIN, X.; FREDRICKSON, J.K.; CHEN, X.; KENNEDY, D.W.; MURRAY, C.J.; ROCKHOLD, M.L.; KONOPKA, A. Quantifying community assembly processes and identifying features that impose them. **The ISME Journal**, London, v. 7, n. 11, p. 2069–2079, 2013.

TEAM, R.D.C. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2008. Disponível em: <<http://www.R-project.org>>. Acesso em: 17 set. 2015.

THOMSON, A.M.; CALVIN, K.V.; CHINI, L.P.; HURTT, G.; EDMONDS, J.A.; BOND-LAMBERTY, B.; FROLKING, S.; WISE, M.A.; JANETOS, A.C. Climate mitigation and the future of tropical landscapes. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 107, n. 46, p. 19633–19638, 2010.

TRINGE, S.G.; VON MERING, C.; KOBAYASHI, A.; SALAMOV, A.A.; CHEN, K.; CHANG, H.W.; PODAR, M.; SHORT, J.M.; MATHUR, E.J.; DETTER, J.C.; BORK, P.; HUGENHOLTZ, P.; RUBIN, E.M. Comparative metagenomics of microbial communities. **Science**, Washington, v. 308, n. 5721, p. 554–557, 2005.

VELLEND, M. Conceptual synthesis in community ecology. **The Quarterly Review of Biology**, Chicago, v. 85, n. 2, p. 183–206, 2010.

VOLKOV, I.; BANAVAR, J.R.; HUBBELL, S.P.; MARITAN, A. Neutral theory and relative species abundance in ecology. **Nature**, London, v. 424, n. 13, p. 1035–1037, 2003.

WANG, J.; SHEN, J.; WU, Y.; TU, C.; SOININEN, J.; STEGEN, J.C.; HE, J.; LIU, X.; ZHANG, L.; ZHANG, E. Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. **The ISME Journal**, London, v. 7, n. 7, p. 1310–1321, 2013.

WARDLE, D.A.; BARDGETT, R.D.; CALLAWAY, R.M.; VAN DER PUTTEN, W.H. Terrestrial ecosystem responses to species gains and losses. **Science**, Washington, v. 332, p. 1273–1278, 2011.

WAWRIK, B.; KERKHOF, L.; KUKOR, J.; ZYLSTRA, G. Effect of different carbon sources on community composition of bacterial enrichments from soil. **Applied and Environmental Microbiology**, Baltimore, v. 71, n. 11, p. 6776–6783, 2005.

WELCH, B.L. The generalization of Student's problem when several different population variances are involved. **Biometrika**, Oxford, v. 34, n. 1, p. 28–35, 1934.

WHITTAKER, R.H. Classification of natural communities. **The Botanical Review**, New York, v. 28, n. 1, p. 1–239, 1962.

WILKE, A.; HARRISON, T.; WILKENING, J.; FIELD, D.; GLASS, E.M.; KYRPIDES, N.; MAVROMMATIS, K.; MEYER, F. The M5nr: a novel non-redundant database containing

protein sequences and annotations from multiple sources and associated tools. **BMC Bioinformatics**, London, v. 13, n. 141, p. 1–5, 2012.

3 LONG-TERM MICROBIAL DYNAMICS AT THE SOYBEAN-RHIZOSPHERE INTERFACE IN AMAZON SOILS

Abstract

An emerging question in Soil Microbial Ecology is whether communities can reach an evolutionary equilibrium after disturbance. Works addressing this issue, often evaluate patterns related to selection, lacking the links with drift, dispersal and speciation processes. This issue acquires importance in the Amazon Biome, due the high rates of deforestation (4,800 km² last year), most for agriculture. The aim of this work was to evaluate the microbial dynamics in a long-term chronosequence of forest to no-till conversion (1 to 20-year) and the role of rhizosphere on this dynamics. We performed a soybean mesocosm experiment and characterized bulk soil and rhizosphere microbial metagenomes. Communities became more similar, the extent the chronosequence advanced, in both fractions. The α -diversity did not vary across fractions, but increased in both along time, while the β -diversity decreased. Acidobacteria, Nitrospirae, δ - and γ -Proteobacteria showed higher abundances in bulk soil, while Bacteroidetes, α - and β -Proteobacteria were more abundant in rhizosphere. Microbial trade-offs occurred mainly from 1- to 10-year in bulk soil and gradually increased in rhizosphere from 1- to 20-year. Several functions changed in bulk soil and just a few in rhizosphere, along the chronosequence. Network connections increased in bulk soil and decreased in rhizosphere, along time. Community assembly in bulk soil fitted to neutral model, with low to intermediate dispersal and higher influence of the environment (selection), driven by aluminum concentration and cations saturation, while in rhizosphere, the assembly fitted to niche-based model, with intermediate to high dispersal and lower environmental influence. The role of drift on microbial dynamics was not significant (< 50%). Along time, the increase of taxa was counterbalanced by enhanced shared-taxa into bulk soil and rhizosphere, leading to microbial homogenization. Taxa and functions in bulk soil changed along the chronosequence, serving as reservoir of biodiversity to the plant, mainly driven by selection. Functions in rhizosphere remained stable along the chronosequence, indicating functional resilience on plant-roots, via taxa trade-offs, mainly driven by dispersal.

Keywords: Bulk soil; Chronosequence; Rhizosphere; Root community; Straw

3.1 Introduction

The thin layer of soil surrounding and sometimes adhered to the roots of superior vascular plants is called rhizosphere. It comprises a fine-roots/soil continuum of 3-5 mm width, which is the adjacent and main source for water and nutrients absorption by the plants. In return, the roots release a variety of exudates, mucilage and other compounds to the rhizosphere, serving as source of carbon and energy to microorganisms, as well as chemotaxis signals, that lead members of microbial communities in the surroundings – the so called bulk soil – to recognize and occupy niches in this region. The magnitude of the recognition, food-web exchange and niche occupancy, is believed to be dependent on the source and quantity of

those substrates released by the roots, which are inwardly related to the plant cultivated. Except for some endophytes that come to soil adhered or even colonizing the seeds of cultivated plants, the source of microbial diversity in the rhizosphere is the bulk soil. Then, besides the effect of the plant, we would expect a complex set of ecological relationships and factors modulating the assembly of communities and their dynamics evolution along time in the soil-rhizosphere interface. Thus, the rhizosphere can be considered as a dynamic “hotspot” of microbial diversity and ecological interactions.

The emphasis in community ecology is mostly on the patterns rather than in the processes that govern community dynamics. Thus, the outcome is often descriptive, lacking the linkage between these patterns and the processes that rule them. Most of the works that cross the line between patterns and processes make inferences based on selection, which is recognized as an important process modulating microbial community evolution. The selection process is evaluated through deterministic interactions among sets of populations in a certain community (MENDES et al., 2014) and between the community and the environmental factors that can modulate its diversity, composition and abundance of species in a certain ecosystem (HOVATTER et al., 2011; JACKSON; FAHRIG, 2014; ROUSK et al., 2010). Despite of that, recently works have emphasized the need to embrace, besides selection, the processes of drift, dispersal and speciation, in order to achieve a more comprehensive framework of the community dynamics in ecosystems (STEGEN et al., 2013), such as the soil-rhizosphere interface.

Most of the ecological models based on selection were primarily developed for intra-specific interactions or interactions among two or a few species. The challenge in working with a complex and species-rich community, such as a soil microbial community, lies in the fact that the possible assembly configurations among species or sets of populations are almost limitless. Evaluating microbial community assembly at the light of ecological processes could make it easier to have a comprehensive picture about this boundless amount of patterns arrangements. Four basic ecological processes can embrace the patterns and all the possible arrangements: selection, dispersal, drift and speciation. In this case, we reduce the outcome from hundreds or thousands of combinations to 12 possible ways, that form the basis of theoretical models (VELLEND, 2010). Thus, it is possible for microbial ecologists to underlie these combinations, resulting in the theoretical community paradigms that better describe the dynamic behavior of microbes, such as soil communities.

The implementation of these ecological studies in microbial ecology have been challenging, given that, despite general theories in ecology fit well to most of the organisms,

these ecological processes and theories were primarily developed to evaluate the assembly of discrete populations of animals and plants, featuring limited number of species and high generation time, which makes it easier to identify and quantify the processes ruling the community dynamic evolution. In the other hand, microbial communities, such as the soil bacterial and archaeal, are major characterized by a limitless number of species, generally with non-discrete and non-normal abundance distribution and higher diversity, compared to the eukaryotes. Furthermore, evolutionary processes act in a faster and broader way, leading to increased difficulty on establishing the limits of interactions in a single community and its dynamics.

We used the framework proposed by Vellend (2010) with some conceptual implementations for microbial community conceived by Stegen et al. (2013) in order to advance our previous understanding about the microbial dynamics in agriculture soils, after the deforestation of the Amazon Rainforest (MENDES et al., 2015; MUELLER et al., 2014; NAVARRETE et al., 2013; RODRIGUES et al., 2013). Unravel the effects of long-term no-till on the ability of soybean roots to select specific members of microbial community (MENDES et al., 2014) was also a goal to be reached in this work. Thus, the aim of this work was to evaluate the microbial community dynamics in long-term no-till cropping and the power of selection of soybean along this chronosequence, using agriculture sites arisen by the deforestation of the Amazon Rainforest. The outcome was a framework of the microbial dynamics along soil fractions and time.

3.2 Material and Methods

3.2.1 Site description and soil sampling

Bulk soil samples were collected in January 2013, in agriculture fields, located in the municipality of Querência, Mato Grosso State, Brazil (Figure 3.1a). These fields were chosen because they are located in the last agriculture frontier of the Southeastern Brazilian Amazon, within the “Alto Xingú” water basin. Site selection was based in previous exploratory visit, as well as the land-use and management historic, obtained through interviews with the farmers and field specialists from the region. The three areas of the chronosequence were in the same landscape, to guarantee that only variation arising from time and management could major influence microbial dynamics.

Figure 3.1 - Description of sampling areas and sampling method. (a) Samples were collected in the municipality of Querência, Mato Grosso State, Brazil. The sampling sites are located in the frontier of Southeastern Amazon (dashed green line perimeter), following a chronosequence of no-till cropping, after deforestation, which includes: 1) first year of no-till cultivation; 2) 10-year successive cropping rotation, with maize/soybean in the summer and millet/ryegrass in the winter and; 3) 20-year successive cropping rotation. Soybean was the culture in the field when we collected the samples. (b) Sampling Cartesian geogrid. Samples were mixed to form composite samples, used in the greenhouse experiment and further analysis

All areas followed basically the same process of preparation to the cultivation. First, the sites were deforested via slash-and-burn. Soon after, they were cultivated with common rice (*Oryza sativa*) for one season and after that, continuously cultivated in a crop rotation system that included millet (*Pennisetum glaucum*) and ryegrass (*Lolium multiflorum*) black oat (*Avena strigosa*) in the winter season (as cover plant), followed by maize (*Zea mays*) and soybean (*Glycine max*) in the summer season, under no-till management. The climate of the region is Am (Köppen-Geiger classification), with annual average temperature of 27°C and annual precipitation of 1400 mm in 2013, composed by well-defined periods of wet and drought (Supplementary Figure SF 3.2).

In order to evaluate the microbial evolution dynamics we established a chronosequence sampling effort, including a recently deforested area, in which soybean was cultivated for the first time (after deforestation and one rice cultivation), hence-forth called 1-year, an area with no-till cropping in rotation since 2003, hence-forth called 10-year and, a last area with no-till cropping in rotation since 1993, hence-forth called 20-year. After deforestation, both areas received liming in the first and each 4th or 5th year, in order to

elevate and keep soil pH around 6, respectively. Fertilizers and pesticides were also regularly applied along time, according to cultivation demand and recommendation.

A total of 144 soil samples (3 areas X 48 samples) were collected from the 0-20 cm profile in a cross-geogrid scheme. Each 4 samples formed one composite sample that was mixed with its neighbor composite sample (Figure 3.1b). The litter layer was removed prior the soil collection and stored for further greenhouse experiment. All samples were transported to the Cellular and Molecular Biology Laboratory of the University of São Paulo, Piracicaba, Brazil, within 48 hours after sampling, for implementation of the greenhouse experiment.

3.2.2 Preparation and conduction of the greenhouse mesocosm experiment

Soil samples collected in the above-described chronosequence were prepared for a mesocosm experiment, where soybean (*Glycine max*), conventional cultivar BRS 232 was used as trap-plant. Farmers in the “Alto Xingú” region largely use that cultivar. Before sowing, the seeds were inoculated with a commercial inoculant containing *Bradyrhizobium elkanii* and *Bradyrhizobium japonicum* (1:1, V/V) in a final concentration of 10^{10} viable cells mL^{-1} of inoculant. Soon after, the seeds were treated with the fungicide Vitavax-Thiram 200 SC[®], according to the manufacturer’s instructions. Seed inoculation and treatment were performed the same way as farmers usually do, to guarantee experimental conditions reliability.

