Assemblage and functioning of bacterial communities in soil and rhizosphere

Yan, Y.

Assemblage and functioning of bacterial communities in soil and rhizosphere

Yan Yan

闫　燕

Assemblage and functioning of bacterial communities in soil and rhizosphere
PhD thesis, Leiden University, The Netherlands.

The research described in this thesis was performed at the Netherlands Institute of Ecology, NIOO-KNAW and at the Institute of Biology of Leiden University.

Cover (封面): Tree Roots (树根), *Vincent van Gogh* (梵高), 1890. Inspiration to rhizosphere.

Lay-out by Yan Yan
Printed by Ipskamp Printing
ISBN: 978-94-028-0205-4

Assemblage and functioning of bacterial communities in soil and rhizosphere

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof.mr.C.J.J.M.Stolker,

volgens besluit van het College voor Promoties

te verdedigen op woensdag 8 juni 2016

klokke 16.15 uur

door

Yan Yan

geboren te Shijiazhuang, Hebei, China

in 1986

Promotiecommissie

Promotor: Prof. dr. J. A. van Veen

Promotor: Prof. dr. P. G. L. Klinkhamer

Co-promotor: Dr. E. E. Kuramae, Nederlands Instituut voor Ecologie-KNAW

Overige leden

Prof. dr. H.P.Spaink

Prof. dr. G.A.Kowalchuk, Universiteit Utrecht

Prof. dr. Jos Raaijmakers

Prof. dr. Xavier Le Roux, Institut national de la recherche agronomique-Lyon, Frankrijk

Dr. Gera Hol, Nederlands Instituut voor Ecologie-KNAW

'God rewards the diligent'

天道酬勤

Confucian Analects

To my dear Parents and Yihao

献给我亲爱的父母和丈夫

Contents

Chapter 1

General introduction

Microbial communities in soil are extremely diverse and determine largely the functioning of terrestrial ecosystems (Philippot et al 2013). The soil microbiome is responsible for a range of key ecosystem functions such as decomposition of organic matter and of polluting compounds, and nutrient cycling (Nielsen et al 2011, Nielsen et al 2015). It also has a strong impact on above ground organisms, in particular plants.

Soil ecosystems consist of a wide variety of microenvironments with distinct characteristics of e.g. pH (Lauber et al 2009), salinity (Wichern et al 2006) and moisture (Brockett et al 2012). These specific habitats influence the resident (micro)organisms and recruit specific members of the soil microbiome, what is said to be one of the major reasons for the huge microbial biodiversity in soil (Fierer and Jackson 2006, Kuramae et al 2012).

Biodiversity has been considered as a critical factor influencing ecosystem processes (Butchart et al 2010). Recent studies on the soil microbiome trying to link biodiversity to ecosystems functioning have shown that, indeed, soil ecosystem functioning depends on microbial diversity (Tardy et al 2014, Wall et al 2015). In general, biodiversity is considered to be advantageous to ecosystem functioning because a more diverse ecosystem might be more resilient than a less diverse system (van Elsas et al 2012). Thus, there is a great challenge in assessing the immense diversity of microbial communities and their response to the environment and changes therein.

Several studies have shown that intensification of soil use by agriculture may reduce the diversity of soil organisms dramatically (de Vries et al 2013, Doran 1980, Garbeva et al 2004, Hartmann et al 2015). Consequently, this has triggered increasing concern that the loss of species in soil may impair ecosystems functions, such as nutrients acquisition by plants and resource recycling between above- and below- ground communities. However, the significance of microbial biodiversity loss is challenged by the concept of functional redundancy (Tilman et al 1997, Wall et al 2015) that assumes that (many) different microbial species can have the same functions in a natural ecosystem, and, therefore, the loss of species may not necessarily alter ecosystem functioning (Nannipieri et al 2003).

Until now, studies about the significance of soil microbial diversity are much less common than similar studies on aboveground macroorganisms especially plants (Martiny et al 2006). So there is a need to increase the studies on microbial diversity in different environments and changes therein and how that will impact the role of the microbiome in soil. Recently, the advent of high-throughput sequencing techniques has opened a promising avenue for improving our understanding of microbial biodiversity and community assemblage processes in specific habitats (Costello et al 2009, Franzosa et al 2015, Ofek-Lalzar et al 2014, Prosser 2015). The main goal of the study described in this thesis was to obtain better understanding of the structuring, diversity and functioning of bacterial communities in soil and and rhizosphere.

1.1 Soil and rhizosphere as habitat of microorganisms

1.1.1 Soil ecosystem

Soil ecosystems support a large trophic complexity within the soil food web with complex interactions with the abiotic environment (De Ruiter et al 1995, Hunt et al 1987). The structure and function of the soil food web is a primary indicator of ecosystem health because the huge functional diversity that exists among soil organisms is vital to ecosystem services, including plant growth and associated energy input to the system (Wall and Six 2015). For instance, there is mounting evidence that the soil food web, in particular the microbial component, can significantly modify belowground-aboveground interactions by regulating nutrient uptake, carbon storage and direct plant effects (Nielsen et al 2015), which in turn influence vegetation production.

The trophic interactions within the soil food web transfer energy between species and so drive ecosystem processes (Pimm et al 1991). It is widely known that microorganisms, such as bacteria and fungi, are mainly responsible for these ecosystem processes (Six et al 2006). Soil animals such as nematodes and detritivorous organisms mainly facilitate the microbial driven processes such as nutrient mineralization (Postma-Blaauw et al 2005, Setala and Huhta 1991). Furthermore, the complexity of the soil food web is essential to maintain the resistance and resilience against environmental changes within terrestrial ecosystems (de Vries et al 2012, de Vries et al 2013). Thus, activities that cause

loss of biodiversity belowground may contribute to a reduction in food web complexity and therefore in soil ecosystem functioning.

1.1.2 Rhizosphere

Since German agronomist and plant physiologist Lorenz Hiltner, more than a century ago, coined the term 'rhizosphere' as the soil compartment influenced by plant roots (Hiltner 1904), our understanding of this hot spot of biological activity in soil has advanced dramatically. Due to the continuous deposition of energy-rich substrate significant enrichment of organisms, mainly bacteria and other microbes, occurs in the rhizosphere. The release of organic carbon from roots into the surrounding soil, which is often termed rhizodeposition (Jones et al 2004), causes dramatic changes in both the biological and physicochemical nature of the soil. The wide range of rhizodeposition mechanisms includes loss of root caps and border cells, of mucilage, release of soluble root exudates, and of volatile organic carbon, as well as transfer of carbon to symbionts and leakage and destruction of root cells (Jones et al 2009). Numerous studies have revealed that rhizodeposition varies dramatically between different plant species, genotypes and plant growth stage (Bais et al 2006, Dennis et al 2010, Jones et al 2004). Consequently, soil microorganisms are confronted by and feed on a wide and dynamic range of rhizodeposits with strong implications for the structure and functioning of rhizosphere communities (de Boer et al 2006, Garbeva et al 2008, Kowalchuk et al 2002).

Many studies have demonstrated that the microbial communities in the rhizosphere are very different to those of the bulk soil (Costa et al 2006, Garbeva et al 2008, Smalla et al 2001). The dynamic processes between plant roots and microbes are a major force of the assembly and structuring of specific communities consisting of species that have the cellular properties to benefit from plant functioning. The root released materials are utilized by microbes as substrate for growth and energy supply (Vandenkoornhuyse et al 2007), but they may also be used for communication between roots and the biota in the rhizosphere (Kowalchuk et al 2006) or may act as antimicrobials for protection against pathogens (Bais et al 2006, Setala and Huhta 1991).

There is ample evidence that the significance of the rhizosphere microbiome is critical to health, productivity and overall conditions (Chaparro et al 2012, Mendes et al 2013, Ziegler et al 2013). There is a broad range of interactions between plant and microbes from beneficial to pathogenic involving a wide spectrum of microorganisms: arbuscular mycorrhizal fungi, plant growth promoting rhizobacteria (PGPR), endophytes and (minor) pathogens. The outcome of the interaction between plant and such microbes may vary depending on plant species, soil type and environmental conditions (van der Putten et al 2013). The primary goal of many plant associated microbiome studies is to drive this interaction towards enhanced benefits for plants by promoting beneficial organisms and reducing pathogens.

Studies concerning plant-associated microorganisms are strongly biased towards those individual species that are culturable and thus can easily be traced. These species comprise, however, only around 5% of the total microbial community, thus most of the important interactions are likely overlooked and remain unnoticed (Mark et al 2005, Matilla et al 2007). In order to be able to assess properly the dynamics of microbial communities in the bulk soil and the rhizosphere a comprehensive profile of the soil microbiome is required.

Yet, due to their enormous diversity, the detailed extent of microbial activities in certain niches, such as the rhizosphere, remains largely unclear. One of the major problems to assess the functioning of the microbiome in soil and rhizosphere is the lack of sound approaches to study microbial biodiversity experimentally.

1.2 Bacterial biodiversity

1.2.1 Experimental approaches to study microbial biodiversity

Understanding the assemblage of microbial communities and their associated diversity requires a broad range of approaches, including assessment of the community structure in nature (Lauber et al 2009), manipulation of natural communities in microcosm studies (Barthes et al 2015), and the assemblage of synthetic communities under controlled conditions (Lebeis et al 2015). Natural communities are dynamic and are the result of filtering processes exerted by

abiotic and biotic factors in the environment over a long period of time (Andrew et al 2012, de Ridder-Duine et al 2005). Most studies nowadays have focused on the richness and diversity of taxa to quantify and characterize soil microbial communities at different levels, from species (Mendes et al 2011, Sunagawa et al 2015, Tilman et al 1997) to family or phylum level (Fierer et al 2005). These studies are based often on manipulation of soil-borne microbial populations already present in nature by making changes in important ecological factors such as moisture and/or temperature conditions so to be able to explain the underlying mechanism in soil ecosystem functioning. Compared to the studies on natural communities, the studies on synthetic communities are technically feasible only with a group of species comprising a subset of the local species that are easy to propagate. A recent example of the assemblage of a synthetic bacterial community concerned the recolonization of previously sterilized Arabidopsis seed surfaces by easily cultural species (Lebeis et al 2015). However, as mentioned before, only a minority of bacteria has been cultured overall (Pedros-Alio 2006). Thus, it is difficult to extrapolate the result of culture-dependent experiments to the real situation in the field. By contrast, an approach that is often used to assess diversity effects is the so-called dilution method (Franklin et al 2001, Garland and Lehman 1999). The assumption on which this approach is based, is that the biodiversity of a soil microbial community can be reduced in comparison to the diversity of the original natural community after inoculation of sterilized soil with a diluted suspension of that soil and subsequent incubation until similar community size. This approach could provide information on assemblage processes closely related to natural processes. As reported in previous dilution to extinction studies (Franklin and Mills 2006, Garland and Lehman 1999), the regrown microbial communities are the result of assembly processes filtered by abiotic or biotic factors. Compared to the use of synthetic communities, the dilution approach allows for the assessment of the direct effects of the extinction of certain species on ecosystem functions.

Previous studies have reported that the reduction of microbial diversity did not change soil processes such as thymidine and leucine incorporation, nitrification and nitrate accumulation (Griffiths et al 2001). Similarly, the associated functioning groups, i.e. carbon mineralization, denitrification and nitrification did not change after reduction of microbial biodiversity affected by the dilution approach (Wertz et al 2006). However, other studies have indicated

that reduction of rare species modified plant-herbivore interactions (Hol et al 2010) and also that reduction of species diversity affected nitrogen cycling (Philippot et al 2013b). Questions that are claimed to be addressed by the dilution approach are, for instance, related to the significance of microbial diversity in natural ecosystems, to the role of rare species in ecosystem functioning, to the interactions among species or to the relationship between community structure and functionality in natural ecosystems. Referring to the last issue, many studies have addressed the taxonomic content of microbial communities in certain habitats and their interaction with aboveground vegetation within the terrestrial ecosystem (Hol et al 2010, Kardol et al 2006, Mendes et al 2011), but the functional genes involved in the plant microbe interactions have remained largely unclear (Mendes et al 2014).

These whole community approaches are feasible also because of the availability of advanced high throughput sequencing technologies. Recently, advances in next-generation DNA-based or RNA-based sequencing technologies have dramatically reduced costs and substantially increased capacity, resulting in an increasing number of comprehensive characterizations of microbial communities in different habitats. (Bulgarelli et al 2015, Lebeis et al 2015, Lundberg et al 2013, Ofek-Lalzar et al 2014). These metagenome surveys based on phylogenetic gene marker amplicon and/or total DNA not only detected microorganisms and their genetic diversity associated with environmental parameters, but also shed light on the functional attributes that microbes enable to perform important contributions to ecosystems (Mendes et al 2014). Thus, next generation sequencing technologies, together with advanced bioinformatics tools to process the huge data sets may provide new insights into the microbial life in natural ecosystem. In this thesis, I combined sensitive molecular approaches including qPCR and next generation sequencing with the old dilution methodology for deciphering the mechanisms of microbial community assemblage in soil and rhizosphere and the functions relevant for plant-microbe interactions.

1.2.2. Functionality of biodiversity

Although several reports in literature provide information that soil microbial communities are linked to ecosystem functioning (Allison et al 2013,

Vogelsang et al 2006), the common opinion is that there is a large functional redundancy within the soil microbiome preponderating the relationship between diversity and functionality (Nannipieri et al 2003). Functional redundancy across different taxa in microbial communities has been suggested as a buffering capacity for biodiversity loss in the ocean (Sunagawa et al 2015). This is consistent with another study on the human gut in which taxonomical composition of the microbiome varied markedly between bacterial groups but gene abundances were evenly distributed (Yu et al 2012). Together this suggests that ecosystem functioning may be independent of the composition of microbial communities.

The analysis of complex microbial communities has been generally limited by technical approaches and sequencing depth. The advanced developments in high throughput sequencing methods by using 16S rRNA gene marker for bacterial species, such as the Roche 454 pyrosequencing technology (Caporaso et al 2011) and the Illumina technology (Caporaso et al 2012), have facilitated comprehensive surveys of the breadth of microbial communities. In addition to the provision of a profile of microbial taxonomical diversity, such data may also help us to understand the significance of functional genes, for instance those related to plant host colonization.

There is ample evidence that the rhizosphere microbiome influences directly and indirectly the composition and productivity (i.e. biomass) of plant communities in terrestrial ecosystems (Schnitzer et al 2011, van der Heijen 2008). Well known examples are the plant growth promoting rhizobacteria which promote plant growth directly by either facilitating nutrient acquisition (Hawkins et al 2000, Miransari 2011) or by inducing plant hormone levels (Zamioudis and Pieterse 2012), or indirectly by suppressing soil- borne plant pathogens (Mendes et al 2011). In contrast, earlier studies have documented that enemy accumulation could cause negative plant-soil feedbacks processes (Bever 1994, Bulgarelli et al 2015, van der Putten et al 2013). However, there is also evidence that plants may recruit beneficial microorganisms in the rhizosphere to get protection against the pathogens presence (Hawkins et al 2000). For example, a study on sugar beet plants attacked by the root pathogen *Rhizoctonia solani* revealed that plants exploited the soil microbiome for protection against pathogen infections (Mendes et al 2011). The ability to progressively enrich for beneficial microbiota is more effective in successive

generations of plants (Miransari 2011). Hence, the diversity and richness of microorganisms in the rhizosphere have been regarded as frontline for plant development in the terrestrial ecosystems.

Previous studies on plant-microbe interactions and the significance of microbial biodiversity to assess the effects of species loss on plant performance, showed that species reduction, and in particular the loss of rare species, reduced plant biomass production (Hol et al 2010). In contrast, as different species may have the same functions related to plant functioning (Nannipieri et al 2003), the reduction of microbial diversity may not influence plant performance. So, in order to understand the fundamental ecological mechanisms of plant-soil feedbacks, we need to better characterize the assemblage as well as the functionality of the microbiome in soil and rhizosphere.

In this study, I used *Jacobaea vulgaris* as model plant for the assessment of the impact of the rhizosphere microbial community on plant growth. *Jacobaea vulgaris*, is one of the most common weeds in the Netherlands (van der Meijden and van der Waalskooi 1979). *Senecio* spp. usually grows in dry and poor soils, and is considered to be toxic weeds for livestock of cattle and horses when ingested (Hartmann 1999). In a previous study, Joosten *et al*. (2009) found a strong negative plant-soil feedback when a sterilized soil was inoculated with its 'own' microbial community compared to plant growth on sterile soil. This negative feedback effect has been also shown in other plant-soil feedback studies, which showed that *Jacobaea vulgaris* biomass in natural soil from different chronosequence fields was lower than in sterilized soil (van de Voorde et al 2012).

1.3 Aim and thesis outline

The aim of this study was to obtain a better understanding of the assemblage of bacterial communities in soil and rhizosphere and the significance of microbial diversity and functional traits for plant-soil feedback processes. The main approach to the assemblage and biodiversity studies was the so-called dilution approach in which the diversity of the bacterial community in soil is manipulated by inoculation and incubation of more or less diluted soil suspensions in pre-sterilized soils.

The main research questions addressed here are:

1) Does the dilution approach reduce the diversity of the bacterial community after inoculation and subsequent incubation of soil suspensions in soil? If so, does soil has a selective power during the assemblage process of bacterial communities? (Chapters 2 and 3)
2) How do the taxonomical and functional diversity of the bacterial community change in the rhizosphere compared to the bulk soil and what are the relative roles of soil and plants in the assembly process? (Chapter 4)
3) What are the main bacterial functional traits that determine the relationship between the bacterial community in soil and plant growth? (Chapter 5)

The experimental approaches to assess the relevance of soil bacterial diversity for the functioning of soil ecosystems are scarce; the dilution method is among the most frequently used. In **Chapter 2**, I revisited this method and shed more light on the assemblage of bacterial communities in soil.

In **Chapter 3** the concept of soil as the driving force for the structuring of the bacterial community was further analyzed by assessing the assemblage of bacterial communities in different soils after inoculation of more or less diluted suspensions in a cross inoculation design.

To further understand the development of the bacterial community in the rhizosphere, *Jacobaea vulgaris* was planted in the incubated soils that were inoculated with the same more or less diluted suspensions as in chapter 2. In **Chapter 4**, I tested which bacterial functional traits were most selected in the rhizosphere. Based on the 16S rRNA gene amplicon sequencing as well as the shotgun metagenome approach the bacterial community composition and its functionality in soil and rhizosphere were compared.

In **Chapter 5**, I report on a study on the impact of the bacterial community in the rhizosphere on plant biomass production linking both species composition and its functionality to plant growth.

The results of the different studies are discussed and summarized in **Chapter 6**.

1.4. References

Allison SD, Lu Y, Weihe C, Goulden ML, Martiny AC, Treseder KK *et al* (2013). Microbial abundance and composition influence litter decomposition response to environmental change. *Ecology* **94:** 714-725.

Andrew DR, Fitak RR, Munguia-Vega A, Racolta A, Martinson VG, Dontsova K (2012). Abiotic factors shape microbial diversity in sonoran desert soils. *Appl Environ Microb* **78:** 7527-7537.

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM (2006). The role of root exudates in rhizosphere interactions with plants and other organisms. *Annual review of plant biology* **57:** 233-266.

Barthes A, Ten-Hage L, Lamy A, Rols JL, Leflaive J (2015). Resilience of aggregated microbial communities aubjected to drought-small-scale studies. Microb Ecol 70: 9-20.

Bever JD (1994). Feedback between plants and their soil communities in an old field community. *Ecology* **75:** 1965-1977.

Brockett BFT, Prescott CE, Grayston SJ (2012). Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. *Soil Biol Biochem* **44:** 9-20.

Bulgarelli D, Garrido-Oter R, Munch PC, Weiman A, Droge J, Pan Y *et al* (2015). Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* **17:** 392-403.

Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW, Almond REA *et al* (2010). Global biodiversity: indicators of recent declines. *Science* **328:** 1164-1168.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ *et al* (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* **108 Suppl 1:** 4516-4522.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N *et al* (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *Isme J* **6:** 1621-1624.

Chaparro JM, Sheflin AM, Manter DK, Vivanco JM (2012). Manipulating the soil microbiome to increase soil health and plant fertility. *Biol Fert Soils* **48:** 489-499.

Costa R, Gotz M, Mrotzek N, Lottmann J, Berg G, Smalla K (2006). Effects of site and plant species on rhizosphere community structure as revealed by molecular analysis of microbial guilds. *Fems Microbiol Ecol* **56:** 236-249.

Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R (2009). Bacterial community variation in human body habitats across space and time. *Science* **326:** 1694-1697.

de Boer W, Kowalchuk GA, van Veen JA (2006). 'Root-food' and the rhizosphere microbial community composition. *New Phytol* **170:** 3-6.

de Ridder-Duine AS, Kowalchuk GA, Gunnewiek PJAK, Smant W, van Veen JA, de Boer W (2005). Rhizosphere bacterial community composition in natural stands of *Carex arenaria* (sand sedge) is determined by bulk soil community composition. *Soil Biol Biochem* **37:** 349-357.

de Vries FT, Liiri ME, Bjornlund L, Bowker MA, Christensen S, Setala HM *et al* (2012). Land use alters the resistance and resilience of soil food webs to drought. *Nat Clim Change* **2:** 276-280.

de Vries FT, Thebault E, Liiri M, Birkhofer K, Tsiafouli MA, Bjornlund L *et al* (2013). Soil food web properties explain ecosystem services across European land use systems. *P Natl Acad Sci USA* **110:** 14296-14301.