Plants were grown in a greenhouse, at CENA-USP, Piracicaba, in order to control the supplement of water, light and temperature as well as to avoid external environmental interference. It consisted of pre-sterilized plastic vases (25 cm height x 25 cm diameter) with a 4 cm pebbles layer at the bottom (for drainage). The vases were filled with 5 kg of the prior collected soil. Then, six seeds were sowed in each vase. The prior collected straw was distributed in the vases according to the quantity found in each sampling point in the field. The experiment was carried out with 36 vases, consisting in six vases with plant and six with no-plant per sampling chronosequence (3 periods X 6 vases X 2 (with or without plants) = 36 vases).

The soil moisture in all vases was corrected for 60% of water holding capacity (WHC = 60%) in the beginning of the experiment, and kept via irrigation with deionized water (each two days). Soybean plants were germinated at 28/20°C (day/night) and 12-hour photoperiod. Ten days after germination, seeds with lower vigor were removed from the vases. Only three plants were kept per vase. The experiment was conducted until stage R1 (50% flowering

plants), comprising the 65th day after sowing. The plants were carefully removed from the vases and transported in dry ice to the Cellular and Molecular Biology Lab, CENA-USP, Piracicaba. Immediately, the roots were briefly shaken to separate the bulk from rhizosphere soil. The soil that remained attached to the roots was considered as rhizosphere soil. The rhizosphere soil was extracted from the roots with the aid of a sterilized brush. Soil samples from the control pots—with no-plant—were considered as bulk soil.

After harvest, portions of soil (500g) and cover straw (200g) from each pot were separated and collected for physicochemical analysis, which was performed at the Soil Chemistry Laboratory, University of São Paulo, Piracicaba, Brazil, according to routine methodology (CLAESSEN et al., 1997; GEE; BAUDER, 1986; KEENEY; NELSON, 1982; TEDESCO et al., 1995) (For more details, see section 2.2.2). Soil enzymatic activity analysis (MELO et al., 2010) was performed in the Biogeochemistry Laboratory, São Paulo State University, Jaboticabal, Brazil.

3.2.3 Metagenomic DNA extraction and high-throughput sequencing

For each sample, metagenomic DNA extraction from 250 mg of soil was performed, using PowerLyzer PowerSoil[®] DNA Isolation Kit (Mo Bio Laboratories, Carlsbad, USA), according to the manufacturer's instructions. DNA quality was verified in 1% Tris-saline buffer (TBS) with sodium boric acid (BRODY; KERN, 2004) agarose gel electrophoresis and NanoDrop 2000[®] spectrophotometer (Thermo Scientific, Waltham, USA), respectively. The quantity of product was calculated on the Qubit[®] fluorometer (Thermo Fischer Scientific, Waltham, USA).

The soil metagenomic libraries were constructed with the Nextera DNA Sample Preparation Kit[®] (Illumina Inc., San Diego, USA), as described by the manufacturer. A total of 36 barcoded samples were obtained, with the Nextera[®] Index Kit (Illumina Inc., San Diego, USA). The product was purified with the Agencourt[®] AMPure[®] XP reagents (Beckman Coulter, Brea, USA). Quantitative PCR of the pooled samples was performed with the KAPA SYBR[®] FAST Universal qPCR Kit (Kappa Biosystems, Woburn, USA) in a StepOne Plus[®] thermocycler (Thermo Fischer Scientific, Waltham, USA). Each reaction consisted of 4 µl of Master Mix, 2 µl of primers, 2 µl of sterile water and 2 µl of DNA diluted 100, 1000 and 2000 times. A standard curve with 6 known DNA concentrations, ranging from 0.0002 pM to 20 pM (provided by the KAPA kit). The high throughput sequencing was performed with MiSeq Reagent Kit V2[®] in a MiSeq[®] Desktop Sequencer (Illumina Inc., San Diego, USA).

The sequencing consisted of a paired end 251-cycle/39-hour run. (All the other steps followed the ones described in the section 2.2.3).

3.2.4 Annotation and data mining of metagenome libraries

Unassembled fastq DNA sequences were annotated with the Metagenomics Rapid Annotation (MG-RAST) pipeline version 3.3.9 (MEYER et al., 2008). Normalized taxonomic profiles were generated using the best hit match criterion from M5nr database (WILKE et al., 2012), while normalized functional profiles were generated from SEED database (OVERBEEK et al., 2013). In order to standardize and remove biases from differences in read lengths and sequencing efforts, matrices of taxonomic- and functional-relative abundances were generated dividing the number of reads of each taxa or functional subsystem by the total number of reads in each sample. The identification of hits was performed in BlastX with a minimum alignment of 50 bp and E-value $< 1 \times 10^{-5}$ as cutoff (MENDES et al., 2014; SMITH et al., 2012).

Shannon's alpha-diversity and Whittaker global beta-diversity were calculated based on taxonomic relative abundance matrix at genus level and functional relative abundance matrix at subsystem Level 2, using the PAST software, version 3 (HAMMER; HARPER; RYAN, 2001). Nonmetric multidimensional scaling (NMDS) plots of Euclidean distance matrices, were used to visualize the structure of community among samples, using the above described taxonomic and functional abundance matrices, on Canoco software, version 5 (LEPŠ; ŠMILAUER, 2005).

All statistical analyses were performed using the following design: 6 samples, two environments (bulk soil and rhizosphere) and three periods along the chronosequence (1-, 10- and 20-year). Abundance data from taxa (M5nr database) and functional subsystems (SEED database) of each soil metagenome were downloaded from MG-RAST and used to generate taxonomic and functional matrices as input for the Statistical Analysis of Metagenomic Profiles (STAMP) software, version 3.0 (PARKS; BEIKO, 2010). The significance across bulk soil and rhizosphere and along the chronosequence in each environment was calculated using two-sided Welch's t-test q-values (WELCH, 1934). The confidence intervals were calculated using the Newcombe–Wilson method (NEWCOMBE, 1998), corrected by the Benjamini–Hochberg false discovery rate (FDR) (BENJAMINI; HOCHBERG, 1995).

In order to identify the major environmental drivers of microbial patterns, a Distance-based Redundancy Analysis (db-RDA) of Euclidean distance matrices, with forward selection

was performed. Furthermore, a Variation Partitioning of Redundancy Analysis (p-RDA), generated by Principal Coordinates Analysis of Neighbor Matrices (PCNM) with forward selection was performed to evaluate the possible single and combined effects of environmental variables on the variation of microbial taxonomic and functional profiles. Latitude and longitude were used as constraining spatial coordinates. Both analyses were performed using the Canoco software, version 5.

To depict the correlations within taxonomic groups and functional categories, their interactions, and their connectivity with environmental factors, all possible rank pairwise Spearman's correlations were computed. The analysis was performed on the Vegan package from the R software, version 3.1.2 (R Development Core Team, 2008), with the Benjamini-Hochberg FDR used to correct the significance values. In order to visualize the interactions among taxonomy, functions and environmental variables, we inputted the pairwise correlations on the Gephi network exploration platform, version 0.8.2 (BASTIAN; HEYMANN; JACOMY, 2009).

To test which paradigm better described general assembly patterns of the microbial communities in bulk soil and rhizosphere along the no-till chronosequence, we calculated rank abundance distributions and immigration rates, based on taxonomic genus level matrices, for each sample, from each land-use system. The species rank abundance for each metagenomic sample were fitted to five different models: the zero sum binomial model and the broken stick (null model), which regard to neutral assembly and the pre-emption, the log-normal and the Zipf models, regarding to niche-based assembly. The broken stick, pre-emption, log-normal and Zipf models were calculated using the script 'radfit' from Vegan package on R software, version 3.1.2. The ZSM model as well as the dispersal rates for each sample were calculated on TeTame software (JABOT; ETIENNE; CHAVE, 2008), version 1.9.. The models were compared based on the Akaike Information Criterion (AIC) — the lower the value, the best fitted was the model.

3.3 Results

3.3.1 Sampling site attributes

In order to understand the contribution of selection process on explaining microbial assembly patterns, we measured or calculated 53 environmental variables, from soil physicochemical attributes (29), to straw characteristics (19) and soil microbial enzymes

activity (5). Follow the main environmental characters that changed along the chronosequence (Table 3.1). For details of the 53 variables evaluated, see supplementary information (Appendix B, Supplementary tables ST 3.1, ST 3.2 and ST 3.3).

Table 3.1 - Environmental variables that have significantly changed along the chronosequence. Tukey HSD test ($P < 0.05$). Letters represent significant differences. Standard deviation is shown.

Variable	Chronosequence	1-year	10-year	20-year
<i>Soil physicochemical</i>				
Cations saturation		47.2±8.1b	82.8±5.2a	79.8±2.5a
Ca ²⁺		15.3±3.0c	29.0±5.2b	35.7±3.4a
Mg ²⁺		9.2±1.8b	19.8±3.4a	17.7±2.0a
N-total		1,311.3±196.3cb	1,479.3±237.4b	1,564.5±144.6a
P		14.0±7.4b	13.2±9.0b	37.7±23.2a
pH _{KCl}		5.0±0.2b	6.0±0.1a	5.8±0.2a
Potential acidity		28.7±5.8a	10.7±4.5b	13.8±2.9b
SOC		20.5±5.4b	23.0±2.0b	26.5±3.3a
SOM		35.5±9.4b	39.5±3.0b	46.0±5.4a
Clay		298.2±19.7b	293.5±25.6b	321.0±35.3a
<i>Soil enzymatic activity</i>				
Acid phosphatase		85.8±8.4a	63.4±8.4b	55.4±8.3c
β-glycosidase		17.7±6.0c	36.8±8.5b	116.0±18.6a
Dehydrogenase		2.4±0.8b	5.1±1.7a	5.6±1.5a
FDA		83.6±11.9b	107.9±19.5a	85.3±12.0b
Urease		9.4±4.2b	18.4±4.7a	9.0±2.4b
<i>Litter</i>				
Straw P-total		0.17±0.05b	0.18±0.02ab	0.21±0.04a
Straw TOC		24.3±1.0b	23.4±4.4ab	29.4±10.6a

All the physicochemical variables that significantly changed along the chronosequence, beside potential acidity, have their values increased. All the enzymes changed their activities along the chronosequence. Acid phosphatase activity decreased along the chronosequence. The activities of beta glycosidase and dehydrogenase increased along time. FDA and urease had higher activities in 10-year chronosequence samples. From all litter variables, only P-total and total organic carbon changed, with increasing values along the chronosequence.

3.3.2 Microbial patterns in the soil-rhizosphere interface along the chronosequence

A total of 10.7 million of 250-bp sequences was obtained by high throughput shotgun metagenomics. The sequences are deposited in MG-RAST, ID = 7830. In order to have a comprehensive picture of the patterns of diversity, abundance and composition of microbial communities along the chronosequence in the interface soil-rhizosphere, we used a set of

multivariate statistical approaches. The taxonomic and functional structure of the microbial communities was visualized by 2-dimension NMDS plots, obtained from Euclidean-distance matrices, using Canoco software, version 5 (Figure 3.2).

Figure 3.2 - Non-metric multidimensional scaling (NMDS) plot of Principal Coordinates Analysis (PCoA), generated by Euclidean distance matrices, with 100 permutations. (a) Taxonomic structure at genus level and (b) functional structure at subsystems level 2, based on shotgun metagenomics data from M5nr database and SEED subsystems database, respectively. Samples were grouped according to soil fraction and time — bulk soil, from dark to light brown and rhizosphere from dark to light green, the extent the chronosequence advanced. In (a) vertical additional axis represented Shannon's alpha-diversity and horizontal additional axis represented Whittaker's Global beta-diversity. Taxonomy was compared through Tukey HSD, $P < 0.05$. Lines connecting samples represented the distance to the nearest neighbor, with Donnelly edge correction

The taxonomic NMDS plot at genus level revealed clear separation of microbial communities from bulk soil and rhizosphere (Figure 3.2a). The communities also evolved along time of no-till cropping, with clear separation of samples from 1-year and the others, in both bulk soil and rhizosphere. Indeed, microbial communities became closely related, the extent the chronosequence advanced, with the 10- and 20-year samples grouped each other in both bulk soil and rhizosphere. Despite of that, the samples within these two fractions did not overlap each other across the chronosequence. The taxonomic variation from the whole set of samples was 41%. The functional categories at subsystems level became closely related the extent the chronosequence advanced (Figure 3.2b). Significant functional differences in the structure of bulk soil and rhizosphere were found only at 1-year. However, the magnitude of the differences in functional structure was meaningless, since the variation from the whole set of samples was 4%.

The Shannon's α -diversity—average total diversity of samples in a certain environment—did not vary across bulk soil and rhizosphere, but across the chronosequence, with samples from 1-year less diverse than samples from 10- and 20-year. The Whittaker's β -diversity—the global diversity shared between different samples of a certain environment—decreased along the chronosequence, both in bulk soil and rhizosphere. It means that, along time, the communities in the same fraction became more similar in both sides of the interface. In summary, we can deduce that the extent the chronosequence advanced more abundant taxa emerged in both bulk soil and rhizosphere samples, but counterbalanced with more taxa been shared in each one of these soil fractions.

The possible shifts in composition were also investigated (Figure 3.3), through the taxonomic metagenomic profiles using the STAMP software, version 2.