Dennis PG, Miller AJ, Hirsch PR (2010). Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *Fems Microbiol Ecol* **72:** 313-327.

De Ruiter PC, Neutel AM, Moore JC (1995). Energetics, patterns of interaction strengths, and stability in real ecosystems. *Science* **269:** 1257-1260.

Doran JW (1980). Soil microbial and biochemical-changes associated with reduced tillage. *Soil Sci Soc Am J* **44:** 765-771.

Fierer N, Jackson JA, Vilgalys R, Jackson RB (2005). Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Appl Environ Microb* **71:** 4117-4120.

Fierer N, Jackson RB (2006). The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* **103:** 626-631.

Franklin RB, Garland JL, Bolster CH, Mills AL (2001). Impact of dilution on microbial community structure and functional potential: Comparison of numerical simulations and batch culture experiments. *Appl Environ Microb* **67:** 702-712.

Franklin RB, Mills AL (2006). Structural and functional responses of a sewage microbial community to dilution-induced reductions in diversity. *Microbial Ecology* **52:** 280-288.

Franzosa EA, Hsu T, Sirota-Madi A, Shafquat A, Abu-Ali G, Morgan XC *et al* (2015). Sequencing and beyond: integrating molecular 'omics' for microbial community profiling. *Nature Reviews Microbiology* **13:** 360-372.

Garbeva P, van Veen JA, van Elsas JD (2004). Microbial diversity in soil: selection of microbial populations by plant and soil type and implications for disease suppressiveness. *Annu Rev Phytopathol* **42:** 243-270.

Garbeva P, van Elsas JD, van Veen JA (2008). Rhizosphere microbial community and its response to plant species and soil history. *Plant Soil* **302:** 19-32.

Garland JL, Lehman RM (1999). Dilution/extinction of community phenotypic characters to estimate relative structural diversity in mixed communities. *Fems Microbiol Ecol* **30:** 333-343.

Griffiths BS, Ritz K, Wheatley R, Kuan HL, Boag B, Christensen S *et al* (2001). An examination of the biodiversity-ecosystem function relationship in arable soil microbial communities. *Soil Biol Biochem* **33:** 1713-1722.

Hartmann M, Frey B, Mayer J, Mader P, Widmer F (2015). Distinct soil microbial diversity under long-term organic and conventional farming. *Isme J* **9:** 1177-1194.

Hartmann T (1999). Chemical ecology of pyrrolizidine alkaloids. *Planta* **207:** 483-495.

Hawkins HJ, Johansen A, George E (2000). Uptake and transport of organic and inorganic nitrogen by arbuscular mycorrhizal fungi. *Plant Soil* **226:** 275-285.

Hiltner L (1904). Über neuere erfahrungen und probleme auf dem gebiete der bodenbakteriologie unter besonderder berücksichtigung der gründüngung und brache. *Arb. Dtsch. Landwirtsch Gesellschaft* **98:** 59-78.

Hol WHG, de Boer W, Termorshuizen AJ, Meyer KM, Schneider JHM, van Dam NM *et al* (2010). Reduction of rare soil microbes modifies plant-herbivore interactions. *Ecol Lett* **13:** 292-301.

Hunt HW, Coleman DC, Ingham ER, Ingham RE, Elliott ET, Moore JC *et al* (1987). The detrital food web in a shortgrass prairie. *Biol Fert Soils* **3:** 57-68.

Jones DL, Hodge A, Kuzyakov Y (2004). Plant and mycorrhizal regulation of rhizodeposition. *New Phytol* **163:** 459-480.

Jones DL, Nguyen C, Finlay RD (2009). Carbon flow in the rhizosphere: carbon trading at the soil-root interface. *Plant Soil* **321:** 5-33.

Joosten L, Mulder PPJ, Klinkhamer PGL, van Veen JA (2009). Soil-borne microorganisms and soil-type affect pyrrolizidine alkaloids in *Jacobaea vulgaris*. *Plant and Soil* **325:** 133-143.

Kowalchuk GA, Buma DS, de Boer W, Klinkhamer PG, van Veen JA (2002). Effects of above-ground plant species composition and diversity on the diversity of soil-borne microorganisms. *Antonie Van Leeuwenhoek* **81:** 509-520.

Kowalchuk GA, Hol WHG, Van Veen JA (2006). Rhizosphere fungal communities are influenced by Senecio jacobaea pyrrolizidine alkaloid content and composition. *Soil Biol Biochem* **38:** 2852-2859.

Kuramae EE, Yergeau E, Wong LC, Pijl AS, van Veen JA, Kowalchuk GA (2012). Soil characteristics more strongly influence soil bacterial communities than land-use type. *Fems Microbiol Ecol* **79:** 12-24.

Lauber CL, Hamady M, Knight R, Fierer N (2009). Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil Bacterial Community Structure at the Continental Scale. *Appl Environ Microb* **75:** 5111-5120.

Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald M *et al* (2015). Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* **349:** 860-864.

Lundberg DS, Yourstone S, Mieczkowski P, Jones CD, Dangl JL (2013). Practical innovations for high-throughput amplicon sequencing. *Nat Methods* **10:** 999-1002.

Mark GL, Dow JM, Kiely PD, Higgins H, Haynes J, Baysse C *et al* (2005). Transcriptome profiling of bacterial responses to root exudates identifies genes involved in microbe-plant interactions. *P Natl Acad Sci USA* **102:** 17454-17459.

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL *et al* (2006). Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4:** 102-112.

Matilla MA, Espinosa-Urgel M, Rodriguez-Herva JJ, Ramos JL, Ramos-Gonzalez MI (2007). Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. *Genome Biology* **8**.

Mendes LW, Kuramae EE, Navarrete AA, van Veen JA, Tsai SM (2014). Taxonomical and functional microbial community selection in soybean rhizosphere. *Isme J* **8:** 1577-1587.

Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al* (2011). Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* **332:** 1097-1100.

Mendes R, Garbeva P, Raaijmakers JM (2013). The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *Fems Microbiol Rev* **37:** 634-663.

Miransari M (2011). Arbuscular mycorrhizal fungi and nitrogen uptake. *Arch Microbiol* **193:** 77-81.

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

Nielsen UN, Ayres E, Wall DH, Bardgett RD (2011). Soil biodiversity and carbon cycling: a review and synthesis of studies examining diversity-function relationships. *Eur J Soil Sci* **62:** 105-116.

Nielsen UN, Wall DH, Six J (2015). Soil Biodiversity and the Environment. *Annu Rev Env Resour* **40:** 63-90.

Ofek-Lalzar M, Sela N, Goldman-Voronov M, Green SJ, Hadar Y, Minz D (2014). Niche and host-associated functional signatures of the root surface microbiome. *Nat Commun* **5**.

Pedros-Alio C (2006). Marine microbial diversity: can it be determined? *Trends Microbiol* **14:** 257-263.

Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH (2013a). Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews Microbiology* 11: 789-799.

Philippot L, Spor A, Henault C, Bru D, Bizouard F, Jones CM *et al* (2013b). Loss in microbial diversity affects nitrogen cycling in soil. *Isme J* **7:** 1609-1619.

Postma-Blaauw MB, de Vries FT, de Goede RGM, Bloem J, Faber JH, Brussaard L (2005). Within-trophic group interactions of bacterivorous nematode species and their effects on the bacterial community and nitrogen mineralization. *Oecologia* **142:** 428-439.

Prosser JI (2015). Dispersing misconceptions and identifying opportunities for the use of 'omics' in soil microbial ecology. *Nature Reviews Microbiology* **13:** 439-446.

Schnitzer SA, Klironomos JN, HilleRisLambers J, Kinkel LL, Reich PB, Xiao K *et al* (2011). Soil microbes drive the classic plant diversity-productivity pattern. *Ecology* **92:** 296-303.

Setala H, Huhta V (1991). Soil Fauna Increase Betula-Pendula Growth - Laboratory Experiments with Coniferous Forest Floor. *Ecology* **72:** 665-671.

Six J, Frey SD, Thiet RK, Batten KM (2006). Bacterial and fungal contributions to carbon sequestration in agroecosystems. *Soil Sci Soc Am J* **70:** 555-569.

Smalla K, Wieland G, Buchner A, Zock A, Parzy J, Kaiser S *et al* (2001). Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: Plant-dependent enrichment and seasonal shifts revealed. *Appl Environ Microb* **67:** 4742-4751.

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G *et al* (2015). Structure and function of the global ocean microbiome. *Science* **348**.

Tardy V, Mathieu O, Leveque J, Terrat S, Chabbi A, Lemanceau P *et al* (2014). Stability of soil microbial structure and activity depends on microbial diversity. *Env Microbiol Rep* **6:** 173-183.

Tilman D, Knops J, Wedin D, Reich P, Ritchie M, Siemann E (1997). The influence of functional diversity and composition on ecosystem processes. *Science* **277:** 1300-1302.

van de Voorde TFJ, van der Putten WH, Bezemer TM (2012). The importance of plant-soil interactions, soil nutrients, and plant life history traits for the temporal dynamics of Jacobaea vulgaris in a chronosequence of old-fields. *Oikos* **121:** 1251-1262.

van der Heijen MGA (2008). The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems (vol 11, pg 296, 2008). *Ecol Lett* **11:** 651-651.

van der Putten WH, Bardgett RD, Bever JD, Bezemer TM, Casper BB, Fukami T *et al* (2013). Plant-soil feedbacks: the past, the present and future challenges. *J Ecol* **101:** 265-276.

van Elsas JD, Chiurazzi M, Mallon CA, Elhottova D, Kristufek V, Salles JF (2012). Microbial diversity determines the invasion of soil by a bacterial pathogen. *P Natl Acad Sci USA* **109:** 1159-1164.

van den Koornhuyse P, Mahe S, Ineson P, Staddon P, Ostle N, Cliquet JB *et al* (2007). Active root-inhabiting microbes identified by rapid incorporation of plant-derived carbon into RNA. *Proc Natl Acad Sci U S A* **104:** 16970-16975.

van der Meijden E, van der Waalskooi RE (1979). Population ecology of senecio-Jacobaea in a Sand Dune System .1. reproductive strategy and the biennial habit. *J Ecol* **67:** 131-153.

Vogelsang KM, Reynolds HL, Bever JD (2006). Mycorrhizal fungal identity and richness determine the diversity and productivity of a tallgrass prairie system. *New Phytol* **172:** 554-562.

Wall DH, Nielsen UN, Six J (2015). Soil biodiversity and human health. *Nature* **528:** 69-76.

Wall DH, Six J (2015). Give soils their due. *Science* **347:** 695-695.

Wertz S, Degrange V, Prosser JI, Poly F, Commeaux C, Freitag T et al (2006). Maintenance of soil functioning following erosion of microbial diversity. *Environ Microbiol* 8: 2162-2169.

Wichern J, Wichern F, Joergensen RG (2006). Impact of salinity on soil microbial communities and the decomposition of maize in acidic soils. *Geoderma* **137:** 100-108.

Yu Y, Wang H, Liu J, Wang Q, Shen TL, Guo WH *et al* (2012). Shifts in microbial community function and structure along the successional gradient of coastal wetlands in Yellow River Estuary. *European Journal of Soil Biology* **49:** 12-21.

Zamioudis C, Pieterse CMJ (2012). Modulation of host immunity by beneficial microbes. *Mol Plant Microbe* In 25: 139-150.

Ziegler M, Engel M, Welzl G, Schloter M (2013). Development of a simple root model to study the effects of single exudates on the development of bacterial community structure. *J Microbiol Meth* **94:** 30-36.

Chapter 2

Revisiting the dilution procedure used to manipulate microbial biodiversity in terrestrial systems

Yan Yan, Eiko E. Kuramae, Peter G.L. Klinkhamer, Johannes A. van Veen

Applied and Environmental Microbiology (2015)

Abstract

It is hard to assess experimentally the importance of microbial diversity in soil for the functioning of terrestrial ecosystems. An approach that is often used to make such assessment is the so-called dilution method. This method is based on the assumption that the biodiversity of the microbial community is reduced after dilution of a soil suspension and that the reduced diversity persists after incubation of more or less diluted inocula in soil. However, little is known how the communities develop in soil after inoculation. In this study, serial dilutions of a soil suspension were made and reinoculated into the original soil previously sterilized by γ-irradiation. We determined the structure of the microbial communities in the suspensions and the inoculated soils using 454-pyrosequencing of 16S rRNA genes. Upon dilution, several diversity indices showed that, indeed, the diversity of the bacterial communities in the suspensions reduced dramatically, with *Proteobacteria* as the dominant phylum of bacteria detected in all dilutions. The structure of the microbial community was changed considerably in soil with *Proteobacteria*, *Bacteroidetes* and *Verrucomicrobia* as the dominant groups in most diluted samples, indicating the importance of soil- related mechanisms operating in the assembly of the communities. We found unique operational taxonomic units (OTUs) even in the highest dilution both in the suspensions and in the incubated soil samples. We conclude that the dilution approach reduces the diversity of microbial communities in soil samples but that it does not allow for accurate predictions on the community assemblage during incubation of (diluted) suspensions in soil.

2.1. Introduction

The significance of biodiversity for terrestrial ecosystem processes continues to be a matter of much debate (Sala et al 2000, Magurran and Henderson 2003, Butchart et al 2010). Compared to the importance of plants and animals, the role of microbial biodiversity is still poorly understood. This lack of knowledge is of great concern as soil microbes, particularly bacteria, represent the major source of biodiversity in terrestrial ecosystems and are known to carry out numerous essential ecosystem functions, including nutrient cycling and facilitating plant nutrition (Philippot et al 2013).

The biggest obstacle to a better understanding of the importance of microbial biodiversity for the functioning of terrestrial ecosystems is the lack of sound experimental approaches to make directed and predictable changes in the diversity of microbial communities in soil. One of the most interesting approaches so far is the so-called dilution method. This method involves the inoculation of sterilized soils with more or less diluted inocula derived from suspensions of the same soil (Salonius 1981, Garland and Lehman 1999, Franklin et al 2001, Griffiths et al 2001, Matos et al 2005, Franklin and Mills 2006, Wertz et al 2006, Hol et al 2010, Philippot et al 2013, Vivant et al 2013). However, previous studies were often limited by the depth and extent of the analytical methodology applied and focused only on the structure of the microbial community after regrowth in the soil. As a consequence, they do not provide information about the community from which the different communities after incubation originated and the process of community assemblage. Therefore, these studies do not allow testing of the assumption that dilution mainly influences the diversity through the reduction of the number of the less abundant, rare, species. In reality, rare species in the original community may have become common after incubation or vice versa.

High-throughput next-generation sequencing technologies have allowed researchers to use deeper sampling depths by providing large numbers of reads by cost-effective means to detect microbial phylogenetic diversity (Margulies et al 2005). This has provided new insights into the details of microbial communities in natural ecosystems (Sogin et al 2006, Huber et al 2007, Neufeld et al 2008) and human body (Turnbaugh et al 2008). One of the exciting possibilities provided by this technology is the ability to estimate accurately the

assembly processes and structure of microbial communities, including the long tail of less abundant microbes, that is evident in graphs of relative abundances of microbial species, which may lead to a better understanding of the relevance of microbial biodiversity in soil.

The major aim of this study was to determine the changes and the associated variation in the composition of a soil microbial community brought about by inoculation of serial dilutions of suspensions of that soil and to detect how the microbial community structure develops during regrowth in soil. This analysis will allow evaluation of the suitability of the dilution approach as a tool for the manipulation of microbial biodiversity and for the separation of rare from abundant species. It will also lead to a better understanding of the selective pressure of the soil environment on the assembly of microbial communities. We addressed three basic questions: 1) does the dilution procedure reduce the diversity of the microbial community after inoculation and subsequent incubation of soil suspensions in soil? 2) does the composition of the microbial community change during incubation in soil? 3) is the dilution procedure effective in separating more and less abundant species so to allow an assessment of their specific roles? In order to answer these questions, we established a range of microbial communities through the inoculation of serial dilutions of microbial suspensions from nonsterilized soil samples into the same soil after sterilization.

2.2. Materials and methods

2.2.1. Soil sampling and treatment

Thirty liters of soil was collected at a depth of around 15 cm from dune sandy soil in Meijendel, The Netherlands (52°9'N, 4°22'E). Soil organic matter content (%) was 9.11 ± 0.36 (n=6), soil pH was 7.4 ± 0.005 (n=6), NO_3^- content (mg/kg) was 30.43 ± 0.85 (n=6), NH_4^+ content (mg/kg) was 2.23 ± 0.25 (n=6), P content (mg/kg) was 15.16 ± 0.41 (n=6). The soil had a sandy texture, with more than 99% of the grains greater than 75 µm. The soil was sieved, homogenized and aliquots of 500 g were stored in plastic bags (Whirl-pak sampling bag, 720 ml, Sigma-aldrich). The bags containing soil were gamma irradiation sterilized (> 25 kGray, Isotron, Ede, the Netherlands). One bag of

soil was kept separately to serve as inoculum. Sterility was checked by spreading 0.5 g sterilized soil from the inoculum bag onto TSA and PDA media. No bacterial and fungal growth on agar plates for 6 replicates was observed in the sterilized soil samples after 6 days. Three gamma irradiation sterilized soil bags were inoculated with autoclaved demineralized water to be used as control. A subsample of the fresh soil was taken to determine soil moisture (24 h, 105 ^{o}C).

Soil suspensions for inoculation were made by mixing 20 g fresh soil and 190 ml autoclaved demineralized water with a blender for 2 minutes. This procedure was repeated 3 times and in between the blender was cooled down on ice for 2 minutes. This was called the 10^{-1} dilution. 100 ml of 10^{-1} dilution was transferred to a bottle containing 900 ml of autoclaved demineralized water and followed the bottle shaking by hands for 1 min. This procedure was repeated for several times until 10^{-6} and 10^{-9} dilutions were made. Subsequently, 25 ml of each dilution were added to 500 g of soil in the bags, and additional autoclaved demineralized water was added to bring the moisture level of the inoculated soil at around 20%, which is roughly similar to the average level at the prevailing climatic conditions at the side from where the soil was taken. In total, 39 bags of soil (i.e. six replicated samples of three dilutions in duplicates plus three controls) were used. We kept the six replicate samples (and the duplicates) per dilution separated throughout the experiment in order to be able to assess the variance caused by the dilution procedure. The remaining suspensions were centrifuged at 3,000 g for 10 min at 4 ^{o}C, and the pellets were stored at -20 ^{o}C for further analysis. After inoculation, the soil bags were incubated at 20 ^{o}C using sterilized cotton plug caps to ensure gas exchange. The soils were turned over regularly once a week to homogenize microbial growth. The aim was to reach similar microbial abundances in the different dilution treatments. After 8 weeks of incubation under laminar flow conditions, soil samples were taken to determine the microbial abundance in all treatments by quantitative real time PCR (qPCR) using Eub 338 (Lane 1991) and Eub 518 (Muyzer et al 1993) primer set for 16S rRNA gene. Total DNA was extracted from the incubated soil using the MoBio Power Soil Extraction Kit according to the supplier's instructions. Each 25 µl reaction consisted of 12.5 µl Sybr green mix (Bioline, GC-Biotech) with 4 mg/ml bovine serum albumin (BSA) in a total volume of 25 µl, 5 µM of each primer, 5 µl template DNA (5 ng/µl). For bacteria, the standard curves were generated using 10-fold dilution series from 10^{8} to 10^{3} of

plasmid DNA. PCRs were run on a Rotor-Gene 3000 (Qiagen) and started with 15 min at 95 ^{o}C, followed by 40 amplification cycles each of 95 ^{o}C for 60 sec, 53 ^{o}C 50 sec and 72 ^{o}C 60 sec. A subsample of soil from each bag was stored at -20 ^{o}C for further analysis. Triplicate reaction mixtures per DNA sample and the appropriate set of standards were used. For qPCR assays, there was a linear relationship between the log of the plasmid DNA copy number and the calculated threshold cycle (C_T value). PCR efficiencies were 99%, and correlation coefficients (R^2) for standard curves were 0.99. Bacterial abundance was similar for all dilution treatments after 8 weeks of incubation as determined by quantitative real time PCR (Fig. 2.1). I also measured fungal abundance by quantitative real-time PCR using the primer of 5.8S and internal transcribed spacer 1 (ITS1) genes. For fungi, the standard curves were generated using 10-fold dilution series from 10^8 to 10^3 of plasmid DNA obtained from fungi. Because of the difficulties in assessing fungal abundance by quantitative real-time PCR due to heterogeneity in ribosomal operon number per fungal species/phylum, we decided to ignore the fungal community in the rest of our analyses. The primers we used for pyrosequencing target both bacteria and archaea. There were no significant numbers of archaea sequences; therefore we did not include archaea in our analyses.

2.2.2. DNA extraction, PCR reaction and 16S rRNA gene fragment pyrosequencing

Total DNA was extracted from the soil suspensions and from incubated soil to determine the composition of the respective microbial communities by 454-pyrosequencing. DNA was extracted using the MoBio Power Soil Extraction Kit according to the supplier's manual (MO BIO Laboratories, Carlsbad, CA, USA). Total DNA concentration was qualified on ND-1000 spectrophotometer (Nanodrop Technology, Wilmington, DE). For DNA concentrations below 5 ng/µl, *i.e.* five samples of 10^{-6} and four samples of 10^{-9} suspension, nested PCR was performed. The general bacterial primer 27F and 1492R (Lane 1991) were used for the first

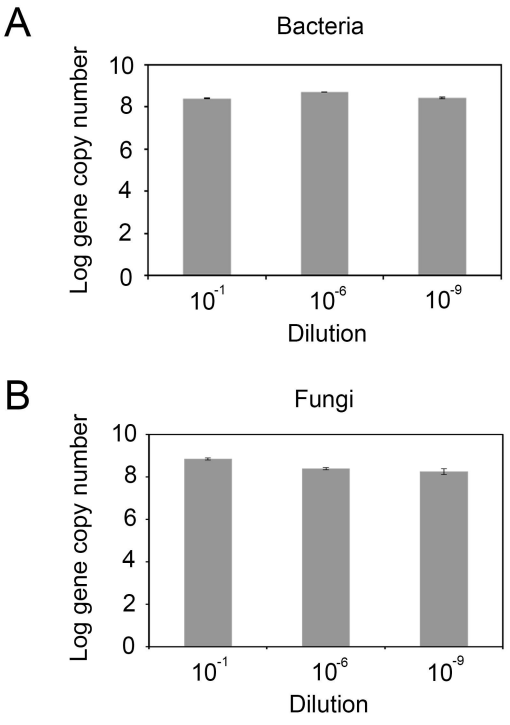

Figure 2.1. Bacterial abundance after incubation among dilution treatments as estimated by real time PCR (Mean ± SE, n=5).

amplification, and then 2 µl of the amplified products from the first round was used as template for the second round PCR using barcoded primers 515F and 806R (Bergmann et al 2011). Five ng/µl of DNA/sample of the diluted samples was used as template for the first round of nested PCR with the PCR program of 95 ^{o}C for 5 min followed by 25 cycles each of 95 ^{o}C for 30 s, 55 ^{o}C 1 min and 72 ^{o}C 10 min. For PCR reactions using barcoded primers were performed using 5 µM of each forward (515F) and reverse (806R) primers, 5 mM dNTPs (Invitrogen, Carlsbad, CA), 1 unit of *Taq* polymerase (Roche, Indianapolis, IN), and 5 ng/µl of sample DNA as the template in a total volume of 25 µl. The PCR conditions for the barcoded primer were similar to the first PCR round except for 25 cycles with 52 ^{o}C annealing temperature. To control for contamination during PCR preparation, one negative control (water in place of DNA) was included for all PCR reactions. PCR products of each subsample from the barcoded primers were generated in six replicates and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega). Equimolar purified

PCR products that were quantified by picogreen assays were mixed and sequenced using Roche Genome Sequencer FLX Titanium 454 sequencing platform (Macrogen, Seoul, South Korea).

2.2.3. Data analysis

The raw sequence data were processed using the QIIME v.1.6.0 pipeline (Caporaso et al 2010). Low quality sequences less than 150 bp in length or with average quality score of less than 25 were removed. After denoising the sequences using Denoiser, version 0.91 (Reeder and Knight 2010), and checking for chimeras using USEARCH, Operational Taxonomic Units (OTUs) were identified using the UCLUST 1.2.21 algorithm (Edgar 2010) with a phylotype defined at the 97% sequence similarity level. The resulting OTUs were aligned against the Ribosomal Database Project (RDP) database (Cole et al 2009).

Alpha diversity calculation was performed based on the rarefied OTU table to compare the diversity among samples at the given level of a sampling effort (Hughes and Hellmann 2005). The OTU table was rarefied to 1,535 reads by "single rarefaction" QIIME script since this number was the lowest number of reads for all samples. The average reads from the three sterilized controls were used as baseline that was subtracted from the reads of the other 36 samples. The OTU table after subtraction of the control was used for further statistical analysis. Chao1 richness and Simpson and Shannon diversity and evenness indices were determined with the "vegan" package (Dixon 2003) in R (The R Foundation for Statistical Computing). The percentage of coverage was calculated by Good's method using the following formula: $\% = [1-(n/N)] \times 100$, where n means the number of phylotypes represented by singletons and N is the total number of sequences (Good 1953).

To compare the communities from the different dilution treatments, Nonmetric MultiDimensional Scaling (NMDS) plots were used to visualize the structure among samples at genus level. The plots were generated from Bray-Curtis similarity index matrices of all samples. NMDS was calculated by using the PAST software (Hammer et al 2001).

2.3. Results

2.3.1. Effect of dilution and incubation on bacterial community diversity

Several indices were used to assess the diversity in the soil suspension dilutions and in the associated soil communities after incubation on the basis of OTU detection (Table 2.1). Remarkably, all indices for the diluted inocula of 10^{-6} and 10^{-9} were significantly higher after incubation than the indices of the associated suspensions, while the indices were lower for the 10^{-1} dilution after incubation in soil. Good's estimator of coverage increased with increasing dilution, indicating that microbial species were lost through dilution.

Table 2.1. Estimators of sequence library diversity, evenness and coverage in soil suspensions at three dilutions and the related samples after incubation in soil.

Treatment	Dilution	S.obs	S.chao-1	Shannon	Simpson	Evenness	Good's estimator of coverage
Suspension	10^{-1}	131.00±3.27	169.15±7.80	3.986±0.036	0.966±0.002	0.41±0.01	97.61±0.12
Soil	10^{-1}	107.20±1.27	134.37±2.96	3.719±0.019	0.954±0.002	0.38±0.01	97.56±0.11
P		*	*	*	*	*	
Suspension	10^{-6}	44.80±7.98	53.09±10.33	2.383±0.416	0.774±0.124	0.24±0.07	99.32±0.18
Soil	10^{-6}	70.09±2.13	89.64±4.46	3.208±0.040	0.934±0.004	0.35±0.01	97.95±0.21
P		*	*	*	NS	*	
Suspension	10^{-9}	17.00±2.17	19.54±2.46	1.462±0.293	0.623±0.128	0.25±0.05	99.77±0.03
Soil	10^{-9}	55.83±1.14	81.82±3.37	2.633±0.042	0.867±0.006	0.25±0.01	97.27±0.24
P		*	*	*	*	NS	

Estimators were calculated for each dilution treatment of soil suspensions (n = 5-6) and incubated soil samples (n = 11-12) as well as significant comparisons ($P < 0.05$) among phylogenetic profile (species level). S.obs is the observed number of OTUs. NS means not significant.

2.3.2. Effect of dilution and incubation on bacterial community composition

After the OTUs were classified according to the RDP database, the soil microbial community consisted of 18 phyla (Fig. 2.2). Phylum-level taxonomic assignments indicated that *Proteobacteria*, followed by *Actinobacteria*, *Bacteroidetes*, *Acidobacteria*, *Verrucomicrobia*, *Planctomycetes* and *Firmicutes* dominated the microbial communities in the original non-diluted (10^{-1}) soil suspension (> 90% of all sequences). The variance in the abundance of the seven dominating phyla among the replicated suspension samples increased from the low-dilution treatments to the high dilution treatments (Table 2.2 and

Fig. 2.3). The same was true for the incubated samples while in general the variance among the replicates of the incubated samples was lower than the variance among the replicated samples of the soil suspensions (Table 2.2 and Fig. 2.4).

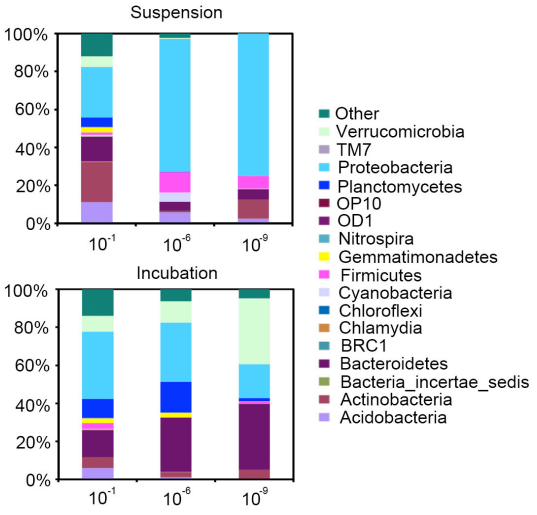

Figure 2.2. Bacterial community composition at phylum level of soil suspensions and incubated soil samples in relative abundances.

To test the selective power of soil further, we analyzed the major phyla at the family level. Visually, we noticed that the diversity of the communities in the incubated soil samples, and in particular, those which were incubated with the highest, i.e. 10^{-9} dilution, suspensions differed strongly from the diversity of the inoculated suspension (Fig. 2.5). Thus, we compared the diversity of the communities in both the suspensions and the inoculated soil samples. Remarkably, for most phyla we found that the Shannon diversity index was significantly higher in the incubated soil samples than in the corresponding suspensions of the 10^{-6} and 10^{-9} dilution (Table 2.3), while they were lower in the soil samples than in the associated suspensions at the 10^{-1} dilutions.

Table 2.2. Coefficients of variation (%) for each phylum measured in soil suspensions at the three dilution levels and in related soil samples after incubation

	Suspension			Soil		
Dilution	10^{-1}	10^{-6}	10^{-9}	10^{-1}	10^{-6}	10^{-9}
Proteobacteria	13.94	46.66	26.66	6.05	20.73	50.69
Actinobacteria	24.95	146.26	205.76	10.97	60.29	77.59
Bacteroidetes	10.03	107.73	192.35	20.30	20.24	31.30
Acidobacteria	29.84	67.97	222.38	13.31	43.77	65.31
Verrucomicrobia	39.95	188.99	223.61	26.50	28.35	43.43
Planctomycetes	39.94	121.61	-	24.56	32.27	107.52
Firmicutes	14.97	64.43	135.96	31.49	109.87	96.17

The table depicts the coefficient of variation (CV) of each phylum based on absolute reads in soil suspensions and incubated soil samples. CV (%) = Standard deviation/mean*100. "-" data are not present.

Table 2.3. Shannon diversity of major phyla

Phylum	Suspension	Soil	*P*	Suspension	Soil	*P*	Suspension	Soil	*P*
	10^{-1}			10^{-6}			10^{-9}		
Acidobacteria	1.60±0.04	1.16±0.04	*	0.73±0.17	0.85±0.07	NS	0.00±0.00	0.54±0.15	*
Actinobacteria	2.49±0.03	2.34±0.03	*	0.96±0.25	1.78±0.07	*	0.01±0.01	1.46±0.16	*
Bacteroidetes	1.49±0.05	1.29±0.04	*	0.61±0.18	1.27±0.06	*	0.00±0.00	1.16±0.07	*
Firmicutes	1.17±0.05	1.04±0.04	NS	0.94±0.27	0.23±0.12	*	0.22±0.19	0.52±0.11	NS
Verrucomicrobia	1.15±0.07	1.23±0.03	NS	0.07±0.07	0.96±0.06	*	0.00±0.00	0.81±0.07	*
Alphaproteobacteria	1.95±0.01	1.88±0.02	NS	1.05±0.19	1.69±0.04	*	0.66±0.07	1.37±0.12	*
Betaproteobacteria	1.39±0.02	1.50±0.03	NS	0.65±0.21	0.75±0.14	NS	0.41±0.12	0.91±0.08	*
Deltaproteobacteria	1.16±0.03	1.31±0.08	NS	0.02±0.02	0.78±0.13	*	0.00±0.00	0.74±0.09	*
Gammaproteobacteria	1.18±0.02	0.94±0.04	*	0.79±0.09	0.95±0.07	NS	0.33±0.16	0.47±0.11	NS

Diversity was calculated for each dilution of soil suspensions (n = 5-6) and incubated soil samples (n = 11-12) as well as the level of significance ($P < 0.05$) for each major phylum based on the phylogenetic profile at the family level. NS means not significant.

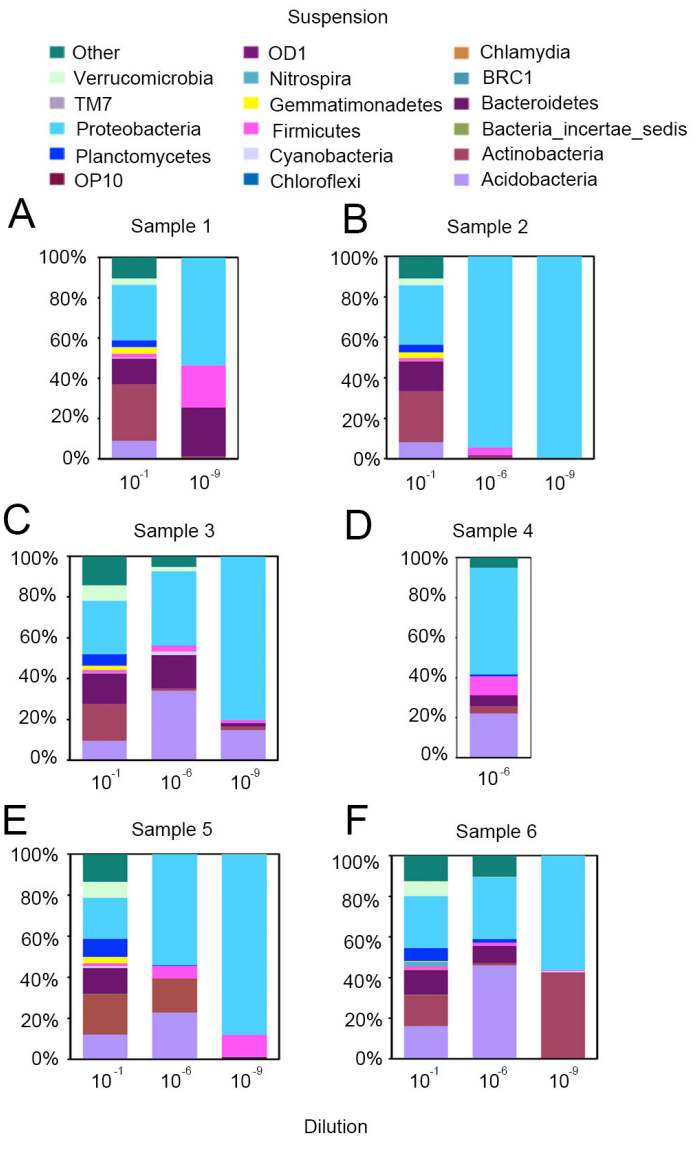

Figure 2.3. Bacterial community composition at phylum level from six samples of soil suspensions. Dilution level is shown below each bar. 10^{-1} and 10^{-9} dilutions of sample 4 are not available due to technical issues.

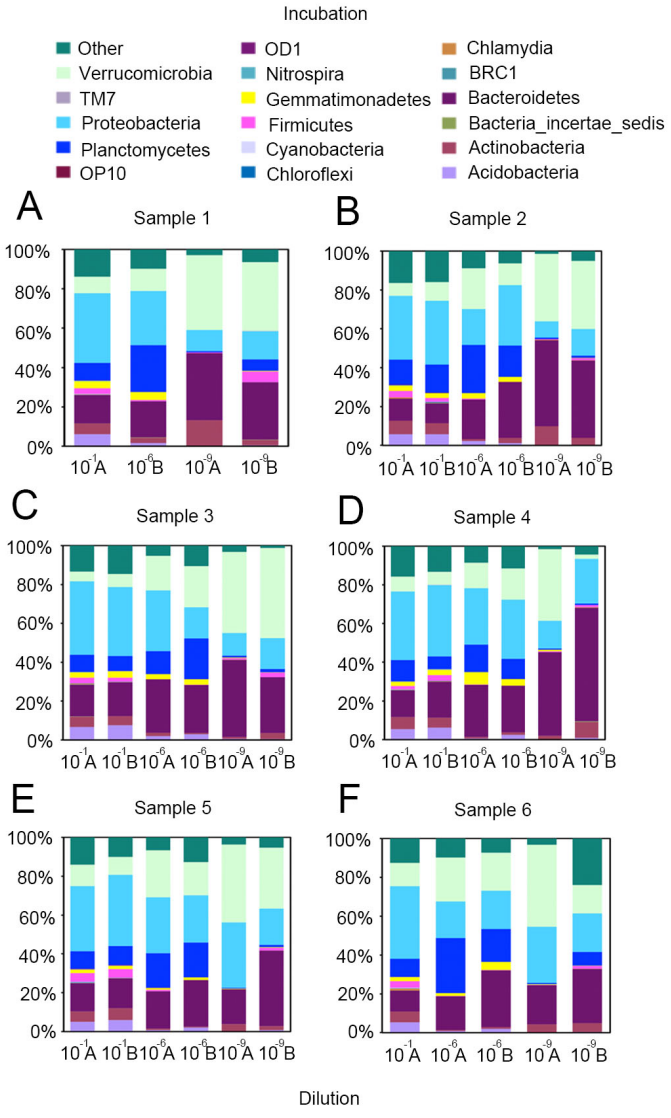

Figure 2.4. Bacterial community composition at phylum level of six replicate samples of the incubated soil. Dilution level is shown below each bar. A, B indicates the duplicates from the same dilution level. 10^{-1}B, 10^{-6}A of sample 1 and 10^{-1}B of sample 6 are not available due to technical issues.

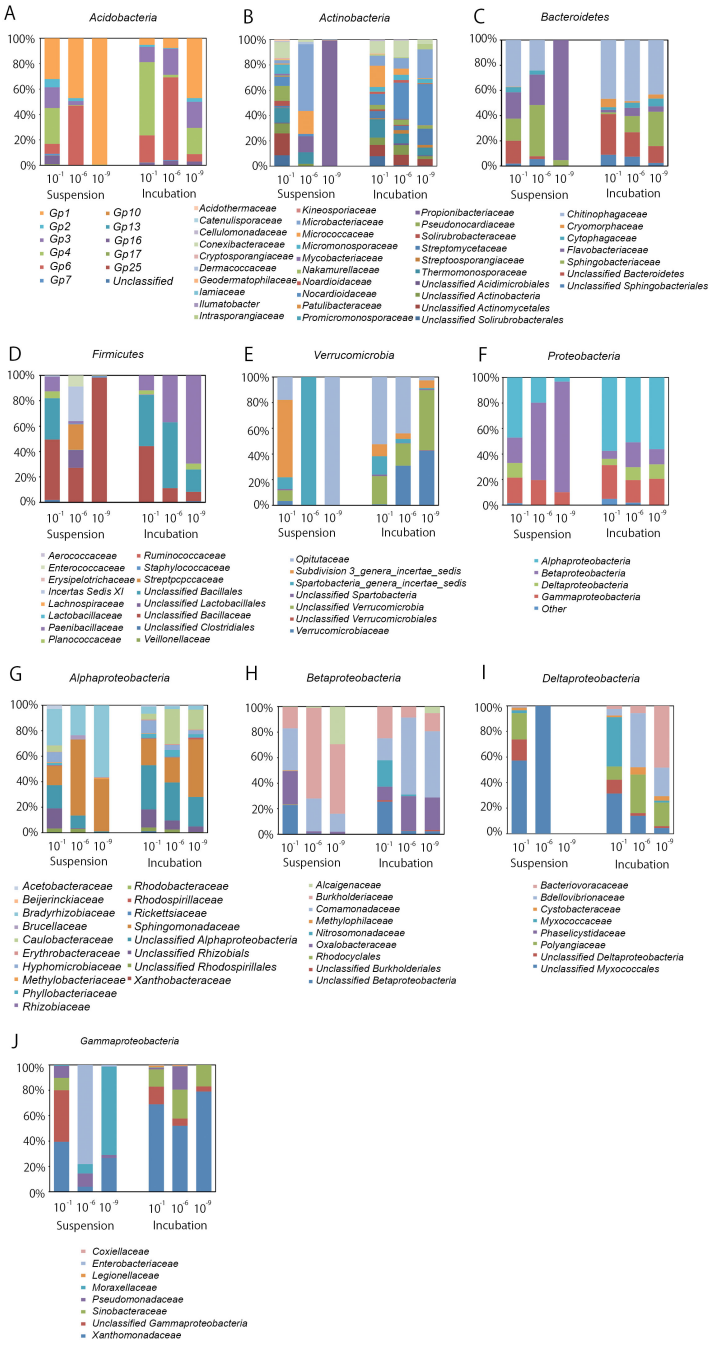

Figure 2.5. Bacterial community composition at the family level of soil suspensions and incubated soil samples.

To compare the overall community structure of the different dilution treatments and differences before/after incubation, the taxonomic abundance profiles were used to compute Bray-Curtis similarity matrix, coordinated into two dimensions by using NMDS (Fig. 2.6). Samples were grouped according to before/after incubation. This analysis revealed clear differences in the microbial community structure between before and after incubation. The community structures of the soil samples after incubation were more similar to each other than to the associated suspension samples. This may hint to selective processes in the soil leading to more equal communities. One-way analysis of similarities (ANOSIM) showed that the dilution treatment had a significant ($R = 0.28$, $P < 0.001$) overall effect on the structure of the bacterial community in the suspension and the soil samples after incubation.