Figure 3.3 - Relative abundance and difference in mean proportions of soil bacterial and archaeal communities at taxonomic phylum level, based on shotgun metagenomics row normalized data from M5nr database. Error bars show standard deviation of 6 samples in each soil fraction and time. Only phyla with average relative abundance higher than 0.1% were accounted. (a) Relative abundance of phyla, in bulk soil (brown) and rhizosphere (green) along the no-till chronosequence. (b) Welch's t-test with corrected q-values calculated using the Benjamini-Hochberg false discovery rate approach (FDR). Only the phyla that significantly shifted after long-term land-use conversion were shown (q-value < 0.05). Samples were grouped according to soil fraction and time — bulk soil (from dark to light brown) and rhizosphere (from dark to light green), the extent the chronosequence advanced. (c) Relative abundance of classes inside the phylum Proteobacteria, in bulk soil and rhizosphere, along the no-till chronosequence

Almost the phyla that significantly shifted, presented lower relative abundance in rhizosphere than in the bulk soil, along the chronosequence (Figure 3.3a). Exception made to Proteobacteria and Bacteroidetes, which had higher abundance in rhizosphere 1-year. We also

evaluated the microbial dynamics along the chronosequence (Figure 3.3b). The community dynamics of bulk soil differed from the dynamics of the rhizosphere. Along the chronosequence in the bulk soil, we noticed that the populations of Acidobacteria and Firmicutes significantly decreased after 10-year no-till cropping. A significant decrease on Acidobacteria also observed after 20-year no-till, besides the decrease of Actinobacteria relative abundance. Despite of that, several phyla significantly enlarged their populations along the bulk soil chronosequence, with highlights to Proteobacteria, Planctomycetes, and Gemmatimonadetes, which gradually enhanced from 1- to 10- and from 1- to 20-year.

Interestingly, the number of phyla that significantly changed on rhizosphere fraction gradually increased along the chronosequence, with shift in nine, 12 and 16 groups, from 1- to 10-year, 10- to 20-year and 1- to 20-year, respectively. More than that, the direction of the evolution was always pointed to the advance of the chronosequence, what means that almost the groups enlarged their populations on rhizosphere the extent the chronosequence advanced, with exception to Acidobacteria and Proteobacteria. In summary, almost the taxa have significantly changed, evolved positively the extent the chronosequence advanced. The main taxonomic shifts in the bulk soil occurred in the first 10 years, while the shifts in the rhizosphere increased in the last 10 years (10- to 20year), computing largely to the bigger differences found in general (1- to 20-year), compared to the first 10 years (1- 10-year).

Depicting the variability of the phylum Proteobacteria at the class level (Figure 3c), we noticed that Alpha- and Betaproteobacteria presented higher abundance in rhizosphere, while Delta- and Gammaproteobacteria decreased. Looking to the microbial evolution dynamics, the relative abundance of Alphaproteobacteria reduced along the chronosequence, while the relative abundance of Beta-, Delta- and Gammaproteobacteria increased with time, in both bulk soil and rhizosphere. Despite of that, Alphaproteobacteria was always the most abundant class inside Proteobacteria, apart from time or soil fraction — 53.7% of sequences inside Proteobacteria and 23.5% of the total number of sequences—followed by Betaproteobacteria (8.2% and 18.8%, respectively).

We also investigated the changes in microbial functional categories along the chronosequence in bulk soil and rhizosphere (Figure 3.4).

Figure 3.4 - Relative abundance and difference in mean proportions of soil functional categories, at subsystems level 2, based on shotgun metagenomics raw normalized data from SEED database. Error bars show standard deviation of 6 samples in each soil fraction and time. (a) Relative abundance of all functional categories, in bulk soil (brown) and rhizosphere (green) along the no-till chronosequence. (b) Welch's t-test with corrected q-values calculated using the Benjamini-Hochberg false discovery rate approach (FDR). Only the functions that significantly shifted after long-term land-use conversion were shown ($q\text{-value} < 0.05$). Samples were grouped according to soil fraction and time — bulk soil (from dark to light brown) and rhizosphere (from dark to light green), the extent the chronosequence advanced

No significant changes were found in any category, when comparing the relative abundance of bulk soil and rhizosphere metagenomic profiles, but for 'stress response', that was higher in rhizosphere 1-year. Along the chronosequence, several functional shifted in the bulk soil, but just a few in rhizosphere. The magnitude and evolution dynamics of the functional changes were also different. In the bulk soil, more functional categories shifted from the 1- to 10-year (11), with just a few from 10- to 20-year (3), resulting in a decrease of functional categories that changed along the whole chronosequence (10), compared to the first 10 years of no-till cropping. Functional categories such as 'aminoacids and derivatives' and 'nucleosides and nucleotides' increased along the chronosequence, while 'respiration', 'virulence, disease and defense' and 'potassium metabolism' decreased. In the rhizosphere, just a few functional categories shifted along the chronosequence. Similar to bulk soil, the

relative abundance of sequences related to ‘respiration’, ‘virulence, disease and defense’ and ‘potassium metabolism’ decreased after 20-year no-till cropping. Interestingly, no differences in functional categories were found between 1- and 10-year rhizosphere. In summary, the microbial functions in bulk soil evolved while the functions in rhizosphere were kept, the extent the chronosequence advanced.

3.3.3 Relationship between soil microbial patterns and environmental factors

In order to enumerate the main environmental drivers of the shifts in assembly and functional categories along the chronosequence and, between bulk soil and rhizosphere, we performed the Euclidean distance-based Redundancy Analysis (db-RDA) (Figure 3.5a), with stepwise forward selection of the 53 explanatory variables measured. The analyses were performed at genus level and functional subsystems level 2, to have a deeper and more comprehensive picture of the main factors driving microbial assembly and functional categories, as well as their relative contribution to the patterns found.

Figure 3.5 - Environmental factors modulating soil microbial taxonomic assembly in bulk soil and rhizosphere, along the chronosequence. (a) Distance based Redundancy Analysis (db-RDA), with forward selection of explanatory variables generated from Euclidean distance matrices, with 1000 Monte-Carlo permutations and corrected by FDR. Samples were grouped according to soil fraction and time — bulk soil (from dark to light brown) and rhizosphere (from dark to light green), the extent the chronosequence advanced. The cations saturation (V%) explained 31.5% (pseudo-F = 15.6; $P_{\text{corrected}} = 0.001^{***}$) of the microbial variation due the environment, followed by aluminum concentration (4%; pseudo-F = 2.0; $P_{\text{corrected}} = 0.045^{*}$). (b) Variation partitioning of RDA generated by Principal Coordinates Analysis of Neighbor Matrices (PCNM), with forward selection and FDR correction. Test of the simple effect of physicochemical variables, litter and soil enzymes, and their combined effect on the modulation of microbial dynamics

The forward-selected variables explained together 35.5% of the total variation in soil microbial taxonomic composition, being 33.0% of the explanation on the first axis of RDA. The cations saturation (V%) was the main driver of the differences on community composition, explaining 31.5% of the total variation (pseudo-F = 15.6; $P_{\text{adjusted}} < 0.001^{***}$), followed by aluminum concentration, which explained 4.0% (pseudo-F = 2.0; $P_{\text{adjusted}} = 0.044^*$). The increment on cations saturation in soil solution along the years, and consequent decreasing of Al^{3+} concentration, enabled samples of 1-year to group each other separately from the rest of the chronosequence. Despite of that, those environmental factors were not able to separate samples from bulk soil and rhizosphere. The total variation across samples due to environmental factors was 24.6% (Figure 3.6a).

Figure 3.6 - Environmental factors modulating soil microbial functional structure in bulk soil and rhizosphere, along the chronosequence. (a) Distance based Redundancy Analysis (db-RDA), with forward selection of explanatory variables generated from Euclidean distance matrices, with 1000 Monte-Carlo permutations and corrected by FDR. Samples were grouped according to soil fraction and time — bulk soil (from dark to light brown) and rhizosphere (from dark to light green), the extent the chronosequence advanced. The cations saturation (V%) explained 11.1% (pseudo-F = 4.3; $P_{\text{corrected}} = 0.001^{***}$) of the functional variation due the environment, followed by aluminum saturation (m%) (5.6%; pseudo-F = 2.2; $P_{\text{corrected}} = 0.019^*$). (b) Variation partitioning of RDA generated by Principal Coordinates Analysis of Neighbor Matrices (PCNM), with forward selection and FDR correction. Test of the simple effect of physicochemical variables, litter and soil enzymes, and their combined effect on the modulation of microbial dynamics

From the total variation on soil functional subsystems structure, the forward-selected explanatory variables account for 16.8%. The cations saturation was the major driver of the variation and explained 11.1% of the total variation (pseudo-F = 4.3; $P_{\text{adjusted}} < 0.001^{***}$),

followed by aluminum saturation with 5.6% (pseudo-F = 2.2; $P_{\text{adjusted}} = 0.019^*$). Despite of that, the total variation across samples due to environmental factors was approximately 0.2%, making negligible their contribution for explaining the functional patterns.

We found different patterns of taxonomic diversity, composition and abundance along the chronosequence and between bulk soil and rhizosphere. Despite of that, the magnitude in which these patterns have change did not follow a systematic gradient. An abrupt change in microbial diversity occurred from 1- to 10-year, in both bulk soil and rhizosphere. The structure of taxa also shifted and became more constraining in 10-year, regardless the soil fraction. The abundance of taxa in the bulk soil shifted mainly from 1- to 10-year while the abundance in rhizosphere shifted progressively along the chronosequence. In general, bulk soil and rhizosphere became more closely related after 10-year of no-till cropping and stabilized from 10- to 20-year.

Based on these patterns we suppose that different drivers were modulating the microbial assembly and functions on the two different soil fractions evaluated and also along the chronosequence. We argued also, that the relative weight of these drivers on modulating diversity, composition and abundance could change along soil fraction and time as well. To underlie these patterns we performed a Variation partitioning Redundancy Analysis (pRDA) (Figure 3.5b and 3.6b). We found that the contribution of the environmental variables to explain the variation on the taxonomic and functional structure of microbial community was lower in rhizosphere than in bulk soil. For taxonomy, the contribution of soil physicochemical variables decreased along the chronosequence in bulk soil in increased in rhizosphere. The influence of straw was increased the extent the chronosequence advanced in the bulk soil and decreased in the rhizosphere. Despite of that, we found a combined effect of soil physicochemistry and straw in 10- and 20-year.

Depicting the taxonomic variability along the chronosequence and between bulk soil and rhizosphere through the PCNM (Table 3.2), we found that the main drivers of taxonomic dynamics along the chronosequence and fractions were: soil organic carbon and beta-glycosidase activity in the bulk soil 1-year; Zn^{2+} , longitude and acid phosphatase activity in bulk 10-year; moisture, Mg^{2+} and straw organic matter in bulk 20-year; P concentration, N-NH_4^+ and beta-glycosidase activity in rhizosphere 1-year; Mg^{2+} and dehydrogenase activity in rhizosphere 10-year, and; N-NO_3^- and straw N-total in rhizosphere 20-year.

Table 3.2 - Correlation between microbial phyla and environmental variables, along the chronosequence and soil fractions. Values extracted from the PCNM analysis

Area	Fraction	F-value	P-value	Variables	Positive (+) or negative (-) correlation with taxonomic groups at phylum level
<i>Bulk soil</i>					
1-year	a+b+c	5.2	0.022*	SOC	Actinobacteria (+), α -Proteobacteria (-), γ -Proteobacteria (-), Firmicutes (-)
	a+c	4.6	0.010*		
	b+c	2.8	0.150 ^{NS}	β -glycosidase	α -Proteobacteria (+), γ -Proteobacteria (+), δ -Proteobacteria (+), Actinobacteria (-)
10-year	a+b+c	11.1	0.019*	Zn, Longitude	α -Proteobacteria (+), β -Proteobacteria (+), γ -Proteobacteria (+), Actinobacteria (-)
	a+c	12.7	0.022*		
	b+c	9.4	0.008**	Acid phosphatase	Actinobacteria (+), α -Proteobacteria (-), β -Proteobacteria (-), γ -Proteobacteria (-), δ -Proteobacteria (-)
20-year	a+b+c	3.6	0.115 ^{NS}	Moisture	Firmicutes (+), Actinobacteria (+), α -Proteobacteria (-)
	a+c	4.9	0.027*	Mg ²⁺	δ -Proteobacteria (+), Acidobacteria (+), Actinobacteria (-)
	b+c	4.6	0.066 ^{NS}	Straw RCOM	α -Proteobacteria (+), β -Proteobacteria (+), Verrucomicrobia (+), δ -Proteobacteria (+), Actinobacteria (-), Firmicutes (-)
<i>Rhizosphere</i>					
1-year	a+b+c	4.3	0.035*	P	δ -Proteobacteria (+), Planctomycetes (+), Actinobacteria (-)
	a+c	6.1	0.009**	N-NH ₄ ⁺	α -Proteobacteria (+), γ -Proteobacteria (+), Firmicutes (-)
	b+c	5.7	0.011*	β -glycosidase	δ -Proteobacteria (+), Acidobacteria (+), Firmicutes (+), β -Proteobacteria (-)
10-year	a+b+c	3.7	0.084	Mg ²⁺	Actinobacteria (+), δ -Proteobacteria (-)
	a+c	4.3	0.059		
	b+c	2.1	0.189	Dehydrogenase	β -Proteobacteria (-)
20-year	a+b+c	4	0.021*	N-NO ₃ ⁻	Actinobacteria (+), α -Proteobacteria (+), Cyanobacteria (+), β -Proteobacteria (-), γ -Proteobacteria (-), Acidobacteria (-), Planctomycetes (-)
	a+c	3.6	0.033*		
	b+c	2	0.200 ^{NS}	Straw N-total	Acidobacteria (+), δ -Proteobacteria (+), α -Proteobacteria (-), Actinobacteria (-), Cyanobacteria (-)

a = soil physicochemical factors
b = overlap
c = enzymes + litter factors

3.3.4 Three-dimension dynamics of the soil microbial community along the chronosequence

We were interested in visualize the evolution dynamics of the interactions among the three dimension of ecological patterns—taxonomy, functions and environmental variables—which interact each other through ecological processes. To achieve that, we calculated all the possible pairwise correlations along these three dimensions and generated visual networks of the significant and strong connections (Figure 3.7). We also calculated the topological properties of each network (Table 3.1).