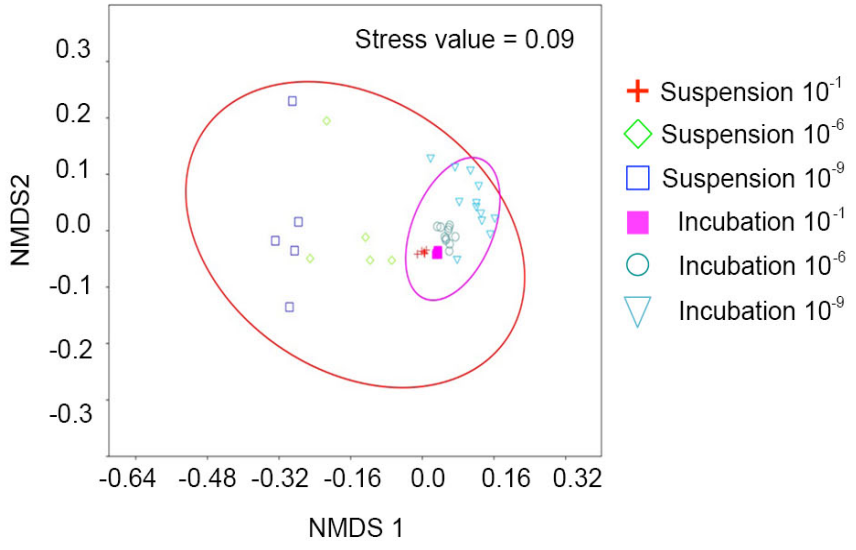

Figure 2.6. NMDS of Bray-Curtis similarity matrix among soil suspensions and incubated soil samples.

2.3.3. Effect of dilution on rare/abundant OTUs

A possibility to determine if the dilution approach is appropriate to separate rare species from abundant ones is to make Venn diagrams to assess the shared and

unique OTUs between dilution treatments in the soil suspensions (Fig. 2.7) and incubated soil (Fig. 2.7). We found 954, 77 and 10 unique OTUs in the 10^{-1}, 10^{-6} and 10^{-9} dilution samples of the soil suspensions, respectively, and 386, 96 and 88 unique OTUs in the respective dilution treatments of the incubated soil samples. To identify the unique OTUs in the different treatments, the phylogenetic affiliation was done at the genus level. From the unique OTUs that were assigned to the genus, a total of 158, 38, 10 unique genera were detected in 10^{-1}, 10^{-6} and 10^{-9} dilutions of soil suspensions, respectively (Table 2.4) and 84, 33 and 34 unique genera were detected in 10^{-1}, 10^{-6} and 10^{-9} of the incubated soil samples, respectively (Table 2.5).

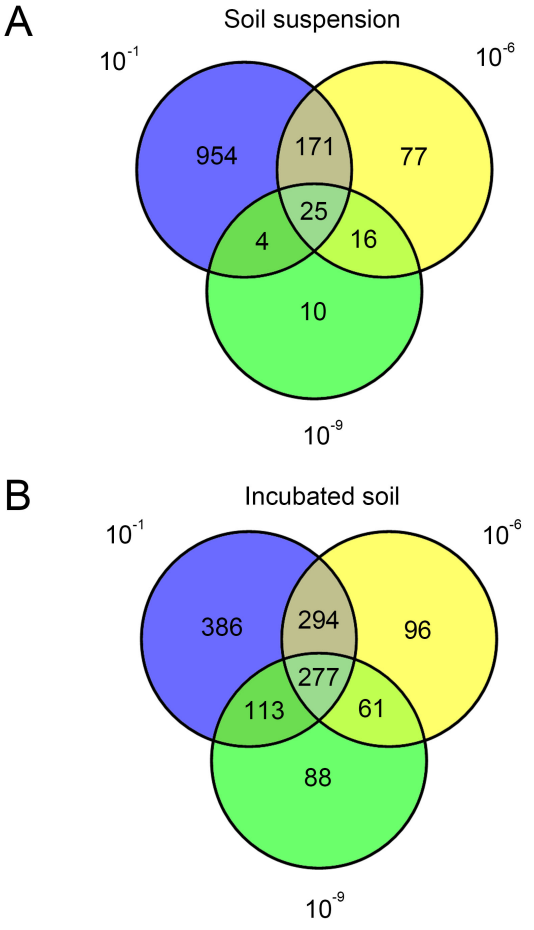

Figure 2.7. Venn diagram of shared and unique OTUs in each dilution of (A) soil suspensions and (B) incubated soil samples.

Table 2.4. The phylogenetic affiliation of the unique OTUs in the six replicate samples of the three dilutions of a soil suspension.

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-1	Other	Other	Other	Other	Other	0.0984	0.0083
10^-1	Acidobacteria	Other	Other	Other	Other	0.0003	0.0003
10^-1	Acidobacteria	Acidobacteria_Gp1	Gp1	Other	Other	0.0063	0.002
10^-1	Acidobacteria	Acidobacteria_Gp10	Gp10	Other	Other	0.0001	0.0001
10^-1	Acidobacteria	Acidobacteria_Gp13	Gp13	Other	Other	0.0001	0.0001
10^-1	Acidobacteria	Acidobacteria_Gp16	Gp16	Other	Other	0.0005	0.0005
10^-1	Acidobacteria	Acidobacteria_Gp17	Gp17	Other	Other	0.0007	0.0002
10^-1	Acidobacteria	Acidobacteria_Gp2	Gp2	Other	Other	0.0012	0.0005
10^-1	Acidobacteria	Acidobacteria_Gp3	Gp3	Other	Other	0.007	0.0011
10^-1	Acidobacteria	Acidobacteria_Gp4	Gp4	Other	Other	0.0081	0.0044
10^-1	Acidobacteria	Acidobacteria_Gp6	Gp6	Other	Other	0.0054	0.0017
10^-1	Acidobacteria	Acidobacteria_Gp7	Gp7	Other	Other	0.0007	0.0003
10^-1	Actinobacteria	Actinobacteria	Other	Other	Other	0.0093	0.0006
10^-1	Actinobacteria	Actinobacteria	Acidimicrobiales	Other	Other	0.0013	0.0004
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Other	Other	0.02	0.0032
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Acid Othermaceae	Acid Othermus	0.0022	0.0007
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Catenulisporaceae	Catenulispora	0.0004	0.0003
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Cellulomonadaceae	Cellulomonas	0.0014	0.0005
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Cryptosporangiaceae	Cryptosporangium	0.0013	0.0002
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Geodermatophilaceae	Other	0.0013	0.0005
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Geodermatophilaceae	Modestobacter	0.0001	0.0001
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	Phycicoccus	0.0009	0.0003
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Kineosporiaceae	Other	0.0004	0.0003
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Other	0.005	0.0011
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Leucobacter	0.0001	0.0001
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Micromonosporaceae	Other	0.0134	0.0016
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Micromonosporaceae	Actinoplanes	0.0005	0.0002
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Micromonosporaceae	Rugosimonospora	0.0014	0.0007
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nakamurellaceae	Humicoccus	0.0016	0.0005
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nocardiaceae	Nocardia	0.0001	0.0001
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nocardiaceae	Rhodococcus	0.0003	0.0003
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Other	0.0017	0.0004
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Actinopolymorpha	0.0003	0.0002
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Aeromicrobium	0.0036	0.0006

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Kribbella	0.0046	0.0012
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides	0.0009	0.0004
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Promicromonosporaceae	Promicromonospora	0.0004	0.0003
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae	Other	0.0042	0.0008
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae	Amycolatopsis	0.0096	0.0019
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae	Kutzneria	0.0008	0.0003
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae	Pseudonocardia	0.0104	0.003
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Streptosporangiaceae	Other	0.0007	0.0004
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Streptosporangiaceae	Nonomuraea	0.0004	0.0002
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Streptosporangiaceae	Streptosporangium	0.0001	0.0001
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Thermomonosporaceae	Other	0.0154	0.0012
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Thermomonosporaceae	Actinoallomurus	0.0091	0.0013
10^-1	Actinobacteria	Actinobacteria	Solirubrobacterales	Other	Other	0.0043	0.0005
10^-1	Actinobacteria	Actinobacteria	Solirubrobacterales	Conexibacteraceae	Conexibacter	0.0067	0.0014
10^-1	Actinobacteria	Actinobacteria	Solirubrobacterales	Solirubrobacteraceae	Solirubrobacter	0.0062	0.0023
10^-1	Bacteria_incertae_sedis	Ktedonobacteria	Ktedonobacterales	Ktedonobacteraceae	Ktedonobacter	0.0004	0.0002
10^-1	Bacteroidetes	Other	Other	Other	Other	0.0149	0.0025
10^-1	Bacteroidetes	Flavobacteria	Flavobacteriales	Cryomorphaceae	Fluviicola	0.0003	0.0002
10^-1	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Other	0.0004	0.0003
10^-1	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Chryseobacterium	0.0003	0.0002
10^-1	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Flavobacterium	0.002	0.0007
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Other	Other	0.0022	0.0004
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	Other	0.0148	0.0019
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	Ferruginibacter	0.0012	0.001
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	Flavisolibacter	0.0003	0.0002
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	Sediminibacterium	0.0032	0.0006
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	Terrimonas	0.0012	0.0003
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Cytophagaceae	Other	0.0003	0.0002
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Cytophagaceae	Cytophaga	0.0041	0.0012
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Cytophagaceae	Dyadobacter	0.0005	0.0002
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Cytophagaceae	Spirosoma	0.0005	0.0004
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Sphingobacteriaceae	Other	0.0001	0.0001
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Sphingobacteriaceae	Mucilaginibacter	0.0014	0.0004
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Sphingobacteriaceae	Pedobacter	0.0004	0.0003

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-1	BRC1	BRC1_genera_incertae_sedis	Other	Other	Other	0.0003	0.0002
10^-1	Chlamydiae	Chlamydiae	Chlamydiales	Other	Other	0.0008	0.0003
10^-1	Chlamydiae	Chlamydiae	Chlamydiales	Parachlamydiaceae	Other	0.0004	0.0002
10^-1	Chlamydiae	Chlamydiae	Chlamydiales	Parachlamydiaceae	Neochlamydia	0.0001	0.0001
10^-1	Chlamydiae	Chlamydiae	Chlamydiales	Simkaniaceae	Simkania	0.0003	0.0002
10^-1	Chloroflexi	Other	Other	Other	Other	0.0001	0.0001
10^-1	Chloroflexi	Anaerolineae	Anaerolineales	Anaerolineaceae	Other	0.0004	0.0002
10^-1	Cyanobacteria	Cyanobacteria	Chloroplast	Bacillariophyta	Other	0.0001	0.0001
10^-1	Cyanobacteria	Cyanobacteria	Chloroplast	Chlorophyta	Other	0.0004	0.0003
10^-1	Cyanobacteria	Cyanobacteria	Family I	GpI	Other	0.0001	0.0001
10^-1	Cyanobacteria	Cyanobacteria	Family V	GpV	Other	0.0001	0.0001
10^-1	Firmicutes	Bacilli	Bacillales	Other	Other	0.0009	0.0002
10^-1	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0.0003	0.0002
10^-1	Firmicutes	Bacilli	Bacillales	Bacillaceae	Oceanobacillus	0.0001	0.0001
10^-1	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Other	0.0001	0.0001
10^-1	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Ammoniphilus	0.0001	0.0001
10^-1	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Cohnella	0.0003	0.0002
10^-1	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus	0.0011	0.0004
10^-1	Firmicutes	Clostridia	Clostridiales	Other	Other	0.0003	0.0002
10^-1	Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales	Gemmatimonadaceae	Gemmatimonas	0.0261	0.0015
10^-1	Nitrospira	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	0.0029	0.0005
10^-1	OD1	OD1_genera_incertae_sedis	Other	Other	Other	0.0001	0.0001
10^-1	OP10	OP10_genera_incertae_sedis	Other	Other	Other	0.0005	0.0002
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Other	0.0345	0.0077
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Blastopirellula	0.0001	0.0001
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Gemmata	0.0033	0.0014
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Pirellula	0.0044	0.0011
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Planctomyces	0.0005	0.0002
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Schlesneria	0.0005	0.0002
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Singulisphaera	0.0032	0.0013
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Zavarzinella	0.002	0.0006
10^-1	Proteobacteria	Other	Other	Other	Other	0.0037	0.0006
10^-1	Proteobacteria	Alphaproteobacteria	Other	Other	Other	0.0173	0.0009
10^-1	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Other	0.0001	0.0001

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-1	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Asticcacaulis	0.0005	0.0002
10^-1	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Caulobacter	0.0008	0.0004
10^-1	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Phenylobacterium	0.0041	0.0011
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Other	Other	0.008	0.0012
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Other	0.0025	0.0006
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Devosia	0.0001	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Hyphomicrobium	0.0016	0.0006
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Prosthecomicrobium	0.0001	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	Methylobacterium	0.0003	0.0003
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	Microvirga	0.0004	0.0003
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Mesorhizobium	0.0008	0.0004
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	0.0003	0.0002
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	Other	0.0001	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	Labrys	0.0004	0.0003
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Other	Other	0.0037	0.0009
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	Other	0.0022	0.0004
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	Acidisoma	0.0005	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	Roseomonas	0.0001	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	Stella	0.0001	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Other	0.0004	0.0003
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Inquilinus	0.0001	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Erythrobacteraceae	Other	0.0003	0.0002
10^-1	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Erythrobacteraceae	Porphyrobacter	0.0001	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Other	0.0013	0.0006
10^-1	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	0.0005	0.0003
10^-1	Proteobacteria	Betaproteobacteria	Other	Other	Other	0.004	0.0008
10^-1	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	0.0007	0.0003
10^-1	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Other	0.0034	0.0008
10^-1	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Other	0.0003	0.0002
10^-1	Proteobacteria	Betaproteobacteria	Methylophilales	Methylophilaceae	Methylophilus	0.0003	0.0002
10^-1	Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae	Other	0.0001	0.0001
10^-1	Proteobacteria	Deltaproteobacteria	Other	Other	Other	0.005	0.0012
10^-1	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bacteriovoracaceae	Peredibacter	0.0001	0.0001
10^-1	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bdellovibrionaceae	Bdellovibrio	0.0003	0.0002

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-1	Proteobacteria	Deltaproteobacteria	Myxococcales	Other	Other	0.0062	0.0013
10^-1	Proteobacteria	Deltaproteobacteria	Myxococcales	Cystobacteraceae	Other	0.0007	0.0002
10^-1	Proteobacteria	Deltaproteobacteria	Myxococcales	Myxococcaceae	Corallococcus	0.0007	0
10^-1	Proteobacteria	Deltaproteobacteria	Myxococcales	Polyangiaceae	Other	0.0034	0.0005
10^-1	Proteobacteria	Deltaproteobacteria	Myxococcales	Polyangiaceae	Byssovorax	0.0017	0.0002
10^-1	Proteobacteria	Deltaproteobacteria	Myxococcales	Polyangiaceae	Chondromyces	0.0011	0.0002
10^-1	Proteobacteria	Gammaproteobacteria	Other	Other	Other	0.0034	0.0009
10^-1	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Alkanindiges	0.0004	0.0003
10^-1	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Cellvibrio	0.0008	0.0006
10^-1	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Sinobacteraceae	Alkanibacter	0.0011	0.0003
10^-1	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Sinobacteraceae	Nevskia	0.0003	0.0003
10^-1	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Sinobacteraceae	Steroidobacter	0.0011	0.0002
10^-1	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Other	0.0092	0.0022
10^-1	Verrucomicrobia	Other	Other	Other	Other	0.0047	0.0007
10^-1	Verrucomicrobia	Opitutae	Opitutales	Opitutaceae	Other	0.0018	0.0007
10^-1	Verrucomicrobia	Opitutae	Opitutales	Opitutaceae	Opitutus	0.0084	0.0021
10^-1	Verrucomicrobia	Spartobacteria	Other	Other	Other	0.0007	0.0003
10^-1	Verrucomicrobia	Spartobacteria	Spartobacteria_genera_incertae_sedis	Other	Other	0.0061	0.0011
10^-1	Verrucomicrobia	Subdivision3	Subdivision3_genera_incertae_sedis	Other	Other	0.005	0.0006
10^-1	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Other	0.0013	0.0003
10^-1	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Prosthecobacter	0.0005	0.0001
10^-1	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Verrucomicrobium	0.0001	0.0001
10^-6	Other	Other	Other	Other	Other	0.0258	0.0151
10^-6	Acidobacteria	Other	Other	Other	Other	0.0001	0.0001
10^-6	Acidobacteria	Acidobacteria_Gp1	Gp1	Other	Other	0.0022	0.0022
10^-6	Acidobacteria	Acidobacteria_Gp3	Gp3	Other	Other	0.0013	0.0008
10^-6	Acidobacteria	Acidobacteria_Gp4	Gp4	Other	Other	0.0014	0.0009
10^-6	Acidobacteria	Acidobacteria_Gp6	Gp6	Other	Other	0.0007	0.0005
10^-6	Bacteroidetes	Other	Other	Other	Other	0.0026	0.0026
10^-6	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Flavobacterium	0.0001	0.0001
10^-6	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	Other	0.006	0.006
10^-6	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Sphingobacteriaceae	Pedobacter	0.0001	0.0001
10^-6	Cyanobacteria	Cyanobacteria	Chloroplast	Bacillariophyta	Other	0.0025	0.0025
10^-6	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus	0.0022	0.0022
10^-6	Firmicutes	Bacilli	Lactobacillales	Other	Other	0.0004	0.0004

Table 2.5. The phylogenetic affiliation of the unique OTUs in the soil samples after incubation of three dilutions of the soil suspension dilutions of a soil suspension.

	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-6	Firmicutes	Bacilli	Lactobacillales	Aerococcaceae	Aerococcus	0.0003	0.0003
10^-6	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	0.0013	0.0011
10^-6	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Lactococcus	0.0014	0.0014
10^-6	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	0.0007	0.0007
10^-6	Firmicutes	Clostridia	Clostridiales	Veillonellaceae	Veillonella	0.0003	0.0003
10^-6	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Other	0.0007	0.0007
10^-6	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Pirellula	0.0017	0.0017
10^-6	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Singulisphaera	0.0004	0.0004
10^-6	Proteobacteria	Other	Other	Other	Other	0.0033	0.0031
10^-6	Proteobacteria	Alphaproteobacteria	Other	Other	Other	0.0001	0.0001
10^-6	Proteobacteria	Alphaproteobacteria	Rhizobiales	Other	Other	0.0005	0.0005
10^-6	Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	Other	0.0004	0.0004
10^-6	Proteobacteria	Alphaproteobacteria	Rhizobiales	Brucellaceae	Pseudochrobactrum	0.0003	0.0003
10^-6	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	0.0001	0.0001
10^-6	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	Other	0.001	0.001
10^-6	Proteobacteria	Betaproteobacteria	Burkholderiales	Alcaligenaceae	Other	0.0004	0.0003
10^-6	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Janthinobacterium	0.0001	0.0001
10^-6	Proteobacteria	Deltaproteobacteria	Other	Other	Other	0.0001	0.0001
10^-6	Proteobacteria	Gammaproteobacteria	Other	Other	Other	0.0012	0.0005
10^-6	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Other	0.0005	0.0005
10^-6	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Other	0.0001	0.0001
10^-6	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Dyella	0.0001	0.0001
10^-6	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Stenotrophomonas	0.0004	0.0004
10^-6	TM7	TM7_genera_incertae_sedis	Other	Other	Other	0.0001	0.0001
10^-6	Verrucomicrobia	Spartobacteria	Spartobacteria_genera_incertae_sedis	Other	Other	0.0003	0.0003
10^-9	Actinobacteria	Actinobacteria	Actinomycetales	Dermacoccaceae	Dermacoccus	0.0004	0.0004
10^-9	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Vagococcus	0.0003	0.0003
10^-9	Firmicutes	Erysipelotrichi	Erysipelotrichales	Erysipelotrichaceae	Turicibacter	0.0001	0.0001
10^-9	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	Methylobacterium	0.0003	0.0003
10^-9	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingopyxis	0.0001	0.0001
10^-9	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingopyxis	0.0004	0.0004
10^-9	Proteobacteria	Betaproteobacteria	Burkholderiales	Alcaligenaceae	Alcaligenes	0.005	0.005
10^-9	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Herbaspirillum	0.0012	0.001
10^-9	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	0.0001	0.0001
10^-9	Verrucomicrobia	Opitutae	Opitutales	Opitutaceae	Opitutus	0.0001	0.0001

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-1	other	other	other	other	other	0.1403	0.0059
10^-1	Acidobacteria	other	other	other	other	0.0002	0.0001
10^-1	Acidobacteria	Acidobacteria_Gp1	Gp1	other	other	0.0031	0.0004
10^-1	Acidobacteria	Acidobacteria_Gp3	Gp3	other	other	0.0072	0.0011
10^-1	Acidobacteria	Acidobacteria_Gp4	Gp4	other	other	0.0346	0.0027
10^-1	Acidobacteria	Acidobacteria_Gp6	Gp6	other	other	0.0128	0.0017
10^-1	Acidobacteria	Acidobacteria_Gp7	Gp7	other	other	0.0002	0.0001
10^-1	Actinobacteria	Actinobacteria	other	other	other	0.0032	0.0005
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	other	other	0.0050	0.0009
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Cryptosporangiaceae	Cryptosporangium	0.0001	0.0001
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Geodermatophilaceae	other	0.0003	0.0003
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Arthrobacter	0.0094	0.0011
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Micromonosporaceae	other	0.0021	0.0005
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nocardiaceae	Nocardia	0.0003	0.0002
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	other	0.0004	0.0001
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae	other	0.0013	0.0004
10^-1	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae	Amycolatopsis	0.0002	0.0002
10^-1	Actinobacteria	Actinobacteria	Solirubrobacterales	other	other	0.0044	0.0007
10^-1	Actinobacteria	Actinobacteria	Solirubrobacterales	Conexibacteraceae	Conexibacter	0.0055	0.0007
10^-1	Bacteroidetes	other	other	other	other	0.0459	0.0041
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	other	other	0.0128	0.0034
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	other	0.0553	0.0037
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	Ferruginibacter	0.0014	0.0004
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	Flavisolibacter	0.0001	0.0001
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	Terrimonas	0.0085	0.0007
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Cytophagaceae	other	0.0002	0.0001
10^-1	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Cytophagaceae	Larkinella	0.0001	0.0001
10^-1	BRC1	BRC1_genera_incertae_sedis	other	other	other	0.0009	0.0002
10^-1	Chlamydiae	Chlamydiae	Chlamydiales	other	other	0.0008	0.0004
10^-1	Chlamydiae	Chlamydiae	Chlamydiales	Parachlamydiaceae	other	0.0023	0.0005
10^-1	Chloroflexi	other	other	other	other	0.0003	0.0002
10^-1	Chloroflexi	Anaerolineae	Anaerolineales	Anaerolineaceae	other	0.0007	0.0002
10^-1	Chloroflexi	Caldilineae	Caldilineales	Caldilineaceae	Caldilinea	0.0004	0.0002
10^-1	Cyanobacteria	Cyanobacteria	Chloroplast	other	other	0.0007	0.0002