Figure 3.7 - Network evolution dynamics, based on Spearman's pairwise correlation matrices. Taxonomic phylum level (blue nodes), except for Proteobacteria at class level, obtained from M5nr database. Functional categories at level 1 (red nodes), from SEED database. Environmental variables (green nodes): soil physicochemical, litter and enzymes. The size of each node is proportional to its number of connections (degree). Each edge is colored according to the target, being that blue edges represent taxonomy-taxonomy connections, red edges represent taxonomy-function or function-function connections and green edges present taxonomy-environment or function-environment connections. Each edge represents a strong (Spearman's $\rho \leq -0.7$, regular edges; Spearman's $\rho \geq 0.7$, bold edges) and significant ($P < 0.05$) correlation

Table 3.2 - Topological characteristics of the networks, along the chronosequence and soil fractions

Network properties	Bulk soil			Rhizosphere		
	1-year	10-year	20-year	1-year	10-year	20-year
Number of nodes ^a	93	99	102	100	95	92
Number of edges ^b	156 (24)	269 (39)	223 (45)	190 (50)	228 (41)	217 (33)
Number of positive edges ^c	89 (13)	134 (18)	89 (15)	103 (23)	114 (16)	122 (15)
Number of negative edges ^d	67 (11)	135 (21)	134 (30)	87 (27)	114 (25)	95 (18)
Network diameter ^e	4	4	3	4	5	5
Modularity ^f	0.69	0.61	0.58	0.66	0.63	0.64
Average path length ^g	1.42	1.67	1.29	1.40	1.59	1.51
Average clustering coefficient ^h	0.16	0.17	0.12	0.14	0.15	0.18
Average degree ⁱ	3.51	5.60	4.46	3.92	4.90	4.88

a Number of features with at least one strong correlation (Spearman's $\rho \geq |0.7|$ and statistically high significant ($P < 0.05$). The values in parentheses show higher significance correlations ($P < 0.01$).

b Number of connections obtained by pairwise correlations between nodes.

c Spearman's pairwise positive and significant correlations ($\rho \geq 0.7$; $P < 0.05$).

d Spearman's pairwise negative and significant correlations ($\rho \leq -0.7$; $P < 0.05$).

e The longest distance between nodes in the network, measured in number of edges.

f The capability of the nodes to form a modular structure, that is, a structure with high density of nodes and clustered topology. Values > 0.4 suggest that the network has a modular structure.

g Average network distance between all pair of nodes or the average length of all edges in the network.

h How nodes are embedded in their neighborhood and the degree to which they tend to cluster together.

i The average number of connections per node in the network, that is, the node connectivity.

All the networks presented modular structure (>0.4) (NEWMAN, 2003). The diameter of the networks decreased on 20-year bulk soil compared to the other years. Differently the diameter increased in rhizosphere from 1- to 10-year. The clustering among neighbors in the network—which is average clustering coefficient—decreased in bulk soil and increased in rhizosphere, the extent the chronosequence advanced. The number of active features increased along the chronosequence in bulk soil and reduced in rhizosphere. Despite of that, the number of strong and significant connections (Spearman's $\rho \geq |0.7|$; $P < 0.05$) increased along the chronosequence in both soil fractions. Interestingly, the number of strong and very significant connections ($P < 0.01$) behaved differently, increasing in bulk soil and decreasing in rhizosphere along the chronosequence. The number of positive connections increased from 1- to 10-year and decreased from 10- to 20-year in bulk soil while in the rhizosphere these connections gradually increased along the chronosequence. The number of negative connections in bulk soil increased from 1- to 10-year and remained stable in 20-year. In the rhizosphere this number was higher in 10-year. The number of very significant negative connections increased constantly along the chronosequence in bulk soil and had right the opposite dynamics in rhizosphere.

The taxonomic groups that present more number of connections in bulk soil were: Alpha-, Beta-, Gamma- and Epsilon-proteobacteria in 1-year; Nitrospirae, Actinobacteria, Spirochaetes, Gamma-, Beta- and Alpha-proteobacteria in 10-year, and; Chloroflexi, Chlorobi and Delta-proteobacteria in 20-year no-till cropping. The most interactive taxonomic groups in the rhizosphere were: Deinococcus-Thermus, Chlorobi, Synergistetes and Actinobacteria in 1-year; Gemmatimonadetes, Chrenarchaeota and Verrucomicrobia in 10-year, and; Beta- and Delta-proteobacteria in 20-year no-till.

The functional categories with higher number of connections in bulk soil were: ‘cofactors, vitamins, prosthetic groups and pigments’ in 1-year; ‘cell wall and capsule’, ‘virulence and disease defense’ and ‘cell division and cell cycle’ in 10-year, and; ‘protein metabolism’, ‘stress response’, ‘photosynthesis’ and ‘sulfur metabolism’ in 20-year no-till cropping. The functional categories over-connected in rhizosphere were: ‘dormancy and sporulation’ and ‘RNA metabolism’ in 1-year; ‘phosphorus metabolism’ and ‘sulfur metabolism’ in 10-year, and; ‘regulation and cell signaling’, ‘membrane transport’ and ‘phosphorus metabolism’ in 20-year no-till cropping.

The main environmental factors modulating the network configuration in bulk soil were: SOM and soil Al saturation in 1-year; soil Mg^{2+} concentration, soil pH and straw pH in 10-year, and; straw P-total in 20-year. In the rhizosphere, higher number of connections were found to: straw Mg-total in 1-year; straw K-total and soil Zn concentration in 10-year, and; soil Al saturation and straw Mg-total in 20-year.

3.3.5 Microbial community assembly dynamics along the chronosequence

Microbial interactions with functions and environmental drivers have changed between bulk soil and rhizosphere, with different drivers and different weights. Also, the three-dimensional dynamics in the distinct soil fractions evolved through different patterns along the chronosequence. Thus, we hypothesized that microbial assembly dynamics should be different in bulk soil and rhizosphere across time. Besides that, the theoretical modeling of microbial assembly could bring to light the ecological processes ruling microbial evolution dynamics in the no-till cropping chronosequence. To confirm this hypothesis, we fitted all samples to theoretical abundance distribution models (Table 3.3).

Table 3.3 - Samples fitting to theoretical ecological models, based on Akaike Information Criterion (AIC) for rank abundance distribution at genus level, based on M5nr database. The tendency to migration was also calculated

Sample	Source	Time	Dispersal Rate ^b	Akaike information Criterion ^a				
				Neutral		Niche-based		
				Broken stick	ZSM	Lognormal	Preemption	Zipf
bPA1	Bulk soil	1-year	0.33	1874.4	1629.0	1633.0	1782.6	1716.7
bPA2	Bulk soil	1-year	0.46	1701.5	1562.3	1549.2	1646.7	1612.6
bPB1	Bulk soil	1-year	0.63	2065.7	1708.2	1717.9	1925.8	1857.2
bPB2	Bulk soil	1-year	0.50	1521.8	1468.8	1450.9	1498.4	1478.4
bPC1	Bulk soil	1-year	0.52	1759.9	1599.7	1593.8	1694.2	1663.6
bPC2	Bulk soil	1-year	0.60	1561.7	1490.0	1473.3	1530.9	1507.8
bSA1	Bulk soil	10-year	0.33	2013.1	1700.1	1710.8	1880.9	1865.5
bSA2	Bulk soil	10-year	0.09	1966.6	1698.4	1711.9	1839.4	1862.8
bSB1	Bulk soil	10-year	0.52	2094.1	1716.0	1739.6	1937.3	1916.2
bSB2	Bulk soil	10-year	0.28	3293.2	2011.3	2262.0	2662.0	2767.5
bSC1	Bulk soil	10-year	0.37	1963.2	1693.8	1699.5	1847.5	1837.2
bSC2	Bulk soil	10-year	0.31	1902.7	1678.4	1674.3	1805.7	1792.4
bTA1	Bulk soil	20-year	0.47	2084.0	1745.5	1770.1	1937.3	1938.5
bTA2	Bulk soil	20-year	0.46	1776.0	1612.8	1607.7	1707.0	1699.7
bTB1	Bulk soil	20-year	0.32	1775.9	1617.5	1605.2	1708.1	1687.4
bTB2	Bulk soil	20-year	0.30	1966.6	1692.4	1694.1	1851.9	1830.1
bTC1	Bulk soil	20-year	0.36	2560.2	1899.8	1996.5	2250.6	2283.1
bTC2	Bulk soil	20-year	0.53	1631.7	1544.5	1529.1	1594.2	1580.6
rPA1	Rhizosphere	1-year	1	1452.9	1402.1	1387.6	1432.3	1406.1
rPA2	Rhizosphere	1-year	1	1402.3	1371.5	1358.4	1386.1	1373.9
rPB1	Rhizosphere	1-year	1	1648.0	1529.7	1522.6	1597.7	1570.6
rPB2	Rhizosphere	1-year	1	1176.1	1178.8	1163.8	1173.3	1167.1
rPC1	Rhizosphere	1-year	1	1230.4	1226.2	1211.7	1226.0	1217.1
rPC2	Rhizosphere	1-year	1	1264.8	1253.3	1237.9	1257.6	1245.0
rSA1	Rhizosphere	10-year	0.73	1726.3	1586.4	1571.6	1664.5	1649.9
rSA2	Rhizosphere	10-year	0.49	1979.8	1695.9	1704.8	1852.8	1833.4
rSB1	Rhizosphere	10-year	0.50	1644.0	1549.7	1538.2	1594.6	1594.7
rSB2	Rhizosphere	10-year	0.78	2023.3	1720.2	1736.5	1887.4	1874.8
rSC1	Rhizosphere	10-year	0.40	2323.6	1798.3	1838.2	2096.3	2060.9
rSC2	Rhizosphere	10-year	0.27	1889.0	1668.2	1665.9	1784.3	1775.2
rTA1	Rhizosphere	20-year	0.63	1906.0	1668.8	1674.0	1804.0	1790.8
rTA2	Rhizosphere	20-year	0.21	1962.0	1687.4	1686.2	1843.8	1813.3
rTB1	Rhizosphere	20-year	0.23	2005.2	1697.2	1703.4	1875.2	1847.3
rTB2	Rhizosphere	20-year	0.85	1728.0	1585.0	1576.8	1667.0	1650.4
rTC1	Rhizosphere	20-year	0.67	1612.4	1532.9	1519.7	1575.7	1566.7
rTC2	Rhizosphere	20-year	0.79	1596.9	1524.1	1508.9	1563.7	1549.0

^a Rank abundance model based on corrected AIC value from Poisson distributions using maximum likelihood estimation. The lowest AIC value for each sample represented the best-fitted model for general community's assembly. AIC corrected values were calculated by the Radfit function on Vegan R Package, with exception of zero-sum model distribution, which was calculated using TeTame Software. Best-fitted models were calculated by the general equation $AIC = -2\log\text{-likelihood} + 2 \times \text{npar}$. ^b Dispersal rate were calculated by Etienne's formula, using TeTame Software. Values of dispersal are between 0 and 1, where 0 means no tendency to migration and 1 means total tendency to migration in a certain community. Values ≥ 0.5 were highlighted in bold. Bold emphasizes the best-fitted model in each metagenomic sample. ZSM - Zero Sum Multinomial.

In bulk soil, 10 samples fitted to neutral model and eight to niche-based, while in rhizosphere six fitted neutral against 12 niche-based. The differences along the chronosequence were even more evident. In the 1-year all the samples of rhizosphere and four from six samples of bulk soil fitted niche-based. In 10-year almost the samples of bulk soil fitted neutral model. In rhizosphere 10- and 20-year as well as in bulk soil 20-year, the samples fitted almost equally to both neutral and niche-based models. The samples of bulk soil presented, in general, low to intermediate rates of dispersal, while samples of rhizosphere presented high rates of dispersal, which means more predisposition to migration from members of rhizosphere community, compared to those from bulk soil community.

3.4 Discussion

3.4.1 Ecological processes governing microbial community assembly

Several studies show that the community composition is strongly related to environmental conditions (ATTARD et al., 2011; FIERER; JACKSON, 2006), which is indicative of selection. In our results, we found that environmental factors acted as selective force, shifting the composition of the microbiome selected by the rhizosphere, and also enhanced the microbial trade-off dynamics into the fractions, along the chronosequence. But, the relative weight of selection of microbiome by plant roots varied across time, the same as found by MENDES et al. (2014). In addition, our results suggested that 20-year no-till cropping was enough to lead microbial community to a dynamic stable state, in both soil fractions evaluated, reducing the role of the plant on trade-offs between bulk soil and rhizosphere.