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-1	Firmicutes	Bacilli	Bacillales	other	other	0.0122	0.0017
10^-1	Firmicutes	Bacilli	Bacillales	Bacillaceae	other	0.0124	0.0012
10^-1	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0.0007	0.0002
10^-1	Firmicutes	Bacilli	Bacillales	Bacillaceae	Oceanobacillus	0.0002	0.0001
10^-1	Firmicutes	Bacilli	Bacillales	Bacillaceae	Tumebacillus	0.0001	0.0001
10^-1	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	Staphylococcus	0.0001	0.0001
10^-1	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	other	0.0001	0.0001
10^-1	Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales	Gemmatimonadaceae	Gemmatimonas	0.0251	0.0020
10^-1	Nitrospira	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	0.0016	0.0004
10^-1	OD1	OD1_genera_incertae_sedis	other	other	other	0.0016	0.0004
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	other	0.0564	0.0046
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Gemmata	0.0089	0.0011
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Pirellula	0.0101	0.0012
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Planctomyces	0.0080	0.0008
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Singulisphaera	0.0068	0.0010
10^-1	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Zavarzinella	0.0012	0.0004
10^-1	Proteobacteria	other	other	other	other	0.0172	0.0019
10^-1	Proteobacteria	Alphaproteobacteria	other	other	other	0.0707	0.0038
10^-1	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	other	0.0039	0.0008
10^-1	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Phenylobacterium	0.0048	0.0007
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	other	other	0.0288	0.0016
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	other	0.0002	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Devosia	0.0148	0.0012
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Hyphomicrobium	0.0024	0.0004
10^-1	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	Microvirga	0.0008	0.0003
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	other	other	0.0058	0.0007
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	Stella	0.0002	0.0002
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	other	0.0004	0.0002
10^-1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Inquilinus	0.0001	0.0001
10^-1	Proteobacteria	Alphaproteobacteria	Rickettsiales	Rickettsiaceae	Rickettsia	0.0006	0.0002
10^-1	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Erythrobacteraceae	Porphyrobacter	0.0003	0.0001
10^-1	Proteobacteria	Betaproteobacteria	other	other	other	0.0057	0.0007
10^-1	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia	0.0044	0.0007
10^-1	Proteobacteria	Deltaproteobacteria	other	other	other	0.0019	0.0006

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-1	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bacteriovoracaceae	Peredibacter	0.0009	0.0002
10^-1	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bdellovibrionaceae	Bdellovibrio	0.0000	0.0000
10^-1	Proteobacteria	Deltaproteobacteria	Myxococcales	other	other	0.0056	0.0009
10^-1	Proteobacteria	Deltaproteobacteria	Myxococcales	Cystobacteraceae	other	0.0002	0.0002
10^-1	Proteobacteria	Deltaproteobacteria	Myxococcales	Polyangiaceae	other	0.0018	0.0003
10^-1	Proteobacteria	Gammaproteobacteria	other	other	other	0.0131	0.0012
10^-1	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	other	0.0003	0.0002
10^-1	Proteobacteria	Gammaproteobacteria	Legionellales	Legionellaceae	other	0.0007	0.0003
10^-1	Proteobacteria	Gammaproteobacteria	Legionellales	Legionellaceae	Legionella	0.0007	0.0003
10^-1	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Alkanindiges	0.0002	0.0002
10^-1	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Sinobacteraceae	other	0.0003	0.0001
10^-1	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	other	0.0522	0.0037
10^-1	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Pseudoxanthomonas	0.0034	0.0007
10^-1	TM7	TM7_genera_incertae_sedis	other	other	other	0.0001	0.0001
10^-1	Verrucomicrobia	Opitutae	Opitutales	Opitutaceae	Opitutus	0.0293	0.0032
10^-1	Verrucomicrobia	Spartobacteria	other	other	other	0.0011	0.0004
10^-1	Verrucomicrobia	Spartobacteria	Spartobacteria_genera_incertae_sedis	other	other	0.0078	0.0017
10^-1	Verrucomicrobia	Subdivision3	Subdivision3_genera_incertae_sedis	other	other	0.0118	0.0009
10^-6	other	other	other	other	other	0.0912	0.0074
10^-6	Acidobacteria	Acidobacteria_Gp1	Gp1	other	other	0.0014	0.0006
10^-6	Acidobacteria	Acidobacteria_Gp3	Gp3	other	other	0.0036	0.0011
10^-6	Actinobacteria	Actinobacteria	Actinomycetales	other	other	0.0010	0.0003
10^-6	Bacteroidetes	other	other	other	other	0.0454	0.0118
10^-6	Bacteroidetes	Flavobacteria	Flavobacteriales	Cryomorphaceae	Fluviicola	0.0031	0.0015
10^-6	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Flavobacterium	0.0151	0.0040
10^-6	Bacteroidetes	Sphingobacteria	Sphingobacteriales	other	other	0.0172	0.0052
10^-6	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	other	0.1008	0.0083
10^-6	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Sphingobacteriaceae	other	0.0008	0.0003
10^-6	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Sphingobacteriaceae	Pedobacter	0.0204	0.0077
10^-6	Cyanobacteria	Cyanobacteria	Chloroplast	Streptophyta	other	0.0001	0.0001
10^-6	Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales	Gemmatimonadaceae	Gemmatimonas	0.0300	0.0046
10^-6	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	other	0.1121	0.0138
10^-6	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Zavarzinella	0.0035	0.0017
10^-6	Proteobacteria	Alphaproteobacteria	Rhizobiales	other	other	0.0091	0.0015

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-6	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	Labrys	0.0003	0.0001
10^-6	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	other	0.0002	0.0002
10^-6	Proteobacteria	Alphaproteobacteria	Rhodospirillales	other	other	0.0030	0.0007
10^-6	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	0.0227	0.0034
10^-6	Proteobacteria	Betaproteobacteria	other	other	other	0.0013	0.0004
10^-6	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Herbaspirillum	0.0001	0.0001
10^-6	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Janthinobacterium	0.0003	0.0003
10^-6	Proteobacteria	Deltaproteobacteria	other	other	other	0.0005	0.0002
10^-6	Proteobacteria	Deltaproteobacteria	Myxococcales	other	other	0.0034	0.0014
10^-6	Proteobacteria	Gammaproteobacteria	other	other	other	0.0025	0.0006
10^-6	Proteobacteria	Gammaproteobacteria	Legionellales	Coxiellaceae	Aquicella	0.0002	0.0001
10^-6	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	0.0001	0.0001
10^-6	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	other	0.0004	0.0003
10^-6	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	0.0081	0.0046
10^-6	Verrucomicrobia	other	other	other	other	0.0300	0.0079
10^-6	Verrucomicrobia	Opitutae	Opitutales	Opitutaceae	other	0.0318	0.0097
10^-6	Verrucomicrobia	Opitutae	Opitutales	Opitutaceae	Opitutus	0.0468	0.0109
10^-6	Verrucomicrobia	Spartobacteria	Spartobacteria_genera_incertae_sedis	other	other	0.0073	0.0041
10^-6	Verrucomicrobia	Subdivision3	Subdivision3_genera_incertae_sedis	other	other	0.0060	0.0015
10^-6	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	Verrucomicrobium	0.0002	0.0002
10^-9	other	other	other	other	other	0.0517	0.0175
10^-9	Acidobacteria	Acidobacteria_Gp1	Gp1	other	other	0.0017	0.0005
10^-9	Actinobacteria	Actinobacteria	other	other	other	0.0011	0.0004
10^-9	Actinobacteria	Actinobacteria	Acidimicrobiales	other	other	0.0004	0.0002
10^-9	Actinobacteria	Actinobacteria	Actinomycetales	other	other	0.0024	0.0004
10^-9	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides	0.0006	0.0004
10^-9	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae	Pseudonocardia	0.0006	0.0004
10^-9	Actinobacteria	Actinobacteria	Actinomycetales	Streptosporangiaceae	other	0.0002	0.0002
10^-9	Actinobacteria	Actinobacteria	Solirubrobacterales	Patulibacteraceae	Patulibacter	0.0001	0.0001
10^-9	Bacteroidetes	other	other	other	other	0.0444	0.0092
10^-9	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Chryseobacterium	0.0001	0.0001
10^-9	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Cloacibacterium	0.0001	0.0001
10^-9	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Flavobacterium	0.0167	0.0080
10^-9	Bacteroidetes	Sphingobacteria	Sphingobacteriales	other	other	0.0082	0.0032

Dilution	Phylum	Class	Order	Family	Genus	Abundance	SE
10^-9	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Chitinophagaceae	other	0.0742	0.0169
10^-9	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Cytophagaceae	Dyadobacter	0.0098	0.0065
10^-9	Bacteroidetes	Sphingobacteria	Sphingobacteriales	Sphingobacteriaceae	Pedobacter	0.0408	0.0072
10^-9	Chlamydiae	Chlamydiae	Chlamydiales	Parachlamydiaceae	other	0.0002	0.0001
10^-9	Deinococcus-Thermus	Deinococci	Deinococcales	Deinococcaceae	Deinococcus	0.0220	0.0135
10^-9	Firmicutes	Bacilli	Bacillales	other	other	0.0021	0.0005
10^-9	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Cohnella	0.0065	0.0034
10^-9	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus	0.0012	0.0004
10^-9	Firmicutes	Clostridia	Clostridiales	other	other	0.0001	0.0001
10^-9	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	other	0.0130	0.0064
10^-9	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Gemmata	0.0007	0.0003
10^-9	Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Planctomyces	0.0014	0.0006
10^-9	Proteobacteria	Alphaproteobacteria	Rhizobiales	other	other	0.0046	0.0007
10^-9	Proteobacteria	Alphaproteobacteria	Rhizobiales	Brucellaceae	Pseudochrobactrum	0.0001	0.0001
10^-9	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	Methylobacterium	0.0001	0.0001
10^-9	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Inquilinus	0.0010	0.0010
10^-9	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	other	0.0021	0.0006
10^-9	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	0.0003	0.0001
10^-9	Proteobacteria	Betaproteobacteria	other	other	other	0.0005	0.0002
10^-9	Proteobacteria	Betaproteobacteria	Burkholderiales	Alcaligenaceae	other	0.0012	0.0011
10^-9	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	other	0.0119	0.0090
10^-9	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bdellovibrionaceae	Bdellovibrio	0.0042	0.0020
10^-9	Proteobacteria	Deltaproteobacteria	Myxococcales	other	other	0.0010	0.0003
10^-9	TM7	TM7_genera_incertae_sedis	other	other	other	0.0001	0.0001
10^-9	Verrucomicrobia	other	other	other	other	0.1487	0.0355
10^-9	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	other	other	0.0012	0.0011
10^-9	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	other	0.0212	0.0075

2.4. Discussion

A number of studies have used the dilution method approach to artificially change microbial diversity (Salonius 1981, Garland and Lehman 1999, Franklin et al 2001, Griffiths et al 2001, Matos et al 2005, Franklin and Mills 2006, Wertz et al 2006, Hol et al 2010, Philippot et al 2013, Vivant et al 2013). This approach is one of the few available methods to manipulate microbial biodiversity of complex natural ecosystems such as the soil. And, indeed, our results show that dilution reduces the microbial biodiversity in the soil suspension and the soil after incubation of more or less diluted suspensions. Previous studies mostly based their conclusions on community measurements with limited resolution, detecting only the more abundant species since those species can be detected in the easiest way. However, compared to the rare biosphere, the abundant members are only a small fraction of microbial diversity (Sogin et al 2006), and thus, in this way, the real microbial biodiversity in these ecosystems may not be accounted for. Furthermore, none of those studies focused on changes in the community structure from the original more or less diluted inocula into different communities after incubation in soil or on the degree of variation in the suspensions after dilution and the consequences of this variation for the variances in the incubated soils. We suspected that the variance among the replicate samples would be considerable, and therefore we determined this variation in the suspension samples and how the variant communities developed during incubation and regrowth in soil. We were especially interested in the possibilities created by the dilution approach to separate abundant and rare species and thus allow experimental studies on the importance of rare (and abundant) microbes in soil ecosystems.

Although less abundant microbes should be more prone to be lost from the original microbial community at increasing dilution, our results show that unique OTUs still show up in the highest-dilution treatment in the suspensions (Fig. 2.7). Most likely certain microbial species are suppressed or masked for amplicon measurements in the low-dilution samples and only show up in the higher-dilution treatments. An issue that may have played an important role in the preparation of the diluted soil suspensions is the adsorption of cells on soil particles. Bakken (1985) claimed that a satisfactory separation of microorganisms and soil particles is not possible, and thus this could have influenced the structure of the microbial communities in the suspensions and, in

particular, the large variation therein. Moreover, also methodological errors may also have played a role in the failure of the sequencing approach to detect all species in the suspensions. For instance, the nested PCR could be a possible source of bias; therefore, the patterns from nested PCRs between samples were compared with the ones from direct PCRs. The patterns obtained from the nested PCRs were similar to those from the direct PCRs in soil suspension. Only minor variations were observed. In this experiment, the PCR products were purified before sequencing to exclude the nonincorporated primers. Thus, we concluded that the nested PCR approach may not have influenced significantly the results.

Similarly, our results indicate that, most likely, rare species that were suppressed in the low-dilution samples may have acquired an opportunity to develop in the higher-dilution samples because the cellular densities were low in those samples after dilution.

The data shown in Table 2.3 and the diagrams of Figure 2.5 clearly indicate that the present methodology, i.e. 454 pyrosequencing, does not allow for a complete view on the species present even in a dilute suspension as these data show that the diversity of the communities of the diluted samples increased during incubation in soil. We do not know what the precise detection limit is of the 454 pyrosequencing technique for observing microbial species in a suspension, but it is fair to assume that the bacterial species that are detected in soil after incubation were present but not detected in the soil suspensions, most probably because of their low abundance. As mentioned, also the data of the Venn diagram (Fig. 2.7) also clearly indicate the presence of species in all suspensions, including the 10^{-9} dilutions, which were not detected by our methodological approach.

The fact that these organisms were detected in soil but not in suspensions may be because these organisms were better adapted to the prevailing conditions of the soil environment (Brazelton et al 2010) than other organisms that were detected in the suspensions but not in soils. These other taxa may have been lost during incubation since they might have had special requirements not available in soil. It is not possible to conclude that these hidden species are rare species, and, thus, the conclusion is warranted that the dilution approach does not guarantee the identification of rare or less abundant versus abundant species.

Although all inoculated organisms returned into the same environment where they came from originally, the actual conditions for the individual organisms could have changed dramatically due to the difference in spatial arrangements and the large heterogeneity in soil. The factors that are responsible for the selection of microbes in soil resulting in the different communities as found in soil versus the communities in the suspensions are not clear. Previous studies have indicated that soil microbial communities were largely influenced by soil moisture (Schimel et al 1999, Brockett et al 2012). In our study, moisture availability after incubation could be a potential clue for the structuring of the community by selecting for individual microbial species with a relatively high moisture stress resistance. Other factors are said to be key to the shaping of bacterial communities in soil (Fierer et al 2003, Eichorst et al 2007, Kuramae et al 2010, Navarrete et al 2013), but the relevance of these factors for the assemblage of the communities from various inocula, as in this study, is not known.

In this study, we have considered several taxonomic diversity indicators. All indicated that the dilution procedure has a strong reducing effect on the microbial diversity (Table 2.1). We have used these different diversity indices because they give different insights into the diversity of complex communities such as soil microbial communities. In contrast to the richness index (Chao estimator), the diversity indices (Shannon and Simpson) focus on both the richness and evenness of a community. Shannon diversity is often sensitive to the presence of rare species, while the Simpson index emphasizes the dominant members (Nagendra 2002). Haegeman et al. (2013) suggests that community diversity is best estimated by Shannon and Simpson indices, whereas Chao estimator was not a reliable estimator of richness in the presence of rare species. Despite the differences in the focus of the diversity indices used here, all indices showed a similar trend. This strongly suggests that the alpha diversity decreases in response to dilution of microbial communities and that this decrease is reflected in the diversity of the communities after incubation in soil.

Interestingly, when we compared the diversity of the different phyla in suspension and in soil after incubation, we observed that the Shannon diversity indices of most phyla decreased from suspension to soil sample for the undiluted (10^{-1}) samples but increased for the most diluted (10^{-9}) samples (Table 2.3). Obviously, there are strong selection mechanisms operating in soil

that lead to a certain homogenization in the communities that are formed after regrowth of the suspensions. That observation is confirmed by the data of Figure 2.6 and Table 2.2, both of which show that the variances in the communities formed in the replicate samples diminished. We are not aware of similar observations presented in literature, but the findings are in line with the wealth of information that indicates that soil is a strong factor shaping the structure of the microbial community inhabiting the soil.

Analysis of the overall microbial community revealed that the community changed through dilution treatment of the soil suspensions and incubated soil at both phylum (Fig. 2.2) and OTU levels (Fig. 2.6). A detailed look at the microbial communities in the original non-diluted (10^{-1}) soil suspension revealed that the core groups comprised the well-known soil microbial phyla of *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, *Acidobacteria*, *Verrucomicrobia*, *Planctomycetes* and *Firmicutes* (Janssen 2006, Roesch et al 2007). During incubation, the same core groups were observed again, but the relative abundances of each group changed substantially. The largest changes in the occurrences of specific phyla were detected for the phylum of *Proteobacteria*, which was highly dominant especially in the higher dilutions but less dominant in the incubated soil samples, for the phylum of *Bacteroidetes*, which decreased slightly with increasing dilution in the soil suspensions but outgrew and increased significantly in the samples after incubation, and for the phylum of *Verrucomicrobia*, which was not detected in the higher soil suspension dilutions but showed up in high numbers after incubation. At the family level, we detected high proportions of *Beta-proteobacteria* represented by *Alcaligenaceae*, *Burkholderiaceae* and *Comamonadaceae* in the highest suspension dilution. Remarkably, their relative abundance decreased during incubation in soil. That was unexpected as *Proteobacteria* are dominant members in various soils, and as they are mostly fast-growing r-strategists, we expected them to be abundantly present in the incubated soil samples. The result may have been caused by the oligotrophic conditions prevailing in our test soil. However, the same observations after incubation in soil were made for *Acidobacteria*, which are generally considered to be soil-adapted oligotrophic organisms (Eichorst et al 2007), and for other well-known soil inhabitants such as *Actinobacteria*. It is interesting to see that groups such as *Verrucomicrobia*, and *Sphingobacteriaceae* and *Chitinophagaceae* families of *Bacterioidetes*

grew out significantly in all dilution treatments during incubation in soil. This contradicts what is known about *Verrucomicrobia*, which is usually considered a low-abundant phylum in soil (Janssen 2006). *Verrucomicrobia* may highly depend on C availability due to their slow-growing life strategy (Bergmann et al 2011); and that, in combination with the observed results, may indicate that *Verrucomicrobia* is a potential indicator of the response of these taxa to environmental factors (Fierer et al 2007). In summary, our results indicate that the dilution procedure leads to reduction of bacterial diversity, but the assembly of the microbial community during incubation in soil cannot be predicted on the basis of the composition of the inoculum. Obviously, soil has a strong selective power in shaping the microbial community, which leads to more uniform structures of the communities even after inoculation of much more variable suspensions. Also the deep sequencing approach applied here did not allow for a complete view of the microbial species present in even highly diluted suspensions. This also hinders the assessment and identification of rare species in a soil sample as even undetected species in the suspensions could develop into abundant populations after only eight weeks of incubation. In future studies we hope to be able to know more about the functional responses of more or less diverse samples and the consequences of these changes for the functioning of the soil ecosystem.

2.5. References

Bakken LR (1985). Separation and Purification of Bacteria from Soil. *Applied and Environmental Microbiology* **49:** 1482-1487.

Bergmann GT, Bates ST, Eilers KG, Lauber CL, Caporaso JG, Walters WA *et al* (2011). The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biology & Biochemistry* **43:** 1450-1455.

Brazelton WJ, Ludwig KA, Sogin ML, Andreishcheva EN, Kelley DS, Shen CC *et al* (2010). Archaea and bacteria with surprising microdiversity show shifts in dominance over 1,000-year time scales in hydrothermal chimneys. *Proceedings of the National Academy of Sciences of the United States of America* **107:** 1612-1617.

Brockett BFT, Prescott CE, Grayston SJ (2012). Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. *Soil Biology & Biochemistry* **44:** 9-20.

Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW, Almond REA *et al* (2010). Global Biodiversity: Indicators of Recent Declines. *Science* **328:** 1164-1168.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7:** 335-336.

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al* (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* **37:** D141-D145.

Dixon P (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14:** 927-930.

Edgar RC (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26:** 2460-2461.

Eichorst SA, Breznak JA, Schmidt TM (2007). Isolation and characterization of soil bacteria that define Teniglobus gen. nov., in the phylum Acidobacteria. *Applied and Environmental Microbiology* **73:** 2708-2717.

Fierer N, Allen AS, Schimel JP, Holden PA (2003). Controls on microbial CO2 production: a comparison of surface and subsurface soil horizons. *Global Change Biology* **9:** 1322-1332.

Fierer N, Bradford MA, Jackson RB (2007). Toward an ecological classification of soil bacteria. Ecology **88:** 1354-1364.

Franklin RB, Garland JL, Bolster CH, Mills AL (2001). Impact of dilution on microbial community structure and functional potential: Comparison of numerical simulations and batch culture experiments. *Applied and Environmental Microbiology* **67:** 702-712.

Franklin RB, Mills AL (2006). Structural and functional responses of a sewage microbial community to dilution-induced reductions in diversity. *Microbial Ecology* **52:** 280-288.

Garland JL, Lehman RM (1999). Dilution/extinction of community phenotypic characters to estimate relative structural diversity in mixed communities. *Fems Microbiology Ecology* **30:** 333-343.

Good IJ (1953). The Population Frequencies of Species and the Estimation of Population Parameters. *Biometrika* **40:** 237-264.

Griffiths BS, Ritz K, Wheatley R, Kuan HL, Boag B, Christensen S *et al* (2001). An examination of the biodiversity-ecosystem function relationship in arable soil microbial communities. *Soil Biology & Biochemistry* **33:** 1713-1722.

Haegeman B, Hamelin J, Moriarty J, Neal P, Dushoff J, Weitz JS (2013). Robust estimation of microbial diversity in theory and in practice. *Isme Journal* **7:** 1092-1101.

Hammer et al HD, Ryan PD. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4: 9. pp.**

Hol WHG, de Boer W, Termorshuizen AJ, Meyer KM, Schneider JHM, van Dam NM *et al* (2010). Reduction of rare soil microbes modifies plant-herbivore interactions. *Ecology Letters* **13:** 292-301.

Huber JA, Mark Welch D, Morrison HG, Huse SM, Neal PR, Butterfield DA *et al* (2007). Microbial population structures in the deep marine biosphere. *Science* **318:** 97-100.

Hughes JB, Hellmann JJ (2005). The application of rarefaction techniques to molecular inventories of microbial diversity. *Environmental Microbiology* **397:** 292-308.

Janssen PH (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Applied and Environmental Microbiology* **72:** 1719-1728.

Kuramae EE, Gamper HA, Yergeau E, Piceno YM, Brodie EL, DeSantis TZ *et al* (2010). Microbial secondary succession in a chronosequence of chalk grasslands. *Isme Journal* **4:** 711-715.

Lane DJ (1991). 16S/23S rRNA sequencing, p. 115-175. *In* EStackebrandt and MGoodfellow (ed), Nucleic acid techniques in bacterial systematics John Wiley & Sons, Chichester, United Kingdom.

Magurran AE, Henderson PA (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature* **422:** 714-716.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA *et al* (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437:** 376-380.

Matos A, Kerkhof L, Garland JL (2005). Effects of microbial community diversity on the survival of Pseudomonas aeruginosa in the wheat rhizosphere. *Microbial Ecology* **49:** 257-264.

Muyzer G, Dewaal EC, Uitterlinden AG (1993). Profiling of Complex Microbial-Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes-Coding for 16s Ribosomal-Rna. *Applied and Environmental Microbiology* **59:** 695-700.

Nagendra H (2002). Opposite trends in response for the Shannon and Simpson indices of landscape diversity. *Applied Geography* **22:** 175-186.

Navarrete AA, Kuramae EE, de Hollander M, Pijl AS, van Veen JA, Tsai SM (2013). Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. *Fems Microbiology Ecology* **83:** 607-621.

Neufeld JD, Li J, Mohn WW (2008). Scratching the surface of the rare biosphere with ribosomal sequence tag primers. *FEMS Microbiology Letters* **283:** 146-153.

Philippot L, Spor A, Henault C, Bru D, Bizouard F, Jones CM *et al* (2013). Loss in microbial diversity affects nitrogen cycling in soil. *Isme Journal* **7:** 1609-1619.

Reeder J, Knight R (2010). Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nature Methods* **7:** 668-669.

Roesch LF, Fulthorpe RR, Riva A, Casella G, Hadwin AKM, Kent AD *et al* (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *Isme Journal* **1:** 283-290.

Sala OE, Chapin FS, Armesto JJ, Berlow E, Bloomfield J, Dirzo R *et al* (2000). Biodiversity - Global biodiversity scenarios for the year 2100. *Science* **287:** 1770-1774.

Salonius PO (1981). Metabolic Capabilities of Forest Soil Microbial-Populations with Reduced Species-Diversity. *Soil Biology & Biochemistry* **13:** 1-10.

Schimel JP, Gulledge JM, Clein-Curley JS, Lindstrom JE, Braddock JF (1999). Moisture effects on microbial activity and community structure in decomposing birch litter in the Alaskan taiga. *Soil Biology & Biochemistry* **31:** 831-838.

Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR *et al* (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America* **103:** 12115-12120.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE *et al* (2008). A core gut microbiome in obese and lean twins. *Nature* **457:** 480-484.

Vivant AL, Garmyn D, Maron PA, Nowak V, Piveteau P (2013). Microbial Diversity and Structure Are Drivers of the Biological Barrier Effect against Listeria monocytogenes in Soil. *Plos One* **8**.

Wertz S, Degrange V, Prosser JI, Poly F, Commeaux C, Freitag T et al (2006). Maintenance of soil functioning following erosion of microbial diversity. *Environ Microbiol* 8: 2162-2169.

Chapter 3

Soil characteristics determine the assemblage of bacterial communities in terrestrial ecosystems

Yan Yan, Peter G.L. Klinkhamer, Johannes A. van Veen, Eiko E. Kuramae

Manuscript

Abstract

In a previous study we showed that the soil had a strong effect on the assemblage of bacterial communities after inoculation of a sterilized soil with bacterial suspensions of that same soil in a dilution series. Here, we continued our investigation on the impact of soil on the assemblage of bacterial communities and checked the concept of the overriding effect of the soil on shaping bacterial communities after inoculation of suspensions obtained from different soils. Diluted suspensions from different soils harboring different bacterial diversities were crossed inoculated into three pre-sterilized soils. We used 16S rRNA gene amplicon sequencing to determine the bacterial community structure of the suspensions and the soils. In a multivariate analysis the different regrown soil bacterial communities after inoculation of different diluted suspensions in a particular soil clustered together while the same suspensions inoculated in different soils were separated. Diversity indices of the suspensions were reduced significantly upon dilution. The strength of selection of the soil on the bacterial communities was stronger for the undiluted 10^{-1} soil samples than of the diluted 10^{-9} soil samples. Permanova tests showed that dilution had a slightly larger effect on the community structure than soil had, and both main effects were larger than their interactions. Meijendel soil was characterized by the highest organic matter, ammonium and nitrate concentrations and pH, while Clue soil was characterized by the highest phosphorus concentration, and Utrecht soil was characterized by the highest C:N ratio. These differences in abiotic environmental factors may explain the variation in bacterial communities across these soils.

Keywords

Microbial biodiversity | Soil selection | Soil chemical factors

3.1. Introduction

Soil-borne bacteria represent an essential component of terrestrial ecosystems, which are key to many vital ecosystem functions (Philippot et al 2013). With an estimate of hundreds to thousands of taxa per gram of soil, their diversity provides the majority of biodiversity in terrestrial ecosystems (Torsvik and Ovreas 2002). Abiotic factors such as pH (Fierer and Jackson 2006), moisture (Brockett et al 2012) and salinity (Crump et al 2004), phosphate availability (Faoro et al 2010), and organic matter content (Verbruggen et al 2010) as well as biotic interactions with local communities of macro- and micro-organisms are known to drive the activity of soil-borne bacteria and to shape their community structure (Garbeva et al 2004, Berg and Smalla 2009).

In order to be able to assess and predict the dynamics of microbial communities in soil, proper understanding of the mechanisms of the assemblage of microbial communities in soil and other natural environments is a long-standing goal of microbial ecology. Microorganisms are dispersed globally and able to propagate in any habitat with suitable environmental conditions (Martiny et al 2006). Upon their arrival in a new environment or upon drastic changes of their current environment, microorganisms may either be assembled into distinct, new, community profiles (Panke-Buisse et al 2015) or functionally adapt to the local habitat without dramatic changes in community composition (Comte and del Giorgio 2010). Thus, bacteria colonizing the soil may be included into a specific microbiome of distinct structure and functionality.

Many studies have addressed the factors responsible for the structuring of microbial communities in soils (Pavon-Jordan et al 2013), but often such studies typically focus on the importance of a single factor without considering the full complexity of all edaphic properties for which an increasing amount of evidence is being generated as critical drivers for the shaping of microbial communities (Kuramae et al 2012, Navarrete et al 2013). Moreover, only few studies have examined the development of communities after addition of diverse inocula in soils, and, when they were done, such studies are often restricted in analysis depth providing little detailed information on taxonomy (Garland and Lehman 1999, Franklin et al 2001, Griffiths et al 2001, Matos et al 2005, Franklin and Mills 2006). Therefore, it is difficult to predict with confidence how soil microbial communities are assembled in different soil habitats.

Advanced sequencing approaches now allow for a more accurate and detailed assessment of the assemblage and the structuring of microbial communities in soils. The major aim of this study was to assess the shaping of bacterial communities after inoculating different suspensions varying in bacterial diversity into different soils, so to detect the impact of soil on the assemblage of microbial communities. We sequenced the 16S rRNA gene marker to provide information about initial inocula and soil bacterial communities after regrowth in soils. A previous study had already shown that the dilution approach is suitable for manipulating the diversity of bacterial communities, and that soil had a strong selective power in shaping the microbial community after inoculation of the different suspensions leading to a rather uniform structure of the regrown microbial community (Chapter 2). In this study, we took this approach one step further by inoculating into three soils two dilutions of the suspensions of these three different soils in a cross-inoculation design. We addressed two basic questions: 1) Does soil determine the composition of the bacterial community after incubation following the inoculation of suspensions with different bacterial diversity? 2) If so, which are the main soil characteristics explaining the structure of the final bacterial community?

3.2. Material and Methods

3.2.1. Soil sampling and treatment

Three field soils were selected across The Netherlands: soils from the surroundings of Utrecht (52°03′N, 5°13′E), from the so-called Clue fields (52°03′N, 5°45′E) and from the Meijendel dunes (52°9′N, 4°22′E). Ten liter of each soil were collected at a depth of around 15 cm in each field. The soil was sieved (5mm) and homogenized, and aliquots of 50 g were stored in plastic bags (Whirl-Pak sampling bag, 100 ml; Sigma-Aldrich). One bag of each soil was kept separately to serve as inoculum. The bags with soil were sterilized by gamma irradiation (> 35 kGy; Isotron, Ede, The Netherlands). As compared to autoclaving and freezing this way of sterilizing soils minimizes the effects on abiotic soil properties. Sterility was checked by spreading 0.5 g of sterilized soil onto Reasoner's 2A (R2A) and potato-dextrose agar (PDA) media and incubated for one week. No bacterial or fungal growth on agar plates for six

replicates was observed in the sterile soil samples after incubation. To control the sterility during the experiment, plates were incubated for the duration of the experiment ay 28 °C and no colonies were observed during the entire incubation period. Three sterilized bags of each soil were inoculated with autoclaved demineralized water to be used as controls. A subsample of the fresh soil was taken to determine soil moisture (24 h, 105°C).

Soil suspensions were made by mixing 20 g fresh soil and 190 ml autoclaved demineralized water with a blender for 2 minutes. This procedure was repeated 3 times and in between the blender was cooled down on ice for 2 minutes. The obtained suspension was called the 10^{-1} dilution. 100 ml of 10^{-1} dilution was transferred to a bottle containing 900 ml of autoclaved demineralized water and subsequently shaken by hand for 1 min. This procedure was repeated for several times until 10^{-9} dilutions were obtained. Subsequently, 2.5 ml of the respective dilutions were added to 25 g of soil in the bags and additional demineralized water was given to bring the moisture level of the inoculated soil at around 20%, which is roughly similar to the average level of the prevailing moisture conditions at the sides from where the soil was taken.

The experiment was designed as a cross inoculation experiment in which we inoculated suspensions of the three soils in each of the sterilized soils. The experiment consisted of twelve treatments in a factorial design, with 108 samples, i.e. three soils × three inocula × two dilutions × six replicates, including 9 sterile controls (three soils × three replicates). The remaining suspensions were centrifuged at 3000 g for 10 min at 4°C and the pellets were stored at -20°C for further analysis. After inoculation, soil bags were incubated at 20°C using sterilized cotton plug caps to ensure gas exchange with 70% humidity in the climate chamber. The soils were turned over regularly once a week to enable homogeneous microbial growth. After 9 weeks of incubation soil samples were taken, under laminar flow conditions, to determine the microbial abundance in all treatments by quantitative real time PCR (qPCR).

Total DNA was extracted from the incubated soil using the MoBio Power Soil Extraction Kit according to the supplier's manual. Total DNA concentration was quantified and qualified on ND-1000 spectrophotometer (Nanodrop Technology, Wilmington, DE). Amplification of the 16S rRNA

gene was performed using the primer set Eub 338 and Eub 518 (Muyzer et al 1993). Each 25 µl reaction solution consisted of 12.5 µl Sybergreen mix (Bioline, GC-Biotech) with 4 mg/ml BSA in a total volume of 25 µl, 5 µM of each primer, 5 µl template DNA (5 ng/µl). For bacteria, the standard curves were generated using 10-fold dilution series from 10^8 to 10^3 of plasmid DNA obtained from *Firmicutes*. Polymerase Chain Reactions were run on a Rotor-Gene 3000 (Qiagen) and started with 15 min at 95°C, followed by 40 amplification cycles each of 95°C for 60 sec, 53°C 50 sec and 72°C 60 sec. A subsample of the soil from each bag was stored at -20°C for further analysis. Triplicate reactions per DNA sample and the appropriate set of standards were used. For qPCR assays, a linear relationship was presumed between the log of the plasmid DNA copy number and the calculated threshold cycle (Ct value). PCR efficiencies were 99% and correlation coefficients for standard curves were $R^2 = 0.99$. Because there were differences of bacterial abundance between the undiluted 10^{-1} samples and diluted 10^{-9} samples after 9 weeks of incubation (Fig. 3.1), the relative abundance of species was used for further analysis and comparison among samples.

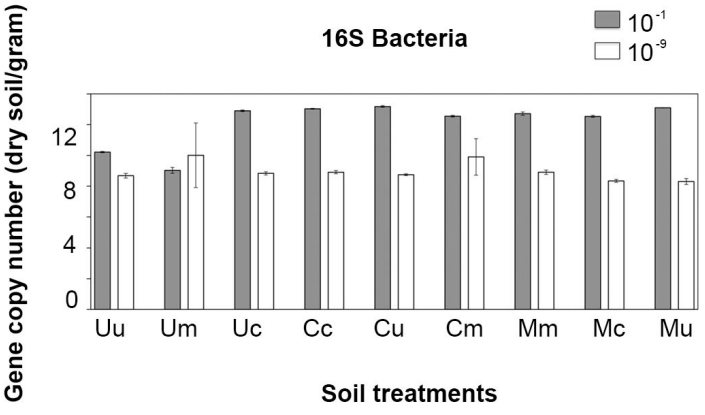

Figure 3.1. Real time PCR of bacterial abundance after 9 weeks incubation. Capital letter means suspension; lower case means incubated soil. Error bars mean standard errors (n=3).

3.2.2. 16S rRNA amplicon sequencing

Total DNA was extracted from the soil suspensions and incubated soil, as described above, to determine the composition of the respective microbial communities by 16S rRNA illumina Miseq. For DNA concentrations below 5 ng/µl, *i.e.* some of the 10^{-9} soil suspension samples, nested PCR was performed and for DNA concentrations above 5 ng/µl, direct PCR was performed. The general bacterial primer set 27F and 1492R (Lane 1991) was used for the first amplification, and subsequently 2 µl of the amplified products from the first round was used as template for the second round PCR using barcoded primers 515F and 806R (Caporaso et al 2012). The PCR program used included incubation at 95°C for 5 min followed by 25 cycles each of 95°C for 30 s, 55°C 1 min and 72°C 10 min. For PCR reactions 5 mM dNTPs (Invitrogen, Carlsbad, CA), 1 unit of *Taq* polymerase (Roche, Indianapolis, IN), and 5 ng/µl of sample DNA as the template in a total volume of 25 µl were used. The PCR conditions for the second round were similar to the first PCR round except for 25 cycles with 52°C annealing temperature. To control for contamination during PCR amplification, one negative control (water instead of DNA) was included for all PCR reactions. PCR products of each subsample from the barcoded primers were generated in six replicates per sample and purified using the Agencourt AMPure XP beads kit. Equimolar purified PCR products that were quantified by fragment analyzer (Advanced Analytical Technologies, GmbH, Germany) were mixed and sequenced using Miseq sequencing from Illumina (Argonne Institute, USA).

3.2.3 Sequence analysis

The RDP extention to PANDASeq (Masella et al., 2012) named Assembler (Cole et al., 2014) was used to merge paired-end reads with a minimum overlap of 10bp and at least a PHRED score of 25. Primer sequences were removed from the per sample FASTQ files using Flexbar version 2.5 (Dodt et al., 2012). Sequences were converted to FASTA format and concatenated into a single file. All reads were clustered into OTUs using the UPARSE strategy by dereplication, sorted by abundance with at least two sequences and clustered using the UCLUST smallmem algorithm (Edgar 2010). These steps were performed with VSEARCH version 1.0.10 (Rognes et al., 2015), which is an

open-source and 64-bit multithreaded compatible alternative to USEARCH. Subsequently, chimeric sequences were detected using the UCHIME algorithm (Edgar et al., 2011) implemented in VSEARCH. All reads before the dereplication step were mapped to OTUs using the usearch_global method implemented in VSEARCH to create an OTU table and converted to BIOM format 1.3.1 (McDonald et al., 2012). Finally, taxonomic information for each OTU was added to the BIOM file by using the RDP Classifier version 2.10 (Cole et al., 2014). All steps were implemented in a Snakemake workflow (Köster and Rahmann, 2012).

Alpha diversity was calculated based on the rarefied OTU table (Hughes and Hellmann 2005). Eight samples were deleted after rarefaction since those samples had substantially lower reads in comparison to the other ones. The average reads from 3 sterilized controls of each soil were used as baseline, which was subtracted from the reads of each soil samples, respectively. The OTU table after subtraction of the control was used for further statistical analysis. Shannon diversity indices were determined with the "vegan" package (Dixon 2003) in R (The R Foundation for Statistical Computing). The percentage of coverage was calculated by Good's method using the formula % = [1-(n/N)] × 100, where n means the number of phylotypes represented by singletons and N is the total number of sequences (Good 1953). Good's method equation gives an estimate of the coverage of an entire sampled community.

To assess the differences between communities of different treatments, Nonmetric Multi-Dimensional Scaling (NMDS) plots were used to visualize the structure among samples at OTU level. Canonical Correlation Analysis (CCA) was used to evaluate the linkages between soil microbial structure and soil chemical characteristics. The plots were generated from Bray-Curtis similarity index matrices of all samples. The effects of soil and dilution on bacterial community composition were tested by a two-way PERMANOVA test for each inoculum, respectively. All the multivariate analyses were performed using the PAST software (Hammer et al 2001).

3.3. Results

3.3.1. Chemical characteristics of three field soils

Soil chemical analysis showed that Meijendel soil had the highest pH and had higher concentrations of NO^{3-}, NH^{4+} and organic matter than the other soils, while Utrecht soil had highest C:N ratio and the lowest pH. Clue soil had the highest phosphorus concentration (Table 3.1).

Table 3.1. Chemical properties of each field soil

Chemical properties	Field soil		
	Utrecht	Clue	Meijendel
OM (%)	4.67±0.18 a	3.97±0.29 a	9.11±0.36 b
NO_3^- (mg/kg)	0.02±0.02 a	6.50±0.51 b	30.43±0.85 c
NH_4^+ (mg/kg)	0.92±0.20 a	1.21±0.18 ab	2.23±0.25 b
Phosphorus (mg/kg)	2.28±0.35 a	80.84±3.56 b	15.16±0.41 c
C:N ratio	20.30±1.22 a	14.81±0.69 ab	12.16±0.26 b
pH (H_2O)	4.61±0.023 a	5.77±0.015 b	7.47±0.005 c

Values are mean ±SE, n = 6. Within columns, means followed by the same letter are not significantly different ($p < 0.05$) based on Tukey HSD test. OM means organic matter. Data was transformed to fit normal distribution when needed.