Since the communities evolved to different holders and both became more likely each other, we believe that they reached an irreversible dynamic steady state, unless dramatic changes in soil management occur (INSAM; DOMSCH, 1988). This dynamic stability is evidenced by the fact that, despite changes in microbial abundance and composition we did not find changes in the abundance and composition of functional categories between soil fractions, along the chronosequence. Thus, we deduce that the trade-off in microbial taxa in the rhizosphere, by selection, was compensated by functional resilience force (ALLISON; MARTINY, 2008), leading to assembly predominantly based on niche. It leads to functional redundancy along the chronosequence, evidencing the functional resilience in the rhizosphere, apart from selection.

Despite the increasing literature regarding the role of selection, dispersal, drift and speciation on soil microbial community dynamics, a common framework explaining the role of these processes and their possible interactions is lacking (STEGEN et al., 2013). Most of the previous studies provide are based on frameworks proposed for macro-ecology (VELLEND, 2010). Furthermore, a few studies have evaluated and discussed how ecological processes modulate the microbial communities along time (DINI-ANDREOTE et al., 2014) and more specifically, how these processes interact to explain microbial trade-offs in the soil-rhizosphere interface (MENDES et al., 2014; SUGIYAMA et al., 2014). In this context, we proposed an ecological framework for the microbiome dynamics in the soil-rhizosphere interface, in a long-term no-till cropping chronosequence (Figure 3.8).

Besides the differences in selection, the relative role of other processes also varied. We know that the variation in the microbial assembly could be related to spatial distance between sites, apart from environment influence, which is an indirect indicative of drift. Furthermore, part of that residue, resulting from the partitioning of the variability on RDA could represent certain degree of drift. But, drift to be important need to be the dominant driver of microbial dynamics (NEMERGUT et al., 2013). In our data, the effect of distance was not significant in both the db-RDA and the partitioning of the PCNM. Indeed, the residual variation was lower than 50% in all the cases, making the drift meaningless to explain the microbial assembly dynamics (Figure 3.5b).

Dispersal is another process modulating microbial dynamics (COTTENIE, 2005; SZÉKELY; LANGENHEDER, 2014). It can interact with selection and drift at the regional and local scales, modulating microbial dynamics (NEMERGUT et al., 2013). The shift from low to moderate rates of dispersal found in bulk soil along the chronosequence, increased the addition of new species to local communities and possibly allowed inferior species to find temporary refuges or change their metabolic strategies (e.g. dormancy). In the other hand, the shift from high to moderate dispersal in rhizosphere, the extent the chronosequence advanced, decreased the dominance of superior competitors in the community.

Differently of that found to macroecology, where the relative influence of stochasticity is thought to increase with resource supply, we observed that it was true only for bulk soil (FINKE; SNYDER, 2008; LANGENHEDER; PROSSER, 2008; RASCHE et al., 2011) and not for rhizosphere. An interaction between resource supply and other physicochemical conditions that impose selection lead the communities to be more selected in a homogenizing way in the rhizosphere, reducing the role of stochasticity in this soil fraction. Furthermore, we observed that increased resource supply increased stochasticity under physicochemical conditions that do not imposed strong selection, as found in the long-term bulk soil. Changes in resource supply should have little influence over levels of stochasticity (DINI-ANDREOTE et al., 2015; MORI et al., 2013; MUTSHINDA; O'HARA, 2011) in environments with constant resource supply.

Figure 3.8 - A framework for microbial dynamics in a no-till bulk soil-rhizosphere chronosequence

3.5 Conclusion

The microbial community in long-term no-till cropping management evolves, shifting taxa along the time. The trade-off is constant in bulk soil and gradually increase in the rhizosphere, the extent the chronosequence advances. Despite of the trade-offs, the microbial communities of bulk soil-rhizosphere interface become more similar along time. In the bulk soil, both taxa and functions shift along the chronosequence, meaning that stochasticity is a major driver the community dynamics, which explain the higher number of samples assembled by neutral model and the tighter connections between taxa, functions and environmental factors. In the rhizosphere the fight for niche occupancy increases the extent the chronosequence advances, characterized by the high taxa trade-off, high homogenizing selection and reducing stochasticity. Despite of that, the functional resilience acts in order to guarantee the maintenance of the niches, with high functional redundancy and niche stability. This explains why, despite of the increase the numbers of connections between taxa, functions and environment, these connections are sparser. We conclude that long-term no-till leads to homogenization of the microbial community in both bulk soil and rhizosphere. The rhizosphere of the plant increases his power of taxa homogenizing selection along the chronosequence, but the functions are kept.

References

- ALLISON, S.D.; MARTINY, J.B.H. Resistance, resilience, and redundancy in microbial communities. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 105, p. 11512–11519, 2008.
- ATTARD, E.; RECOUS, S.; CHABBI, A.; DE BERRANGER, C.; GUILLAUMAUD, N.; LABREUCHE, J.; PHILIPPOT, L.; SCHMID, B.; LE ROUX, X. Soil environmental conditions rather than denitrifier abundance and diversity drive potential denitrification after changes in land uses. **Global Change Biology**, Malden, v. 17, n. 5, p. 1975–1989, 2011.
- BASTIAN, M.; HEYMANN, S.; JACOMY, M. Gephi: an open source software for exploring and manipulating networks. In: INTERNATIONAL AAAI CONFERENCE ON WEBLOGS AND SOCIAL MEDIA, 2009, San Jose. **Proceedings...** Disponível em: <<http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>>. Acesso em: 10 nov. 2014.
- BENJAMINI, Y.; HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society**, London, v. 57, n. 1, p. 289–300, 1995.

BRODY, J.R.; KERN, S.E. Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis. **BioTechniques**, New York, v. 36, n. 2, p. 214–216, 2004.

CLAESSEN, M.E.C.; BARRETO, W.O.; PAULA, J.L.; DUARTE, M.N. **Manual of soil analysis methods**. 2nd. ed. Rio de Janeiro: Embrapa, 1997. 212 p.

COTTENIE, K. Integrating environmental and spatial processes in ecological community dynamics. **Ecology Letters**, Oxford, v. 8, p. 1175–1182, 2005.

DINI-ANDREOTE, F.; STEGEN, J.C.; VAN ELSAS, J.D.; SALLES, J.F. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. **Proceedings of the National Academy of Sciences**, Washington, v. 112, n. 11, p. 1326–1332, 2015.

DINI-ANDREOTE, F.; SILVA, M.C.P.; TRIADÓ-MARGARIT, X.; CASAMAYOR, E.O.; VAN ELSAS, J.D.; SALLES, J.F. Dynamics of bacterial community succession in a salt marsh chronosequence: evidences for temporal niche partitioning. **The ISME Journal**, London, v. 8, p. 1989–2001, 2014.

FIERER, N.; JACKSON, R.B. The diversity and biogeography of soil bacterial communities. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 103, n. 3, p. 626–631, 2006.

FINKE, D.L.; SNYDER, W.E. Niche partitioning increases resource exploitation by diverse communities. **Science**, Washington, v. 321, p. 1488–1490, 2008.

GEE, G.W.; BAUDER, J.W. Particle-size analysis. In: KLUTE, A. (Ed.). **Methods of soil analysis**. Madison: ASA, 1986. p. 383–411.

HAMMER, Ø.; HARPER, D.A.T.; RYAN, P.D. PAST: Paleontological statistics software package for education and data analysis. **Paleontologia Electronica**, College Station, v. 4, n. 1, p. 1–9, 2001.

HOVATTER, S.R.; DEJELO, C.; CASE, A.L.; BLACKWOOD, C.B. Metacommunity organization of soil microorganisms depends on habitat defined by presence of *Lobelia siphilitica* plants. **Ecology**, Ithaca, v. 92, n. 1, p. 57–65, 2011.

INSAM, H.; DOMSCH, K.H. Relationship between soil organic carbon and microbial biomass on chronosequences of reclamation sites. **Microbial ecology**, New York, v. 15, n. 2, p. 177–188, 1988.

JABOT, F.; ETIENNE, R.S.; CHAVE, J. Reconciling neutral community models and environmental filtering: theory and an empirical test. **Oikos**, Copenhagen, v. 117, p. 1308–1320, 2008.

JACKSON, N.D.; FAHRIG, L. Landscape context affects genetic diversity at a much larger spatial extent than population abundance. **Ecology**, Ithaca, v. 95, n. 4, p. 871–881, 2014.

KEENEY, D.R.; NELSON, D.W. Nitrogen - inorganic forms. In: PAGE, A.L. (Ed.). **Methods in soil analysis**. 2nd ed. Madison: ASA; SSSA, 1982. pt. 2: p. 643–698.

LANGENHEDER, S.; PROSSER, J.I. Resource availability influences the diversity of a functional group of heterotrophic soil bacteria. **Environmental Microbiology**, Oxford, v. 10, n. 9, p. 2245–2256, 2008.

LEPŠ, J.; ŠMILAUER, P. Multivariate analysis of ecological data using CANOCO. **Bulletin of the Ecological Society of America**, Ithaca, v. 86, p. 6–6, 2005.

MELO, W.J.; MELO, G.M.P.; ARAÚJO, A.S.F.; MELO, V.P. Avaliação da atividade enzimática em amostras de solo. In: FIGUEIREDO, M.V.B.; BURITY, H.A.; OLIVEIRA, J.P.; SANTOS, C.E. (Ed.). **Biotecnologia aplicada à agricultura**. Recife: EMBRAPA, 2010. p. 153–187.

MENDES, L.W.; KURAMAE, E.E.; NAVARRETE, A.A.; VAN VEEN, J.A.; TSAI, S.M. Taxonomical and functional microbial community selection in soybean rhizosphere. **The ISME Journal**, London, v. 8, p. 1577–1587, 2014.

MENDES, L.W.; TSAI, S.M.; NAVARRETE, A.A.; DE HOLLANDER, M.; VAN VEEN, J.A.; KURAMAE, E.E. Soil-borne microbiome: linking diversity to function. **Microbial Ecology**, New York, v. 70, p. 255–265, 2015.

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. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. **BMC Bioinformatics**, London, v. 9, n. 386, p. 1–8, 2008.

MORI, A.S.; SHIONO, T.; KOIDE, D.; KITAGAWA, R.; OTA, A.T.; MIZUMACHI, E. Community assembly processes shape an altitudinal gradient of forest biodiversity. **Global Ecology and Biogeography**, Malden, v. 22, n. 7, p. 878–888, 2013.

MUELLER, R.C.; PAULA, F.S.; MIRZA, B.S.; RODRIGUES, J.L.M.; NÜSSLEIN, K.; BOHANNAN, B.J.M. Links between plant and fungal communities across a deforestation chronosequence in the Amazon rainforest. **The ISME Journal**, London, v. 8, p. 1548–1550, 2014.

MUTSHINDA, C.M.; O'HARA, R.B. Integrating the niche and neutral perspectives on community structure and dynamics. **Oecologia**, Berlin, v. 166, n. 1, p. 241–251, 2011.

NAVARRETE, A.A.; TAKETANI, R.G.; MENDES, L.W.; CANNAVAN, F.D.S.; MOREIRA, F.M.S.; TSAI, S.M. Land-use systems affect archaeal community structure and functional diversity in western amazon soils. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 35, n. 1, p. 1527–1540, 2011.

NEMERGUT, D.R.; SCHMIDT, S.K.; FUKAMI, T.; O'NEILL, S.P.; BILINSKI, T.M.; STANISH, L.F.; KNELMAN, J.E.; DARCY, J.L.; LYNCH, R.C.; WICKEY, P.; FERRENBURG, S. Patterns and processes of microbial community assembly. **Microbiology and Molecular Biology Reviews**, Washington, v. 77, n. 3, p. 342–356, 2013.

NEWCOMBE, R.G. Improved confidence intervals for the difference between binomial proportions based on paired data. **Statistics in Medicine**, Chichester, v. 17, p. 2635–2650, 1998.

NEWMAN, M.E.J. The structure and function of complex networks. **SIAM Review**, Philadelphia, v. 45, p. 110–122, 2003.

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. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). **Nucleic Acids Research**, Oxford, v. 42, p. 1–9, 2013.

PARKS, D.H.; BEIKO, R.G. Identifying biologically relevant differences between metagenomic communities. **Bioinformatics**, Oxford, v. 26, n. 6, p. 715–721, 2010.

RASCHE, F.; KNAPP, D.; KAISER, C.; KORANDA, M.; KITZLER, B.; ZECHMEISTER-BOLTENSTERN, S.; RICHTER, A.; SESSITSCH, A. Seasonality and resource availability control bacterial and archaeal communities in soils of a temperate beech forest. **The ISME Journal**, London v. 5, n. 3, p. 389–402, 2011.

RODRIGUES, J.L.M.; PELLIZARI, V.H.; MUELLER, R.; BAEK, K.; JESUS, E.D.C.; PAULA, F.S.; MIRZA, B.; HAMAOU, G.S.; TSAI, S.M.; FEIGL, B.; TIEDJE, J.M.; BOHANNAN, B.J.M.; NÜSSLEIN, K. Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 110, n. 3, p. 988–993, 2013.