3.3.2. Effect of dilution and soil on bacterial community diversity

Alpha diversity indices reduced significantly upon dilution for each of the three soil suspensions (Table 3.2). Diversity indices changed substantially after incubation in the different soils. This is especially true for Clue and Meijendel inocula. In general, the diversity of the soil with 10^{-9} diluted inocula increased compared to that of 10^{-9} suspension. Good's estimator of coverage of all samples was above 99% indicating that sequencing depth was enough to detect most species in this study.

Table 3.2. Estimators of sequence library diversity and coverage in soil suspensions and incubated soil samples.

Time	Suspension	Soil	Dilution	Observed	Shannon	Good's estimator of coverage
Suspension	Utrecht		10^{-1}	134±2	3.40±0.04	0.995
Suspension	Utrecht		10^{-9}	100±2	2.28±0.03	0.998
Suspension	Clue		10^{-1}	190±4	3.79±0.08	0.993
Suspension	Clue		10^{-9}	50±2	0.91±0.02	0.997
Suspension	Meijendel		10^{-1}	190±3	3.59±0.11	0.994
Suspension	Meijendel		10^{-9}	80±8	2.17±0.11	0.999
Soil	Utrecht	Utrecht soil	10^{-1}	111±6	3.27±0.07	0.994
Soil	Clue	Utrecht soil	10^{-1}	91±4	3.13±0.06	0.995
Soil	Meijendel	Utrecht soil	10^{-1}	57±4	2.46±0.22	0.998
Soil	Utrecht	Clue soil	10^{-1}	138±4	3.54±0.06	0.995
Soil	Clue	Clue soil	10^{-1}	163±3	3.79±0.06	0.995
Soil	Meijendel	Clue soil	10^{-1}	131±8	3.32±0.06	0.995
Soil	Utrecht	Meijendel soil	10^{-1}	132±3	3.58±0.11	0.997
Soil	Clue	Meijendel soil	10^{-1}	189±1	4.25±0.04	0.995
Soil	Meijendel	Meijendel soil	10^{-1}	164±2	3.86±0.06	0.995
Soil	Utrecht	Utrecht soil	10^{-9}	113±12	2.21±0.42	0.995
Soil	Clue	Utrecht soil	10^{-9}	113±13	2.59±0.31	0.995
Soil	Meijendel	Utrecht soil	10^{-9}	121±19	2.16±0.45	0.996
Soil	Utrecht	Clue soil	10^{-9}	142±19	2.53±0.41	0.994
Soil	Clue	Clue soil	10^{-9}	173±7	3.48±0.22	0.993
Soil	Meijendel	Clue soil	10^{-9}	128±17	2.51±0.22	0.994
Soil	Utrecht	Meijendel soil	10^{-9}	140±9	2.77±0.38	0.990
Soil	Clue	Meijendel soil	10^{-9}	122±6	1.85±0.19	0.988
Soil	Meijendel	Meijendel soil	10^{-9}	100±7	1.56±0.23	0.988

Estimators and statistical significance ($P < 0.05$) were calculated for each dilution treatment of incubated soil samples (n = 5-6) on the basis of the phylogenetic profile at the species level. S.obs is the observed number of OTUs. NS means not significant.

3.3.3. Effects of dilution and soil on bacterial community composition

To explain the variability of the community composition in the different treatments, relative abundances were used to compute the Bray-Curtis similarity matrix (Fig. 3.2). Samples were grouped according to soils and dilutions

(ANOSIM: R = 0.80, $P < 0.001$). We also assessed, by CCA (Fig. 3.2B), which of the soil characteristics could be responsible for the separation of the communities, including organic matter (OM), ammonium and nitrate concentrations and pH, phosphorus concentration, and C:N ratio.

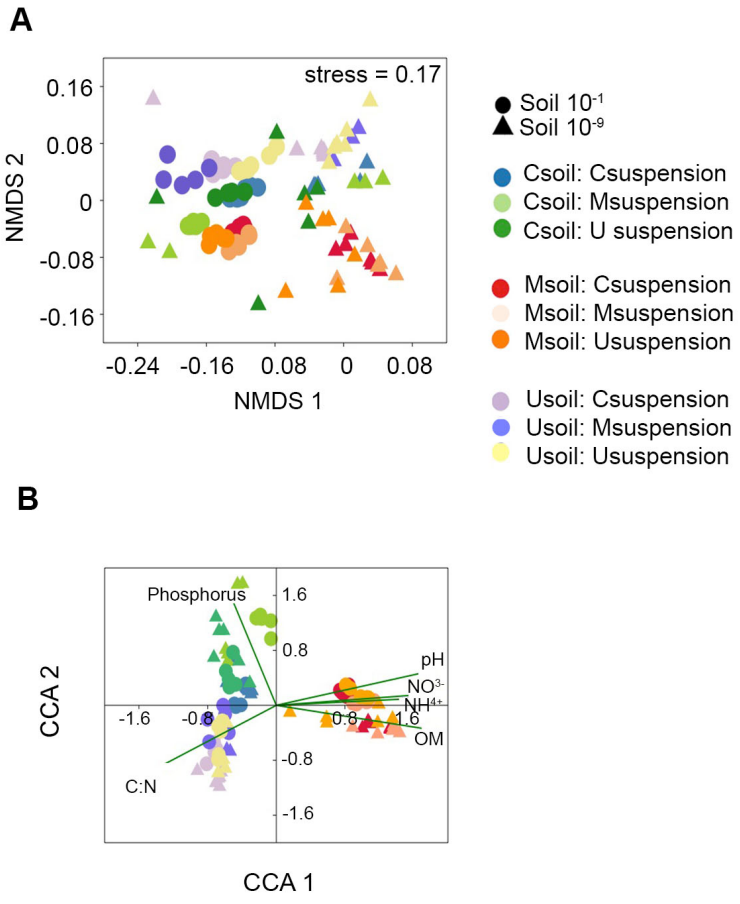

Figure 3.2. Redundancy analysis of bacterial community patterns and soil characteristics from samples. (A) NMDS plot of Bray-Curtis similarity matrix among six replicate samples of the two dilutions of the incubated soil samples. (B) CCA plot of bacterial community patterns and soil chemical characteristics for the incubated soil samples.

Furthermore, we quantified the individual effects of soil and dilution on the regrown bacterial communities with different inoculations. Based on the associated *F*-values, for all three soils, the dilution effect was slightly stronger than the soil effects and both individual effects were larger than their interactions. Permanova test for these two factors yielded statistically

significant results regarding dilutions and soils, as well as their interactions (Table 3.3).

Table 3.3. Results from two-way PERMANOVA analysis using Bray-Curtis similarity showing the effects of soil, dilution and their interaction on the composition of bacterial communities.

Inocula	Factors	Sum of sqrs	df	Mean Square	*F*	p
U suspension	Soil	2.88	2	1.44	10.94	0.0001
	Dilution	1.76	1	1.76	13.38	0.0001
	Interaction	1.41	2	0.70	5.34	0.0001
C suspension	Soil	3.03	2	1.52	22.45	0.0001
	Dilution	2.47	1	2.47	36.64	0.0001
	Interaction	2.15	2	1.08	15.94	0.0001
M suspension	Soil	3.22	2	1.61	17.73	0.0001
	Dilution	2.18	1	2.18	24	0.0001
	Interaction	2.09	2	1.04	11.47	0.0001

The most dominant phyla in the undiluted 10^{-1} suspension samples were *Proteobacteria*, *Acidobacteria*, *Actinobacteria*, *Planctomycetes*, *Bacteroidetes* and *Verrucomicrobia* (Fig. 3.3). Some interesting trends in the relative abundance of phyla were observed. The abundance of *Proteobacteria* dominated in the three soils of the undiluted 10^{-1} soil suspension samples, and *Actinobacteria*, *Deinococcus* and *Proteobacteria* comprised around 80% of the total population of the, diluted, 10^{-9} Utrecht and Meijendel soil suspension samples, respectively (Fig. 3.3A). The communities that developed after inoculation and incubation of a suspension derived from that same soil showed to be highly similar with regard to phyla abundance, whereas communities developed in other soils were less similar. The differences in phylum abundance after incubation were larger for the diluted 10^{-9} samples than for the 10^{-1} undiluted samples. This holds for all three soils.

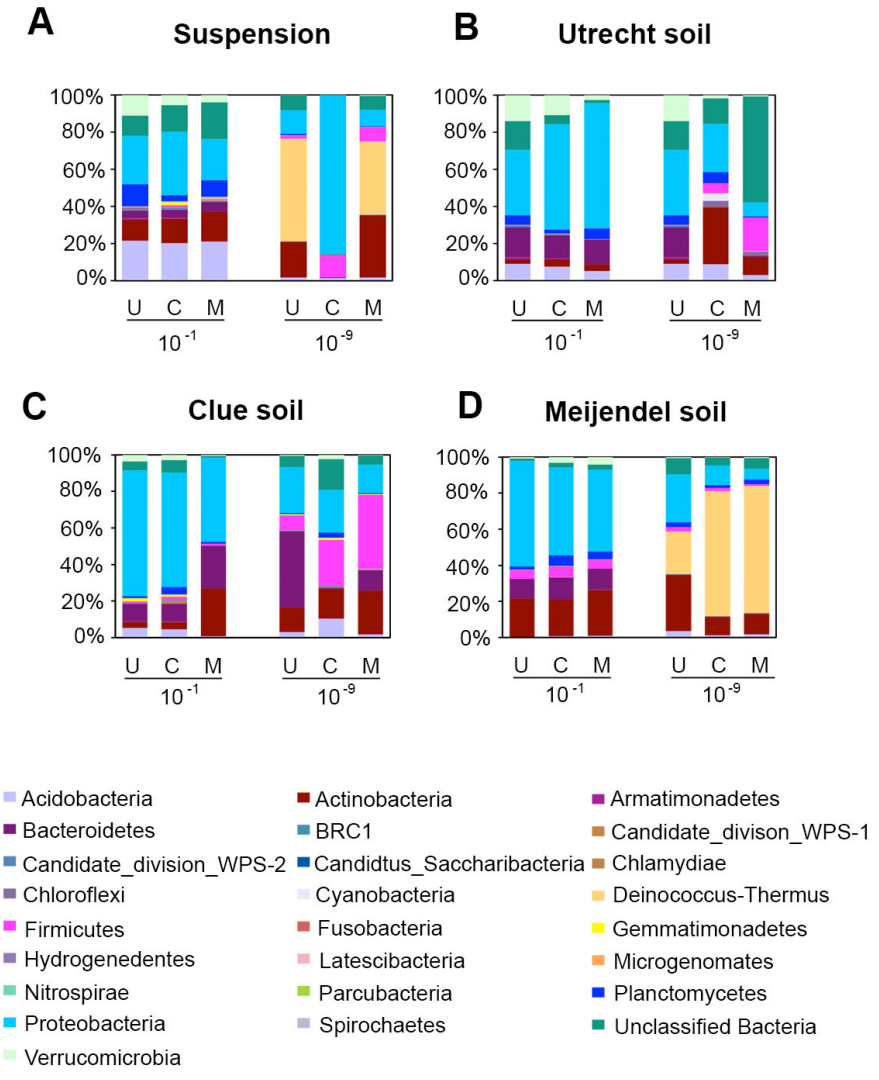

Figure 3.3. Bacterial community composition based on relative abundances at the phylum level of soil suspensions and incubated soil samples. (A) suspension samples, the origins are indicated at the bottom of each plot; (B-D) incubated soil samples from Utrecht soil, Clue soil and Meijendel soil.

3.4. Discussion

In a previous study (Chapter 2) we already found strong indications that soil is a major driving force shaping the structure of bacterial communities that develop after inoculation of suspensions with different bacterial diversities and compositions. The current study was designed to test the concept of the overriding impact of soil on the assemblage of bacterial communities leading to homogeneous community structures even after inoculation of suspensions from different soils and thus different bacterial communities. The composition of the regrown soil bacterial communities was strongly determined by the soil in which they were inoculated such that these communities showed great similarity within each of the three inoculated soils while across the three inoculated soils there were large differences. This result confirmed the earlier formulated concept of the strong impact of soil on the assemblage of bacterial communities. The observations of the soil effect were clear both at the phylum (Fig. 3.3) and the OTU (Fig. 3.2) levels.

Also we observed that dilution had a strong effect on bacterial diversities in the different soils (Table 3.2). Similarly, as was already demonstrated in Chapter 2, diluting soil suspensions to 10^{-9} dilutions led to a significant reduction of the diversity of the bacterial communities. Considering that we inoculated sterile soil, it is fair to assume that the soil itself did not add a substantial inoculum to the community. The observed increased diversity of some of the communities that developed after incubation of diluted 10^{-9} suspensions as compared to the original suspension, may reflect the failure of the technological approach to detect all organisms in a sample, although Good's coverage was large enough to assume with confidence that the largest proportion of the present community was included in the sequence process. Remarkably, the diluted 10^{-9} suspension of Clue soil had a much lower diversity than the other 10^{-9} dilution suspensions (Table 3.2), which was associated with a dominant proportion of *Proteobacteria* in these 10^{-9} dilution Clue suspension (Fig. 3.3). We do not have a straight explanation for this observation; the diversity index for the undiluted Clue suspension was not strongly aberrant from the other soil suspensions. Neither can we explain the observation that the diluted 10^{-9} soil samples that were regrown in Meijendel soil showed remarkably lower diversity indices for both Clue and Meijendel inocula, respectively (Table 3.2). Meijendel soil was characterized by the highest pH

and high proportion of organic matter, nitrate and phosphate. The communities of these soil samples were dominated by *Deinococcus* spp. Neither from our own data nor from literature data on the occurrence of *Deinococcus* spp we can directly relate the high abundance in Meijendel soils directly to these environmental factors.

Generally, a higher similarity in bacterial community composition was found amongst soils after incubation of the more diverse undiluted 10^{-1} samples than that of the less diverse diluted 10^{-9} samples (Fig. 3.2). *Proteobacteria* were dominant in each soil after incubation of undiluted samples, which may explain the relative similarity in the communities after incubation of the undiluted 10^{-1} soil suspensions as compared to the communities that developed after incubation of the diluted 10^{-9} soil suspensions, in which *Proteobacteria* nor any other phylum was consistently dominant (Fig. 3.3). The well known soil bacterial phyla of *Proteobacteria*, *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Deinococcus* and *Verrucomicrobia* (Roesch et al 2007) were observed as dominant members of the communities of each soil.

As the data of Table 3.3 indicate, the effect of dilution is, at least, equal or larger than the effect of soil on the community assemblage process, indicating that the dilution approach not only caused a dramatic reduction in the species diversity but also to rather different communities. The diversity indices of incubated samples were not always highest in their own native soil (Table 3.1). This indicates that only a selected fraction of the original community is able to establish in soil, even in the 'own' soil; more detailed information on the physico-chemical factors that are responsible for this differential species establishment in soil after inoculation of suspensions is needed to predict the outcome of the regrown process on the basis of the original composition of the suspensions.

Logically, on the basis of the afore-described consideration the environmental factors that may determine the assemblage process differ significantly among the soils. Among the soil characteristics that we found to be of significant importance in this study, pH is often regarded as a key factor for shaping bacterial communities (Rousk et al 2010). Remarkably, pH was only indicated as a determining factor in Meijendel soil, which could explain the bacterial communities in Meijendel soil. The abundance of *Acidobacteria* may be another indicator of acidic conditions in soil (Navarrete et al 2013).

Indeed, we observed that the relative abundance of Gp1 and Gp3 of *Acidobacteria* was highest in the most acid Utrecht soils and lowest in the Meijendel soil (Fig. 3.3). Also other soil factors could contribute significantly to the assemblage of the bacterial communities, such as organic matter, nitrite and ammonium, which were found probably to be important for shaping the bacterial communities in Meijendel soil. A previous study reported that the abundance of *Firmicutes* was highly correlated with phosphorus content (Kuramae et al 2012), and, indeed, we found that the abundance of *Firmicutes* was highest in the Clue field soil in which we observed the highest phosphorus content of the three soils. Thus, our study shows that not only pH as suggested by Fierer and Jackson (2006) and Rousk (2010), but also other environmental factors may serve as determinants of the structure of bacterial communities in specific soils.

In conclusion, we have shown that soil characteristics have strong impact on the assemblage of bacterial communities. Soil abiotic factors play a major role in shaping bacterial community structure independent of the diversity of the original suspension inoculated in soil. Indeed, the three soils that were considered in this study modified the bacterial community structure differently by providing specific habitats suitable for the growth of the inocula, which confirmed the concept of the overriding impact of the physicochemical nature of the soil on the assemblage of bacterial communities in terrestrial ecosystems.

3.5. References

Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM (2006). The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* **57:** 233-266.

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

Bhattacharyya PN, Jha DK (2012). Plant growth-promoting rhizobacteria (PGPR): emergence in agriculture. *World Journal of Microbiology & Biotechnology* **28:** 1327-1350.

Brockett BFT, Prescott CE, Grayston SJ (2012). Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. *Soil Biology & Biochemistry* **44:** 9-20.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N *et al* (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *Isme Journal* **6:** 1621-1624.

Cole JR, Wang Q, Fish JA, Chai BL, McGarrell DM, Sun YN et al (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research* 42: D633-D642.

Comte J, del Giorgio PA (2010). Linking the patterns of change in composition and function in bacterioplankton successions along environmental gradients. *Ecology* **91:** 1466-1476.

Crump BC, Hopkinson CS, Sogin ML, Hobbie JE (2004). Microbial biogeography along an estuarine salinity gradient: Combined influences of bacterial growth and residence time. *Applied and Environmental Microbiology* **70:** 1494-1505.

Dixon P (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14:** 927-930.

Dodt M, Roehr JT, ;, Ahmed R, ;, Dieterich C (2012). FLEXBAR—Flexible barcode and adapter processing for next-generation sequencing platforms. *Biology* 1: 895-905.

Edgar RC (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26:** 2460-2461.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME [WWW Document]. URL http://www.drive5.com/uchime/ (accessed 6.20.11).

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

Franklin RB, Garland JL, Bolster CH, Mills AL (2001). Impact of dilution on microbial community structure and functional potential: Comparison of numerical simulations and batch culture experiments. *Applied and Environmental Microbiology* **67:** 702-712.

Franklin RB, Mills AL (2006). Structural and functional responses of a sewage microbial community to dilution-induced reductions in diversity. *Microbial Ecology* **52:** 280-288.

Garbeva P, van Veen JA, van Elsas JD (2004). Microbial diversity in soil: selection microbial populations by plant and soil type and implications for disease suppressiveness. *Annu Rev Phytopathol* **42:** 243-270.

Garland JL, Lehman RM (1999). Dilution/extinction of community phenotypic characters to estimate relative structural diversity in mixed communities. *Fems Microbiology Ecology* **30:** 333-343.

Good IJ (1953). The Population Frequencies of Species and the Estimation of Population Parameters. *Biometrika* **40:** 237-264.

Griffiths BS, Ritz K, Wheatley R, Kuan HL, Boag B, Christensen S *et al* (2001). An examination of the biodiversity-ecosystem function relationship in arable soil microbial communities. *Soil Biology & Biochemistry* **33:** 1713-1722.

Hammer et al HD, Ryan PD. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4: 9. pp.**

Hughes JB, Hellmann JJ (2005). The application of rarefaction techniques to molecular inventories of microbial diversity. *Environmental Microbiology* **397:** 292-308.

Koster J, Rahmann S (2012). Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* 28: 2520-2522.

Kuramae EE, Yergeau E, Wong LC, Pijl AS, van Veen JA, Kowalchuk GA (2012). Soil characteristics more strongly influence soil bacterial communities than land-use type. FEMS *Microbiol Ecol* **79:** 12-24.

Lane DJ (1991). 16S/23S rRNA sequencing, p. 115-175. *In EStackebrandt and MGoodfellow* (ed), Nucleic acid techniques in bacterial systematics John Wiley & Sons, Chichester, United Kingdom.

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL *et al* (2006). Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology* **4:** 102-112.

Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD (2012). PANDAseq: PAired-eND Assembler for Illumina sequences. *Bmc Bioinformatics* 13.

Matos A, Kerkhof L, Garland JL (2005). Effects of microbial community diversity on the survival of Pseudomonas aeruginosa in the wheat rhizosphere. Microbial Ecology **49:** 257-264.

McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D et al (2012). The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *Gigascience* 1.

Muyzer G, Dewaal EC, Uitterlinden AG (1993). Profiling of Complex Microbial-Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes-Coding for 16s Ribosomal-Rna. *Applied and Environmental Microbiology* **59:** 695-700.

Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G (2003). Microbial diversity and soil functions. *European Journal of Soil Science* **54:** 655-670.

Navarrete AA, Kuramae EE, de Hollander M, Pijl AS, van Veen JA, Tsai SM (2013). Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. *FEMS Microbiol Ecol* **83:** 607-621.

Panke-Buisse K, Poole AC, Goodrich JK, Ley RE, Kao-Kniffin J (2015). Selection on soil microbiomes reveals reproducible impacts on plant function. *ISME J* **9:** 980-989.

Pavon-Jordan D, Karell P, Ahola K, Kolunen H, Pietiainen H, Karstinen T *et al* (2013). Environmental correlates of annual survival differ between two ecologically similar and congeneric owls. *Ibis* **155:** 823-834.

Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH (2013). Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews Microbiology* **11:** 789-799.

Roesch LF, Fulthorpe RR, Riva A, Casella G, Hadwin AK, Kent AD *et al* (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* **1:** 283-290.

Rognes, T., Mahé, F., xflouris, 2015. vsearch: VSEARCH version 1.0.16.