ROUSK, J.; BAATH, E.; BROOKES, P.C.; LAUBER, C.L.; LOZUPONE, C.; CAPORASO, J.G.; KNIGHT, R.; FIERER, N. Soil bacterial and fungal communities across a pH gradient in an arable soil. **The ISME Journal**, London, v. 4, p. 1340–1351, 2010.

SMITH, R.J.; JEFFRIES, T.C.; ROUDNEW, B.; FITCH, A.J.; SEYMOUR, J.R.; DELPIN, M.W.; NEWTON, K.; BROWN, M.H.; MITCHELL, J. G. Metagenomic comparison of microbial communities inhabiting confined and unconfined aquifer ecosystems. **Environmental Microbiology**, Oxford, v. 14, n. 1, p. 240–253, 2012.

STEGEN, J.C.; LIN, X.; FREDRICKSON, J.K.; CHEN, X.; KENNEDY, D.W.; MURRAY, C.J.; ROCKHOLD, M.L.; KONOPKA, A. Quantifying community assembly processes and identifying features that impose them. **The ISME Journal**, London, v. 7, n. 11, p. 2069–2079, 2013.

SUGIYAMA, A.; UEDA, Y.; ZUSHI, T.; TAKASE, H.; YAZAKI, K. Changes in the bacterial community of soybean rhizospheres during growth in the field. **PLoS One**, New York, v. 9, n. 6, p. 1–9, 2014.

SZÉKELY, A.J.; LANGENHEDER, S. The importance of species sorting differs between habitat generalists and specialists in bacterial communities. **FEMS Microbiology Ecology**, Amsterdam, v. 87, n. 1, p. 102–112, 2014.

TEAM, R.D.C. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2008. Disponível em: <<http://www.R-project.org>>. Acesso em: 17 set. 2015.

TEDESCO, M.J.; GIANELLO, C.; BISSANI, C.A.; BOHNEN, H.; VOLKWEISS, S.J. **Analysis of soil, plants and other materials**. Porto Alegre: Universidade Federal do Rio Grande do Sul, 1995. 174 p.

VELLEND, M. Conceptual synthesis in community ecology. **The Quarterly Review of Biology**, Chicago, v. 85, n. 2, p. 183–206, 2010.

WELCH, B.L. The generalization of student's problem when several different population variances are involved. **Biometrika**, Oxford, v. 34, n. 1, p. 28–35, 1934.

4 GENERAL CONCLUSIONS

Here, we trespassed the threshold of the pure descriptive patterns inference about microbial dynamics, that is selection, going beyond to obtain a comprehensive framework of the other processes—diversification, dispersal, drift, beside selection—that govern microbial evolutionary dynamics in long-term land-use change. We also evaluated how these processes interact each other along time in no-till cropping and how the plant-roots act to modulate microbial communities assembly. The outcome of this framework will allow us to depict how soil microbial communities evolve along time in agroecosystems, in the light of theoretical ecological paradigms.

APPENDICES

Appendix A - Supplementary information from Chapter 2

Supplementary Table ST 2.1 - Soil physicochemical characteristics from the 0-20 cm profile from three land-use systems, two counties, in two different seasons. Sampling sites are located at the Atlantic Rainforest Biome, Santa Catarina State, Brazil. Results from 38 measured or calculated variables

Sample	Land-use	County	Season	N-total	C-total	H-total	S-total	C:N ratio	N-NH ₄ ⁺	N-NO ₃ ⁻	Moisture
wocf1	Forest	Chapecó	Winter	3.4	43.7	15.6	0.7	12.9	14.0	28.0	29.2
wocf2	Forest	Chapecó	Winter	3.9	54.3	16.7	0.9	13.8	24.5	24.5	25.2
wocf3	Forest	Chapecó	Winter	3.7	47.3	16.6	0.9	12.9	49.0	35.0	21.4
woxf1	Forest	Xanxerê	Winter	3.9	47.3	17.9	0.7	12.0	10.5	21.0	16.3
woxf2	Forest	Xanxerê	Winter	3.8	46.5	18.2	0.7	12.3	35.0	35.0	15.7
woxf3	Forest	Xanxerê	Winter	4.0	47.1	16.4	0.7	11.9	14.0	24.5	21.7
wocp1	Grassland	Chapecó	Winter	2.7	36.4	15.7	0.6	13.3	35.0	42.0	15.8
wocp2	Grassland	Chapecó	Winter	2.6	35.6	15.9	0.6	13.7	31.5	31.5	17.6
wocp3	Grassland	Chapecó	Winter	2.3	32.5	15.4	0.7	14.1	56.0	52.5	21.1
woxp1	Grassland	Xanxerê	Winter	3.3	42.1	18.2	0.8	12.9	28.0	38.5	21.2
woxp2	Grassland	Xanxerê	Winter	3.7	46.3	18.3	0.8	12.5	14.0	24.5	30.4
woxp3	Grassland	Xanxerê	Winter	3.3	44.5	16.9	0.7	13.8	28.0	38.5	30.2
wocd1	No-till	Chapecó	Winter	2.6	30.9	14.3	0.7	12.1	10.5	49.0	21.8
wocd2	No-till	Chapecó	Winter	2.7	33.1	15.3	0.6	12.2	7.0	52.5	23.1
wocd3	No-till	Chapecó	Winter	2.5	30.2	14.1	0.7	12.0	14.0	73.5	22.3
woxd1	No-till	Xanxerê	Winter	2.5	33.8	16.5	0.6	13.7	3.5	14.0	18.6
woxd2	No-till	Xanxerê	Winter	2.4	33.8	16.6	0.7	13.9	14.0	24.5	19.9
woxd3	No-till	Xanxerê	Winter	2.3	32.7	16.5	1.0	14.2	17.5	28.0	18.4
socf1	Forest	Chapecó	Summer	2.6	41.0	16.4	2.5	15.6	91.0	168.0	29.2
socf2	Forest	Chapecó	Summer	3.6	54.6	17.7	2.3	15.0	115.5	185.5	25.2
socf3	Forest	Chapecó	Summer	3.6	55.0	16.6	2.7	15.2	129.5	161.0	21.4

(continue)

Supplementary Table ST 2.1 - Soil physicochemical characteristics from the 0-20 cm profile from three land-use systems, two counties, in two different seasons. Sampling sites are located at the Atlantic Rainforest Biome, Santa Catarina State, Brazil. Results from 38 measured or calculated variables

Sample	Land-use	County	Season	N-total	C-total	H-total	S-total	C:N ratio	N-NH ₄ ⁺	N-NO ₃ ⁻	Moisture
soxf1	Forest	Xanxerê	Summer	3.8	44.5	18.6	5.5	11.7	150.5	87.5	16.3
soxf2	Forest	Xanxerê	Summer	3.3	40.7	16.5	3.2	12.5	126.0	129.5	15.7
soxf3	Forest	Xanxerê	Summer	3.9	47.4	18.6	9.4	12.3	157.5	119.0	21.7
socp1	Grassland	Chapecó	Summer	2.2	37.2	16.2	2.4	16.8	112.0	129.5	15.8
socp2	Grassland	Chapecó	Summer	2.6	39.7	16.9	2.5	15.5	122.5	129.5	17.6
socp3	Grassland	Chapecó	Summer	2.0	34.0	16.2	2.3	16.6	161.0	164.5	21.1
soxp1	Grassland	Xanxerê	Summer	3.6	47.1	17.9	6.0	13.0	112.0	147.0	21.2
soxp2	Grassland	Xanxerê	Summer	4.0	50.8	19.1	3.4	12.8	115.5	164.5	30.4
soxp3	Grassland	Xanxerê	Summer	4.2	51.9	18.4	3.6	12.4	129.5	157.5	30.2
socd1	No-till	Chapecó	Summer	2.4	34.0	14.9	2.6	14.3	140.0	175.0	21.8
socd2	No-till	Chapecó	Summer	2.5	35.0	14.8	2.3	13.9	112.0	178.5	23.1
socd3	No-till	Chapecó	Summer	2.5	35.0	14.3	2.6	14.0	115.5	136.5	22.3
soxd1	No-till	Xanxerê	Summer	2.3	33.3	15.7	4.0	14.6	129.5	147.0	18.6
soxd2	No-till	Xanxerê	Summer	2.1	31.3	17.1	4.4	14.6	157.5	98.0	19.9
soxd3	No-till	Xanxerê	Summer	2.3	33.1	17.5	3.1	14.3	140.0	157.5	18.4

Supplementary Table ST 2.1 - Soil physicochemical characteristics from the 0-20 cm profile from three land-use systems, two counties, in two different seasons. Sampling sites are located at the Atlantic Rainforest Biome, Santa Catarina State, Brazil. Results from 38 measured or calculated variables

Sample	Land-use	County	Altitude	Longitude (UTM)	Latitude (UTM)	pH _{water}	pH _{CaCl₂}	P	K ⁺	SOM	Al ³⁺	Ca ²⁺	Mg ²⁺
ocf1	Forest	Chapecó	619.3	331718.5	7008635.6	4.0	4.5	5.2	86.2	4.2	4.6	0.9	0.3
ocf2	Forest	Chapecó	608.7	331684.2	7008607.4	4.0	4.3	4.7	63.5	5.3	4.0	1.6	0.4
ocf3	Forest	Chapecó	594.3	331656.2	7008597.8	4.1	4.4	5.1	66.9	5.0	4.8	1.1	0.3
oxf1	Forest	Xanxerê	710.0	355616.4	7031154.8	4.4	4.9	5.0	59.9	4.7	2.4	1.5	0.8
oxf2	Forest	Xanxerê	715.3	355588.2	7031158.8	4.6	5.2	5.0	67.2	4.8	1.7	2.8	1.1
oxf3	Forest	Xanxerê	723.3	355560.5	7031164.6	4.5	5.0	4.8	99.2	4.9	2.0	2.3	1.1
ocp1	Grassland	Chapecó	640.7	336913.9	7002706.1	5.3	5.7	8.6	123.0	4.7	0.6	4.0	2.5
ocp2	Grassland	Chapecó	643.0	336899.3	7002730.6	5.5	5.8	4.5	71.8	4.8	0.0	4.0	2.6
ocp3	Grassland	Chapecó	640.3	336885.7	7002757.5	5.4	5.7	5.6	70.3	4.5	0.3	4.1	2.8
oxp1	Grassland	Xanxerê	720.0	353916.6	7031785.2	4.5	5.1	4.6	250.2	5.1	2.2	1.9	1.1
oxp2	Grassland	Xanxerê	723.3	353940.5	7031773.8	4.7	5.3	4.0	253.9	5.2	1.9	2.1	1.1
oxp3	Grassland	Xanxerê	725.3	353961.6	7031766.0	4.4	5.2	4.4	273.1	5.1	2.3	1.9	1.0
ocd1	No-till	Chapecó	681.7	332385.4	7010950.6	5.1	5.7	12.9	347.0	4.6	1.2	5.2	1.6
ocd2	No-till	Chapecó	678.3	332362.4	7010936.7	5.5	6.1	13.0	431.2	4.8	0.0	6.5	1.8
ocd3	No-till	Chapecó	674.3	332337.2	7010923.5	5.3	6.0	9.2	322.5	4.5	0.9	5.9	1.8
oxd1	No-till	Xanxerê	749.0	353335.5	7030566.5	5.3	5.8	15.1	96.6	3.9	0.6	7.4	2.8
oxd2	No-till	Xanxerê	746.7	353336.2	7030597.9	5.6	6.0	16.3	107.6	3.9	0.0	8.6	3.2
oxd3	No-till	Xanxerê	742.3	353338.7	7030623.8	5.3	5.8	16.4	103.0	3.8	0.2	6.9	2.4

Supplementary Table ST 2.1 - Soil physicochemical characteristics from the 0-20 cm profile from three land-use systems, two counties, in two different seasons. Sampling sites are located at the Atlantic Rainforest Biome, Santa Catarina State, Brazil. Results from 38 measured or calculated variables

(continue)												
Sample	Land-use	County	Soil density	Total porosity	Micropores	Macropores	Biopores	P. Resistance	Sand	Clay	Silt	P. A. A. diameter
ocf1	Forest	Chapecó	0.9	0.6	0.3	0.3	0.2	0.6	38.9	49.3	11.8	3.8
ocf2	Forest	Chapecó	0.9	0.6	0.3	0.3	0.2	0.7	35.4	47.7	17.0	5.4
ocf3	Forest	Chapecó	1.0	0.6	0.4	0.3	0.1	1.0	34.3	48.7	17.0	5.3
oxf1	Forest	Xanxerê	0.9	0.7	0.4	0.3	0.1	0.5	21.6	48.0	30.4	5.2
oxf2	Forest	Xanxerê	0.9	0.7	0.4	0.2	0.1	0.5	28.6	47.0	24.4	5.4
oxf3	Forest	Xanxerê	1.0	0.6	0.3	0.3	0.1	0.6	19.2	49.3	31.5	5.7
ocp1	Grassland	Chapecó	1.2	0.6	0.5	0.1	0.0	2.4	29.6	52.7	17.8	5.7
ocp2	Grassland	Chapecó	1.2	0.6	0.5	0.1	0.0	2.2	34.3	49.0	16.7	5.6
ocp3	Grassland	Chapecó	1.1	0.6	0.5	0.1	0.0	2.3	31.4	51.7	16.9	5.5
oxp1	Grassland	Xanxerê	1.1	0.6	0.5	0.1	0.0	2.1	15.6	53.0	31.4	5.7
oxp2	Grassland	Xanxerê	1.0	0.6	0.6	0.0	0.0	1.8	20.4	53.7	25.9	5.6
oxp3	Grassland	Xanxerê	1.0	0.6	0.6	0.0	0.0	2.0	23.0	51.3	25.7	5.8
ocd1	No-till	Chapecó	1.2	0.6	0.5	0.1	0.0	1.2	30.7	58.3	11.0	5.9
ocd2	No-till	Chapecó	1.2	0.6	0.5	0.1	0.1	1.2	29.1	54.0	16.9	5.8
ocd3	No-till	Chapecó	1.1	0.6	0.5	0.1	0.1	1.2	34.9	53.3	11.8	6.0
oxd1	No-till	Xanxerê	1.1	0.6	0.5	0.1	0.0	1.8	37.0	46.0	17.0	5.0
oxd2	No-till	Xanxerê	1.1	0.6	0.5	0.1	0.0	1.6	33.9	43.0	23.1	4.8
oxd3	No-till	Xanxerê	1.1	0.6	0.5	0.1	0.0	1.7	38.4	50.3	11.3	4.8

Supplementary Table ST 2.1 - Soil physicochemical characteristics from the 0-20 cm profile from three land-use systems, two counties, in two different seasons. Sampling sites are located at the Atlantic Rainforest Biome, Santa Catarina State, Brazil. Results from 38 measured or calculated variables

(conclusion)													
Sample	Land-use	County	H + Al	CEC	pH 7.0	Al ₃ ⁺ Saturation	Bases saturation	K CEC	Ca CEC	Mg CEC	Ca/Mg	Ca/K	Mg/K
ocf1	Forest	Chapecó	25.2	26.7	75.8	5.9	0.9	3.8	1.2	3.2	4.3	1.4	
ocf2	Forest	Chapecó	31.3	33.5	65.5	7.2	0.5	5.3	1.4	3.5	9.8	2.6	
ocf3	Forest	Chapecó	28.2	29.7	75.3	5.7	0.6	4.1	1.1	3.6	6.9	1.8	
oxf1	Forest	Xanxerê	16.2	18.7	49.9	14.7	0.8	8.9	5.0	2.0	10.2	5.6	
oxf2	Forest	Xanxerê	12.6	16.7	30.2	28.0	1.0	19.2	7.8	2.4	19.7	8.1	
oxf3	Forest	Xanxerê	14.3	18.0	35.4	20.3	1.4	12.8	6.1	2.1	9.0	4.3	
ocp1	Grassland	Chapecó	6.4	13.2	9.0	51.5	2.4	30.5	18.6	1.6	13.6	8.4	
ocp2	Grassland	Chapecó	5.3	12.2	0.0	56.3	1.5	33.4	21.3	1.6	22.5	14.6	
ocp3	Grassland	Chapecó	6.0	13.0	4.1	53.9	1.4	31.2	21.4	1.5	22.5	15.4	
oxp1	Grassland	Xanxerê	12.9	16.5	37.9	22.5	3.9	12.0	6.6	1.8	3.2	1.7	
oxp2	Grassland	Xanxerê	9.4	13.3	33.9	29.1	4.8	16.2	8.1	2.0	3.6	1.8	
oxp3	Grassland	Xanxerê	10.8	14.3	39.0	25.0	4.9	12.9	7.1	1.8	3.0	1.6	
ocd1	No-till	Chapecó	5.9	13.7	13.7	56.6	6.5	38.1	12.0	3.2	5.9	1.8	
ocd2	No-till	Chapecó	4.1	13.5	0.0	69.9	8.2	48.4	13.3	3.6	5.9	1.6	
ocd3	No-till	Chapecó	4.6	13.2	9.5	64.9	6.3	45.0	13.7	3.3	7.6	2.4	
oxd1	No-till	Xanxerê	5.4	15.8	5.5	65.9	1.5	46.8	17.5	2.7	34.9	13.0	
oxd2	No-till	Xanxerê	4.4	16.4	0.0	73.2	1.7	52.0	19.4	2.7	32.1	11.8	
oxd3	No-till	Xanxerê	5.3	15.0	2.4	64.1	1.7	46.2	16.2	2.9	28.1	9.8	

Units: N-total (mg L⁻¹); C-total (mg L⁻¹); H-total (mg L⁻¹); S-total(mg L⁻¹); C:N ratio (C/N); N-NH₄⁺ (mg L⁻¹); N-NO₃⁻ (mg L⁻¹); Moisture (%); Altitude (m); Longitude (UTM); Latitude (UTM); pHwater (-log [H⁺]); pHCaCl₂ (-log [H⁺]); P (mg L⁻¹); K⁺ (mg L⁻¹); SOM (%); Al³⁺ (cmol_c dm⁻³); Ca²⁺ (cmol_c dm⁻³); Mg²⁺ (cmol_c dm⁻³); Soil density (g dm⁻³); Total porosity; Micropores; Macropores; Biopores; Penetration resistance; Sand; Clay; Silte; Ponderate aggregate average diameter; H + Al (cmol_c dm⁻³); CEC pH 7.0; Al³⁺ Saturation; Bases saturation; K CEC; Ca CEC; Mg CEC; Ca/Mg; Ca/K; Mg/K

(conclusion)

Supplementary Table ST 2.2 - Global and partial contributions of environmental and season + geographic variables on the variance of taxonomic assembly at Genus Level from M5NR non-redundant database. Distance Based Redundancy Analysis (db-RDA), with forward selection. Data show the adjusted explanation of variation from the simple effects of environmental variables and season + geographic variables

Analysis	Forest R^2_{adj}	Grassland R^2_{adj}	No-till R^2_{adj}
Environmental variables			
<i>Global</i>	0.32*	0.75***	0.14
<i>Forward-selected raw</i>	0.21*	0.68***	0.10
Season + geographic variables			
<i>Global</i>	0.18*	0.26*	0.12*
<i>Forward-selected raw</i>	0.06	0.22*	0.08

Supplementary Table ST 2.3 - Global and partial contributions of environmental and season + geographic variables on the variance of taxonomic assembly at Level 2 from SEED subsystems database. Distance Based Redundancy Analysis (db-RDA), with forward selection. Data show the adjusted explanation of variation from the simple effects of environmental variables and season + geographic variables

Analysis	Forest R^2_{adj}	Grassland R^2_{adj}	No-till R^2_{adj}
Environmental variables			
<i>Global</i>	0.18*	0.57**	0.22
<i>Forward-selected raw</i>	0.10*	0.44**	0.15
Season + geographic variables			
<i>Global</i>	0.08*	0.18*	0.12*
<i>Forward-selected raw</i>	0.03	0.11*	0.08

Supplementary Table ST 2.4 – Number of pairwise correlations as inferred by Spearman’s rank abundance coefficient. Pairwise correlations were obtained from taxonomic class level, functional subsystems level 2 and environmental data

Land-use	Pairwise	Taxonomic		Functional		Taxonomic X Functional		Taxonomic X explanatory variables		Functional X explanatory variables			
		Site 1	Site 2	Site 1	Site 2	Site 1	Site 2	Site 1	Site 2	Site 1	Site 2		
Forest	Winter sampling	Pairwise correlations	2,413	2,411	17,931	18,130	13,271	13,358	2,491	2,578	6,860	6,639	
		Significant correlations	185 (33)	241 (28)	1,050 (178)	1,121 (152)	657 (129)	788 (91)	225 (144)	177 (37)	783 (359)	698 (244)	
		Significant positive correlations	122 (25)	165 (22)	551 (93)	557 (72)	307 (55)	355 (39)	103 (67)	83 (19)	388 (187)	374 (134)	
		Significant negative correlations	63 (8)	76 (6)	499 (85)	564 (80)	350 (74)	433 (52)	122 (77)	94 (18)	395 (172)	324 (110)	
	Summer sampling	Pairwise correlations	2,414	2,413	17,561	18,701	13,145	13,574	2,506	2,531	6,893	7,130	
		Significant correlations	136 (9)	136 (19)	883 (120)	1,179 (171)	587 (78)	750 (106)	298 (100)	304 (110)	948 (284)	879 (395)	
		Significant positive correlations	110 (8)	110 (18)	415 (53)	642 (89)	297 (49)	392 (55)	167 (57)	167 (54)	483 (143)	458 (223)	
		Significant negative correlations	26 (1)	26 (1)	468 (67)	537 (82)	290 (29)	358 (51)	131 (43)	137 (56)	465 (141)	421 (172)	
	Grassland	Winter sampling	Pairwise correlations	2,403	2,410	18,910	18,525	13,633	13,505	2,611	2,665	7,135	7,064
			Significant correlations	179 (43)	271 (30)	1,931 (231)	1,441 (151)	1,226 (126)	1,195 (128)	522 (308)	749 (406)	1,582 (960)	1,603 (829)
			Significant positive correlations	108 (17)	208 (25)	939 (110)	709 (83)	643 (68)	596 (70)	231 (140)	403 (240)	782 (486)	801 (402)
			Significant negative correlations	71 (26)	63 (5)	992 (121)	732 (68)	583 (58)	599 (58)	291 (168)	346 (166)	800 (474)	802 (427)
Summer sampling		Pairwise correlations	2,416	2,412	18,690	19,110	13,565	13,714	2,623	2,599	7,306	7,273	
		Significant correlations	167 (38)	130 (13)	1,537 (241)	1,272 (122)	1,006 (149)	759 (76)	627 (306)	371 (238)	1,753 (908)	1,403 (780)	
		Significant positive correlations	109 (20)	75 (10)	755 (120)	626 (67)	533 (88)	379 (42)	295 (138)	173 (106)	870 (451)	733 (373)	
		Significant negative correlations	58 (18)	55 (3)	782 (121)	646 (55)	473 (61)	380 (34)	332 (168)	198 (132)	883 (457)	670 (407)	
No-till		Winter sampling	Pairwise correlations	2,414	2,415	17,731	17,910	13,209	13,277	2,420	2,413	6,628	6,350
			Significant correlations	268 (43)	274 (37)	968 (139)	1,181 (222)	608 (81)	944 (151)	146 (28)	42 (7)	538 (162)	548 (130)
			Significant positive correlations	173 (28)	163 (21)	465 (70)	607 (113)	318 (48)	458 (65)	87 (18)	22 (3)	271 (81)	272 (60)
			Significant negative correlations	95 (15)	111 (16)	503 (69)	574 (109)	290 (33)	486 (86)	59 (10)	20 (4)	267 (81)	276 (70)
	Summer sampling	Pairwise correlations	2,411	2,414	17,934	18,713	13,291	13,574	2,450	2,473	6,675	6,937	
		Significant correlations	147 (22)	298 (35)	901 (124)	1240 (176)	548 (71)	1067 (157)	285 (105)	149 (0)	750 (280)	779 (251)	
		Significant positive correlations	70 (9)	165 (22)	492 (77)	633 (89)	285 (35)	554 (86)	145 (49)	69 (0)	409 (161)	371 (110)	
		Significant negative correlations	77 (13)	133 (13)	409 (47)	607 (87)	263 (36)	513 (71)	140 (56)	80 (0)	241 (119)	408 (141)	

Appendix B - Supplementary information from Chapter 3

Supplementary Figure SF 3.1 - Values of (a) precipitation (mm) and (b) temperature (°C) along the year of 2013, in the municipality of Querência, Mato Grosso State, Brazil. Source: INMET

Supplementary Table ST 3.1 - Soil physicochemical results from the 0-20 cm profile from three chronosequences. Results from 29 measured or calculated variables
(continue)

Sample	N-total	N-NH ₄ ⁺	N-NO ₃ ⁻	Moisture	SOM	SOC	Sandy	Silt	Clay	Texture	pH H ₂ O	pH KCl	P	K ⁺
PA1	1183	17	10	21.29	45	26	680	18	302	md-arg.	5.5	4.8	16	14
PA2	1659	6	10	24.97	49	28	672	27	302	md-arg.	5.9	5.1	8	11
PA3	1365	14	11	18.81	34	19	680	17	303	md-arg.	6	5.2	24	10
PA4	1393	6	11	25.30	35	20	689	34	277	md-arg.	5.9	4.9	7	20
PA5	1484	7	17	28.36	37	22	679	19	302	md-arg.	5.7	4.6	8	9
PB1	1407	7	21	22.38	34	20	691	32	277	md-arg.	5.9	4.9	12	18
PB2	1218	6	11	21.69	27	15	695	27	277	md-arg.	5.9	5	28	13
PB3	1141	3	6	26.36	29	17	683	14	302	md-arg.	5.9	4.9	6	18
PB4	1400	4	24	16.68	36	21	688	10	302	md-arg.	5.9	4.9	4	8
PB5	1743	14	17	23.85	32	18	689	33	278	md-arg.	6.4	5.5	4	11
PC1	1281	7	13	36.30	31	18	666	30	303	md-arg.	6	5.2	11	5
PC2	1120	16	19	23.99	27	16	648	24	328	md-arg.	6.1	5.3	9	8
PC3	1323	27	13	36.42	24	14	673	25	303	md-arg.	5.8	5.1	23	5
PC4	1589	4	12	16.41	38	22	649	23	327	md-arg.	6.2	5.2	12	6
PC5	1106	4	10	26.64	27	16	662	10	328	md-arg.	6.2	5.6	3	13

Supplementary Table ST 3.1 - Soil physicochemical results from the 0-20 cm profile from three chronosequences. Results from 29 measured or calculated variables
(continue)

Sample	N-total	N-NH ₄ ⁺	N-NO ₃ ⁻	Moisture	SOM	SOC	Sandy	Silt	Clay	Texture	pH H ₂ O	pH KCl	P	K ⁺
SA1	1008	2	16	26.67	44	26	696	27	277	md-arg.	6.6	6	20	14
SA2	1680	5	19	15.59	39	23	701	21	277	md-arg.	6.4	6	15	11
SA3	1477	3	14	12.76	44	26	679	18	303	md-arg.	6.8	6.1	12	11
SA4	1540	7	20	28.92	44	26	694	17	290	md-arg.	6.5	6	32	12
SA5	1729	3	23	13.40	43	25	702	21	277	md-arg.	6.7	5.9	10	8
SB1	1568	5	15	22.42	35	20	696	28	277	md-arg.	6.7	5.8	2	7
SB2	1554	6	28	22.37	40	23	694	29	277	md-arg.	6.6	6	11	5
SB3	1743	4	16	21.21	34	20	689	33	277	md-arg.	6.7	5.8	12	6
SB4	1407	2	18	19.73	40	23	682	16	302	md-arg.	6.7	6.2	17	6
SB5	1743	7	15	21.92	39	23	699	24	277	md-arg.	6.8	5.9	12	7
SC1	1540	2	13	24.20	38	22	645	29	326	md-arg.	6.6	6	5	8
SC2	1526	4	24	21.49	41	24	654	19	327	md-arg.	6.9	6	26	6
SC3	1449	6	14	22.10	44	26	650	22	328	md-arg.	6.8	6.2	8	8
SC4	1519	3	17	27.12	44	26	638	35	327	md-arg.	6.9	6.1	10	7
SC5	1897	5	42	18.11	45	26	652	20	328	md-arg.	6.8	6.1	16	8

Supplementary Table ST 3.1 - Soil physicochemical results from the 0-20 cm profile from three chronosequences. Results from 29 measured or calculated variables
(continue)

Sample	N-total	N-NH ₄ ⁺	N-NO ₃ ⁻	Moisture	SOM	SOC	Sandy	Silt	Clay	Texture	pH H ₂ O	pH KCl	P	K ⁺
TA1	1512	2	6	22.52	46	27	649	24	327	md-arg.	6.8	6	18	10
TA2	1666	4	12	20.13	53	31	629	30	341	md-arg.	6.6	5.8	27	14
TA3	1561	20	16	23.38	45	26	636	25	339	md-arg.	6.7	5.9	17	8
TA4	1484	19	20	24.98	48	28	628	19	353	arg.	6.6	5.9	24	10
TA5	1505	2	5	23.34	45	26	652	22	327	md-arg.	6.7	6.1	19	10
TB1	1659	2	6	29.44	47	27	617	31	352	arg.	6.5	5.6	61	14
TB2	1659	2	9	29.69	49	28	615	33	352	arg.	6.6	5.7	73	19
TB3	1764	11	7	27.30	46	27	612	35	352	arg.	6.5	5.8	32	12
TB4	2100	1	7	93.61	55	32	603	21	377	arg.	6.4	5.6	32	12
TB5	1638	2	3	26.23	39	23	610	14	377	arg.	6.8	6	19	9
TC1	1596	5	6	23.15	44	25	692	31	277	md-arg.	6.1	5.5	25	7
TC2	1295	2	11	22.55	37	21	707	16	277	md-arg.	6.7	5.9	22	6
TC3	1477	11	7	32.45	40	23	710	13	277	md-arg.	6.3	5.7	32	5
TC4	1589	3	10	25.87	43	25	709	14	277	md-arg.	6.7	5.9	27	7
TC5	1862	26	24	24.77	51	30	689	22	289	md-arg.	6.3	5.8	30	8

Supplementary Table ST 3.1 - Soil physicochemical results from the 0-20 cm profile from three chronosequences. Results from 29 measured or calculated variables
(continue)

Sample	Ca ²⁺	Mg ²⁺	Al ³⁺	H+Al	SB	CEC	Cations sat.	Al sat.	B	Cu ²⁺	Fe ²⁺	Mn ²⁺	Zn ²⁺
PA1	14	7	1	38	223	604	37	6	0	0	30	0	3
PA2	21	11	0	24	332	574	58	0	0	4	34	0	4
PA3	19	12	0	26	317	577	55	0	0	0	51	8	4
PA4	12	6	1	33	196	525	37	7	11	0	63	0	0
PA5	11	5	4	36	173	537	32	18	10	0	33	9	0
PB1	13	7	1	33	220	550	40	6	9	0	61	8	0
PB2	13	9	0	29	229	524	44	0	0	0	92	8	4
PB3	11	8	1	24	214	457	47	6	0	0	83	8	3
PB4	12	8	1	35	211	558	38	6	0	3	37	0	0
PB5	19	13	0	17	336	509	66	0	0	0	59	11	3
PC1	16	11	0	24	268	511	53	0	0	0	101	10	3
PC2	15	10	0	24	255	497	51	0	0	0	65	9	0
PC3	16	10	0	16	263	419	63	0	6	0	59	8	0
PC4	20	13	0	31	338	650	52	0	0	0	65	20	4
PC5	19	12	0	14	323	462	70	0	0	0	48	9	0

Supplementary Table ST 3.1 - Soil physicochemical results from the 0-20 cm profile from three chronosequences. Results from 29 measured or calculated variables
(continue)

Sample	Ca ²⁺	Mg ²⁺	Al ³⁺	H+Al	SB	CEC	Cations sat.	Al sat.	B	Cu ²⁺	Fe ²⁺	Mn ²⁺	Zn ²⁺
SA1	31	22	0	12	543	664	82	0	7	0	33	14	9
SA2	31	23	0	7	545	614	89	0	0	0	28	11	24
SA3	30	22	0	10	529	633	84	0	0	0	24	13	10
SA4	30	19	0	16	496	652	76	0	0	0	31	14	11
SA5	28	21	0	14	499	638	78	0	7	0	36	11	9
SB1	21	14	0	9	360	446	81	0	7	0	36	11	9
SB2	24	18	0	5	429	481	89	0	0	0	27	12	11
SB3	22	15	0	9	378	464	81	0	24	3	28	10	16
SB4	28	23	0	9	508	594	85	0	28	0	25	11	11
SB5	26	22	0	9	490	577	85	0	8	0	24	13	12
SC1	33	22	0	17	559	732	76	0	0	0	35	11	16
SC2	34	20	0	14	542	681	80	0	12	0	32	12	13
SC3	35	22	0	7	577	646	89	0	6	0	32	11	15
SC4	35	24	0	12	598	720	83	0	16	0	32	14	13
SC5	35	24	0	14	603	741	81	0	15	5	34	9	18

Supplementary Table ST 3.1 - Soil physicochemical results from the 0-20 cm profile from three chronosequences. Results from 29 measured or calculated variables
(conclusion)

Sample	Ca ²⁺	Mg ²⁺	Al ³⁺	H+Al	SB	CEC	Cations sat.	Al sat.	B	Cu ²⁺	Fe ²⁺	Mn ²⁺	Zn ²⁺
TA1	38	19	0	12	577	698	83	0	23	5	36	19	23
TA2	37	18	0	16	565	721	78	0	14	4	35	10	17
TA3	31	17	0	16	489	645	76	0	12	4	44	18	19
TA4	38	20	0	14	593	731	81	0	21	5	43	22	23
TA5	40	20	0	12	607	729	83	0	21	3	34	21	20
TB1	37	17	0	16	549	705	78	0	8	5	38	8	18
TB2	38	19	0	17	588	761	77	0	6	4	37	10	23
TB3	39	19	0	14	595	734	81	0	0	3	34	12	23
TB4	42	20	0	24	633	876	72	0	0	5	38	11	27
TB5	35	18	0	16	542	698	78	0	0	3	29	9	20
TC1	35	19	0	12	544	665	82	0	0	3	26	20	29
TC2	29	14	0	10	439	543	81	0	0	0	26	13	25
TC3	33	16	0	12	491	612	80	0	0	15	26	14	42
TC4	32	19	0	9	522	608	86	0	0	6	30	19	37
TC5	36	19	0	12	563	685	82	0	0	3	33	12	25

Supplementary Table ST 3.2 - Straw physicochemical properties of composite samples, from three chronosequences. Results from 19 measured or calculated variables

Samples	PA	PB	PC	SA	SB	SC	TA	TB	TC
Straw pH (CaCl ₂)	6.40	6.40	6.40	6.40	6.30	6.30	6.40	6.40	6.30
Straw density	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
Straw moisture	42.56	43.70	46.16	43.47	46.25	46.12	43.77	46.84	48.33
Straw TOM	46.69	48.38	52.45	39.02	50.98	58.89	39.36	62.26	80.30
Straw COM	44.14	41.56	45.30	32.94	44.07	49.93	30.22	55.51	72.77
Straw RCOM	2.55	6.82	7.15	6.09	6.90	8.96	9.14	6.75	7.53
Straw TC	25.94	26.89	29.14	21.69	28.32	32.72	21.87	34.59	44.61
Straw TOC	24.53	23.09	25.17	18.11	24.48	27.73	16.79	30.85	40.43
Straw TMR	53.31	51.62	47.55	60.98	49.02	41.11	60.64	37.74	19.70
Straw IMR	43.07	42.61	35.61	49.71	39.16	32.03	51.16	27.92	13.49
Straw SMR	10.24	9.01	11.94	11.27	9.86	9.08	9.48	9.82	6.21
Straw N-total	2.09	1.62	1.41	1.73	1.64	1.45	1.48	1.84	1.88
Straw P-total	0.12	0.16	0.24	0.16	0.17	0.20	0.16	0.23	0.23
Straw K-total	0.09	0.09	0.09	0.07	0.07	0.09	0.09	0.09	0.12
Straw Ca-total	1.34	1.24	1.23	1.20	1.23	1.32	1.39	1.47	1.51
Straw Mg-total	0.30	0.30	0.33	0.23	0.24	0.28	0.23	0.28	0.31
Straw S-total	0.10	0.12	0.07	0.11	0.06	0.11	0.11	0.08	0.12
Straw C-total:N-total ratio	12.41	16.60	20.67	12.54	17.27	22.57	14.78	18.80	23.73
Straw C-org:N-total ratio	11.74	14.25	17.85	10.47	14.93	19.12	11.34	16.77	21.51

Supplementary Table ST 3.3 – Enzymatic activity of bulk soil samples, from three chronosequences. Results from five measured enzymes

Sample	Urease	Dehydrogenase	Beta Glycosidase	FDA-hydrolysis	Acid phosphatase
PA1	7.6	2.1	18.8	85.0	87.3
PA2	15.4	2.4	15.9	76.4	76.6
PA3	5.4	1.6	12.1	88.1	55.8
PA5	8.8	2.1	10.7	80.8	61.2
PB1	3.6	1.6	14.3	68.0	86.2
PB2	6.8	2.5	29.4	87.6	90.7
PB3	3.9	3.0	11.1	53.5	98.3
PB5	14.8	1.5	7.0	91.2	96.5
PC1	11.2	1.8	14.2	81.5	98.0
PC2	11.7	3.9	13.7	103.4	76.3
PC3	9.7	1.2	15.0	69.0	72.4
PC5	6.8	2.1	13.2	73.4	70.6
SA1	11.0	5.3	27.6	107.9	58.1
SA2	14.7	4.0	26.9	103.3	52.5
SA3	11.2	3.3	39.0	103.7	69.7
SA5	17.9	4.6	33.8	108.1	66.1
SB1	22.1	3.4	35.7	95.0	66.5
SB2	23.0	8.0	44.1	143.2	76.4
SB3	10.5	4.1	45.4	107.2	68.4
SB5	14.1	2.6	29.1	96.7	68.2
SC1	18.5	5.9	47.8	86.5	60.0
SC2	21.0	4.2	38.9	111.5	67.2
SC3	15.7	3.4	35.9	93.1	61.8
SC5	15.8	4.9	34.6	103.7	79.9
TA1	9.9	5.6	94.2	96.6	42.0
TA2	9.0	4.0	92.2	74.6	60.0
TA3	5.6	3.2	77.0	75.5	61.6
TA5	6.0	4.5	105.1	68.4	48.0
TB1	7.4	3.5	116.7	102.7	58.7
TB2	13.4	7.0	128.1	84.6	49.2
TB3	5.4	6.7	127.0	88.0	38.8
TB5	10.0	6.7	124.6	126.3	58.7
TC1	7.5	6.7	134.7	80.1	64.8
TC2	6.9	6.9	129.8	73.1	57.9
TC3	10.0	5.2	140.3	98.0	86.5
TC5	4.2	4.8	110.0	93.9	57.9