Rousk J, Baath E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG *et al* (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J* **4:** 1340-1351.

Torsvik V, Ovreas L (2002). Microbial diversity and function in soil: from genes to ecosystems. *Current Opinion in Microbiology* **5:** 240-245.

Chapter 4

Functional traits determine rhizosphere selection of bacterial communities

Yan Yan, Eiko E. Kuramae, Mattias de Hollander, Peter G.L. Klinkhamer, Johannes A. van Veen

In revision

Abstract

We studied the selective effects of the rhizosphere on bacterial communities of different diversity by comparing the composition and the functional traits of these communities in soil and rhizosphere. Differences in diversity were established by inoculating into sterilized soils diluted suspensions of the same soil. We used 16S rDNA amplicon sequencing to determine the taxonomical structure of the bacterial communities and a shotgun metagenomics approach to investigate the potential functional diversity of the communities. We found clear differences between the soil and the rhizosphere bacterial community of each dilution at the OTU level. In many cases, the species diversity within a phylum differed significantly between soil and rhizosphere. Network analysis revealed stronger interactions among bacterial OTUs in the rhizosphere than in the soil. The enrichment processes in the rhizosphere selected microbes with particular functional genes related to transporters, Embden Meyerhof Parnas pathway and hydrogen metabolism. The species with particular functional traits that were over-represented in the rhizosphere samples differed between soil and rhizosphere samples. This selection was not random across bacteria with these functional traits. Overall this suggests selection on the bacterial community of the rhizosphere based on functional traits.

Keywords

Microbial diversity | Rhizosphere selection| Community structure| Microbial functions

4.1. Introduction

Loss of biodiversity can have significant consequences for ecosystem processes (Sala et al 2000, Magurran and Henderson 2003, Butchart et al 2010), for example the productivity and stability of ecosystems (Worm and Duffy 2003, McGill et al 2007). Whether or not this effect holds true for microbial communities, which are assumed to have a high degree of functional redundancy, is still a matter of debate. Soil microbes represent the majority of biodiversity in terrestrial ecosystems and are largely responsible for the maintenance of soil quality and functioning (Philippot et al 2013). Deeper knowledge of soil microbial biodiversity and the link with functionality could lead to a better understanding of the importance of biodiversity for the functioning of terrestrial ecosystems.

One of the most fascinating hotspots of activity and diversity in soils is the rhizosphere. The composition of microbial communities and their activities in the rhizosphere have a large impact on the growth and health of plants (Mendes et al 2011, Berendsen et al 2012). The microbial community in the rhizosphere is mainly derived from the surrounding soil community. Therefore, changes in the soil community, for example those brought about by disturbances, are expected to have significant effects on the assembly and final composition of the rhizosphere community.

Although there is an increasing amount of literature that deals with the influence of stochastic and deterministic factors, including soil and plant characteristics, on microbial community assemblage at various taxonomic levels (Langenheder and Szekely 2011, Mendes et al 2011, Stegen et al 2012), the relative contribution of soil and plant characteristics to the process of microbial community assemblage at different functional levels is not yet known. Difficulties in experimental assessment constitute the major obstacle in understanding how microbial diversity is created and affected by factors such as soil and plants. In this study, as in many others (Salonius 1981, Garland and Lehman 1999, Franklin et al 2001, Matos et al 2005, Franklin and Mills 2006, Wertz et al 2006, Hol et al 2010, Philippot et al 2013, Vivant et al 2013), the experimental approach is based on the assumption that the diversity of the microbial community in the soil can be altered by inoculating diluted suspensions in a pre-sterilized soil. Although this method has been used frequently in the past, little is known about how the assembly of bacterial

communities in the soil and in the rhizosphere proceeds after inoculation. More specifically, until now, we have ignored the question of whether or not functional characteristics of the microbial community play a role in the selection of microbial species in soil and rhizosphere and if so, how. Recent advances in high-throughput sequencing now allow for the assessment of both the taxonomic composition and function of the rhizosphere microbiome (Bulgarelli et al 2015), which enables us to address this question.

The major aim of this study was to acquire a better understanding of microbial community selection at both the taxonomic and functional level in soil and rhizosphere. In order to obtain communities differing in diversity, we inoculated serial dilutions of suspensions into original sterilized soil. After an established incubation period, plants were potted in the various soil samples. The plant species we used in this study, *Jacobaea vulgaris*, is one of the most common weeds in The Netherlands. We applied 16S rRNA gene amplicon sequencing to analyze the community structure in the diverse soil and rhizosphere samples and a total DNA shotgun sequencing approach to assess their potential functions. In Chapters 2 and 3 we found that the soil has a strong impact on the assemblage of bacterial communities after incubation of various diluted inocula. We hypothesized that plants will exert a further selection at both taxonomic and functional trait levels. In particular, we studied whether species selection in the rhizosphere exerts an effect on functional traits of the microbes and, if so, whether this selection is random across species with these traits or species-specific.

4.2. Materials and Methods

4.2.1. Soil sampling and plant selection

Thirty liters of soil were collected at a depth of 15 cm from a dune soil in Meijendel, The Netherlands. Soil organic matter content (%) was 9.11 ± 0.36 (n=6), soil pH was 7.4 ± 0.005 (n=6), NO^{3-} content (mg/kg) was 30.43 ± 0.85 (n=6), NH^{4+} content (mg/kg) was 2.23 ± 0.25 (n=6), P content (mg/kg) was 15.16 ± 0.41 (n=6). The soil was sieved and homogenized and stored in 500 g aliquots in plastic bags. One bag of soil was kept separately to prepare the

inoculum. All the soil was sterilized by γ-irradiation (> 25 kGray, Isotron, Ede, the Netherlands). The sterility was tested by spreading 0.5 g of the soil from the inoculum-bag onto TSA and PDA media. No bacterial and fungal growth was observed on agar plates with the sterilized soil after 6 days for 6 replicates. Three sterilized soil bags inoculated with sterilized water were used as a control for the community assemblage during the entire experimental period. A subsample of the fresh soil was used to determine soil moisture (24 h, 105 ºC). For the dilution treatments, a 10 % suspension of untreated soil in sterilized water (10^{-1}) was sequentially diluted to obtain further dilutions 10^{-6} and 10^{-9} and these were added to the bags with the sterilized soil. The 10^{-1} suspension was considered to be the undiluted treatment.

Jacobaea vulgaris was selected as the plant species. Seeds were collected in Meijendel (52°9'N, 4°22'E), The Netherlands. One seed was propagated by tissue culture (Joosten et al 2009). Since tissue culture has often been defined as the "sterile" plant, it was reasonable to use the "clean" cloned plants for the further experiments. After 8 weeks of incubation of the inoculated soils, tissue culture plants were potted in 0.5 L pots containing the incubated soil. Samples were taken from the bulk soil at the moment of planting. After 6 weeks of plant growth, plants were harvested and gently shaken to remove the loosely adhered soil after which rhizosphere soil samples were collected by removing the remnant soil with a fine sterile brush. Samples were stored at -20 ºC for further analysis. The design of the experiment included 3 dilutions with 6 replicates each for both the incubated bulk soil and rhizosphere soil samples. Given that during plant growth the soil was only isolated (by a layer of tin foil) from the atmosphere, we considered the possibility that this could constitute an unknown source of bacteria. However, we assumed that this would not have a major effect on our results as we know that the bulk soil had a full grown community of over 10^9 cells per gram of soil after the 8-week pre-incubation period following inoculation with the (un-) diluted suspensions (Chapter 2). The impact of bacterial and extracellular DNA left in soil after sterilization prior to inoculation was accounted for by subtracting the OTUs found in the non-inoculated samples from those detected in the inoculated samples (Chapter 2).

4.2.2. DNA extraction, PCR reaction and 16S rDNA gene fragment sequencing

Total DNA was extracted from the incubated bulk soil and rhizosphere soil to determine the composition of the respective microbial communities by 454-pyrosequencing of the 16S rDNA genetic marker. The DNA was extracted using the MoBio Power Soil Extraction Kit according to the supplier's manual (MO BIO Laboratories, Carlsbad, CA, USA). Total DNA concentration was quantified on an ND-1000 spectrophotometer (Nanodrop Technology, Wilmington, DE). PCRs were performed using 5 µM of each forward (515F) and reverse (806R) bar-coded primers (Bergmann et al 2011), 5 mM dNTPs (Invitrogen, Carlsbad, CA), 1 unit of *Taq* polymerase (Roche, Indianapolis, IN), and 5 ng/µl of sample DNA as the template in a total volume of 25 µl with a PCR program of 95 ºC for 5 min, followed by 25 cycles each of 95 s for 30 s, 52 ºC 1 min and 72 ºC for 10 min. To detect any contamination during PCR preparation, negative controls (water in place of DNA) were included for all PCR reactions. PCR products of each subsample from the barcoded primers were generated in six replicates and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega). Equimolar purified PCR products that were quantified by picogreen assays were mixed and sequenced using Roche Genome Sequencer FLX Titanium 454 sequencing platform (Macrogen, Seoul, Korea).

4.2.3. Amplicon sequence analysis

The raw data was processed using the QIIME v.1.6.0 pipeline (Caporaso et al 2010). Low quality sequences below 150 bp in length or with an average quality score below 25 were removed. After denoising the sequences using Denoiser 0.91 (Reeder and Knight 2010), and testing for chimeras using USEARCH (Edgar et al 2011), Operational Taxonomic Units (OTUs) were identified using the UCLUST 1.2.21 algorithm (Edgar 2010) with a phylotype defined at the 97% sequence similarity level. The resulting OTUs were aligned against the Ribosomal Database Project database (Cole et al 2009).

4.2.4. Metagenomics library preparation for DNA shotgun sequencing

Shotgun metagenomic analyses were conducted on the soil DNA extracts following the illumine Pair-End Prep kit protocol with sequencing performed using 2×300 bp sequencing run on the Illumina Miseq2000 (Macrogen Inc. Company, South Korea). Paired end reads were trimmed using Sickle (Joshi and Fass, 2011) with a minimum PHRED score of 30 and at least 150 bp in length. Next, a co-assembly of all data was made with Spades 3.1.1 (Bankevich et al 2012) at different k-mer lengths of 31,91,101 and 121. On the final assembly, genes were predicted using Prodigal 2.61 (Hyatt et al 2010) and converted from GFF (General Feature Format) to GTF (General Transfer Format) using cufflinks 2.1.1 (Trapnell et al 2010). Per sample reads were mapped to contigs using BamM 1.4.1 (Imelfort 2015) that uses BWA 0.7.12-r1039 (Li and Durbin 2009) and samtools 1.2 (Li et al 2009). Next, the number of reads per sample mapping to genes was calculated using featureCounts (Liao et al 2014). To annotate the set of genes, hmmsearch 3.0 (Finn et al 2015) was used to screen the FOAM (Prestat et al 2014) set of Hidden Markov Models (release 1.0). Scripts provided by FOAM were used to select the best hit in the database. For each gene the best KO hits were added to the count matrix of featureCounts as a single column. Next the KO column was aggregated using the Python Pandas library (McKinney 2015). Hits to multiple KO terms were split. Finally for each FOAM level a count matrix was made. The full analysis pipeline has been implemented in a Snakemake workflow (Koster and Rahmann 2012).

4.2.5. Statistical analysis

Alpha diversity calculations were performed based on the rarefied OTU table to compare the diversity among samples at a given level of sampling effort (Hughes and Hellmann 2005). The OTU table was rarefied to 1,535 reads by "single rarefaction" QIIME script since this number was the lowest number of reads for all samples. The average sequence reads from 3 sterilized controls were used as a baseline that was subtracted from the reads of all samples. The OTU table after this subtraction was used for further statistical analysis. We determined Chao1 richness, Simpson and Shannon diversity indices with the "vegan" package (Dixon 2003) in R (The R Foundation for Statistical

Computing). The percentage coverage was calculated by Good's method using the formula: % coverage = [1-(n/N)] × 100, where n is the number of phylotypes represented by singletons and N is the total number of sequences (Good 1953).

Principal coordinates analysis (PCoA) matrices were used to visualize the community structure among samples, using the generated taxonomic and functional abundance matrices. The PCoA plots were generated from Bray-Curtis similarity index matrices of all samples and created using the PAST software program (Hammer et al 2001). Differences in bacterial community composition among treatments were tested by analysis of similarities (ANOSIM). Differential abundance of taxa and functional genes were performed using the "ggtern" package in R to rank taxa down to the genus level and level 2 of functional data (FOAM Database) according to the contributions of the dilution groups. The functions that were selected differed significantly between soil and rhizosphere for at least two dilutions and the differences between soil and rhizosphere were in the same direction for all three dilutions.

Network analyses were performed to gain a better understanding of the microbial interactions in the soil and rhizosphere. Correlations amongst all OTUs were calculated with the Sparse Correlations for Compositional data algorithm (SparCC) (Friedman and Alm 2012) implemented in mothur (Schloss et al 2009). The OTUs with less than three sequences were filtered since they were poorly represented. Only correlations with values above 0.5 or below -0.5 and a statistically significant *P*-value lower than 0.05 were represented in the network using R (R development Core Team, 2008), which were then visualized with the interactive platform Gephi (Bastian MHS 2009).

All the analyses in this study were based on OTUs, except for diversity analysis within particular phyla that were based on the family level.

4.3. Results

4.3.1. Diversity of the bacterial community in soil and rhizosphere

Remarkably, dilution had a stronger effect on the diversity indices than had the rhizosphere selection and in most cases the number of species detected and the diversity indices were similar or higher in the rhizosphere than in the bulk soil (Table 4.1).

Table 4.1. Estimators of microbial diversity and coverage in incubated soils and rhizosphere.

Treatment	Dilution	S.obs	S.chao-1	Shannon	Simpson	Good's estimator of coverage
Soil	10^{-1}	107.20±1.27	134.37±2.96	3.719±0.019	0.954±0.002	97.56±0.11
Rhizosphere	10^{-1}	113.88±2.34	141.85±6.99	3.747±0.049	0.952±0.004	97.77±0.14
Soil	10^{-6}	70.09±2.13	89.64±4.46	3.208±0.040	0.934±0.004	97.95±0.21
Rhizosphere	10^{-6}	85.25±1.28	110.78±4.38	3.334±0.046	0.928±0.005	98.24±0.10
Soil	10^{-9}	55.83±1.14	81.82±3.37	2.633±0.042	0.867±0.006	97.27±0.24
Rhizosphere	10^{-9}	76.36±2.45	95.83±3.92	3.209±0.097	0.916±0.012	98.23±0.16
Two-way ANOVA	Dilution	F=301.5 ***	F=65.26 ***	F=46.11 ***	F=104.8 ***	
	Soil/rhizosphere	F=82.49 ***	F=14.01 ***	F=7.032 **	F=28.85 ***	
	Interaction	F=7.613 **	F=1.332	F=10.9 ***	F=14.15 ***	

Estimators and statistical significance were calculated for each dilution treatment of soil and rhizosphere samples (n = 5-6) based on phylogenetic profiles at the species level. S.obs is the observed number of OTUs. NS means not significant. Results from two-way ANOVA comparisons of estimators diversity are given the *F*-statistic and *P* value (indicated by asterisk: *** $P < 0.001$; ** $P < 0.01$).

The dominant phyla detected in this experiment had contrasting reactions to the presence of plants; the (Shannon) diversity indices for dominant phyla were higher in the rhizosphere than in the bulk soil, and vice versa (Table 4.2). The strongest differences between the diversity indices of the soil *versus* the rhizosphere samples were found in the undiluted 10^{-1} inocula. The rhizosphere samples showed more statistically significant differences within various phyla than did the soil samples for the diluted inocula. Good's estimator of coverage was above 97%.

Table 4.2. Shannon diversity within major phyla in incubated soil and rhizosphere samples.

Phylum/Family	Soil	Rhizosphere	*P*	Soil	Rhizosphere	*P*	Soil	Rhizosphere	*P*
	10^{-1}			10^{-6}			10^{-9}		
Acidobacteria	1.16±0.04	1.39±0.04	*	0.85±0.07	0.75±0.10	NS	0.54±0.15	0.57±0.11	NS
Actinobacteria	2.34±0.03	1.75±0.08	*	1.78±0.07	1.19±0.13	*	1.46±0.16	1.38±0.15	NS
Bacteroidetes	1.29±0.04	1.14±0.05	*	1.27±0.06	1.08±0.08	NS	1.16±0.07	1.31±0.05	NS
Firmicutes	1.04±0.04	0.91±0.04	NS	0.23±0.12	0.90±0.07	*	0.52±0.11	0.92±0.06	*
Verrucomicrobia	1.23±0.03	1.34±0.09	*	0.96±0.06	0.98±0.10	NS	0.81±0.07	0.77±0.11	*
Alphaproteobacteria	1.88±0.02	2.14±0.01	*	1.69±0.04	1.95±0.03	*	1.37±0.12	1.99±0.04	*
Betaproteobacteria	1.50±0.03	1.25±0.01	*	0.75±0.14	0.91±0.08	NS	0.91±0.08	0.47±0.10	*
Deltaproteobacteria	1.31±0.08	1.16±0.11	NS	0.78±0.13	0.87±0.08	NS	0.74±0.09	0.87±0.12	NS
Gammaproteobacteria	0.94±0.04	1.11±0.07	*	0.95±0.07	0.72±0.09	NS	0.47±0.11	0.66±0.12	NS

Diversity and statistical significance ($P < 0.05$) was calculated for each dilution of incubated soil and rhizosphere samples (n = 5-6) within the major phyla based on phylogenetic profiles at the family level. NS means not significant.

4.3.2. Effects of dilution, soil and plant on bacterial community composition

After aligning OTUs with the RDP database, we identified the most dominant phyla in all samples, *i.e., Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, *Acidobacteria*, *Verrucomicrobia*, *Planctomycetes* and *Firmicutes* (Fig. 4.1A). Information on the most relevant patterns in the relative abundances at the phylum and family level is provided in Figure 4.1 A and B.

To visualize differences in community structure between the six groups (three dilutions for the incubated soil and the rhizosphere), taxonomic abundances were used to compute the Bray-Curtis similarity matrices (Fig. 4.2A). Rhizosphere samples were clearly separately from the incubated soil samples (ANOSIM, R = 0.36, $P < 0.001$). A PCoA representing the taxonomic compositions of the soil samples showed a strong separation of the three dilutions (Fig. 4.2C; R= 0.80, $P < 0.001$). In contrast, rhizosphere samples of the three dilutions were more clustered together although still distinct (Fig. 4.2E; R = 0.49, $P < 0.001$).

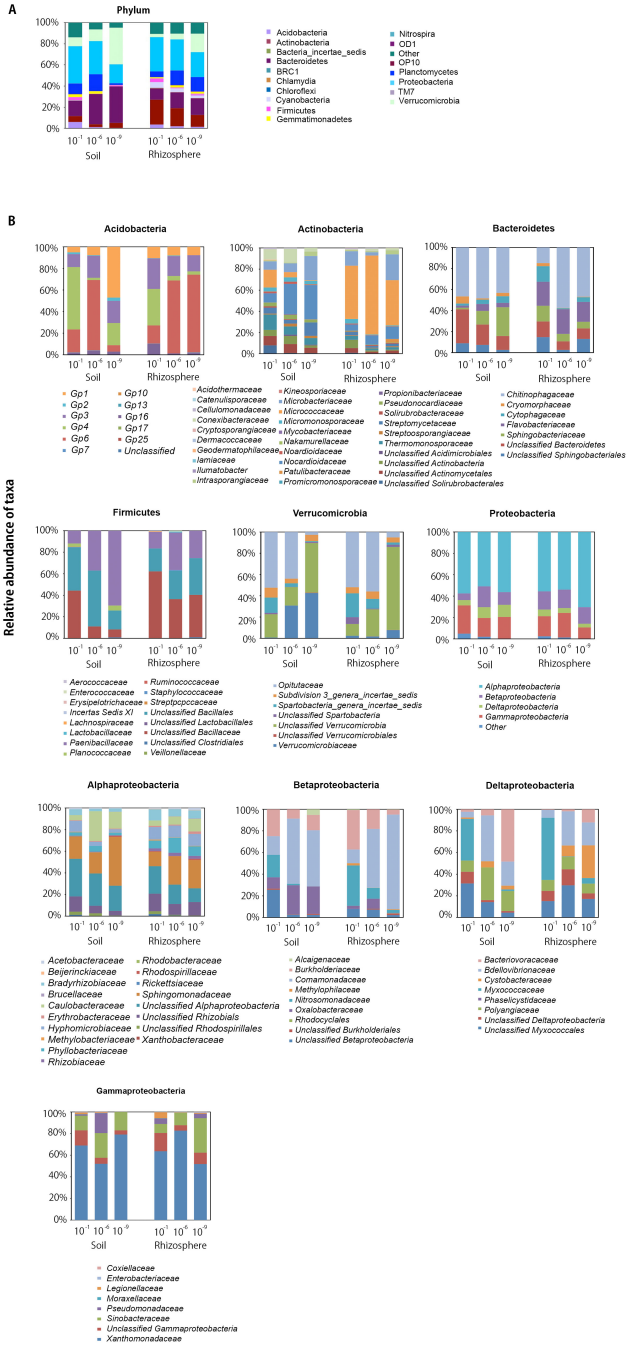

Figure 4.1. Profiles of soil and rhizosphere bacterial communities at the phylum (A) and family level (B) expressed as relative abundances.

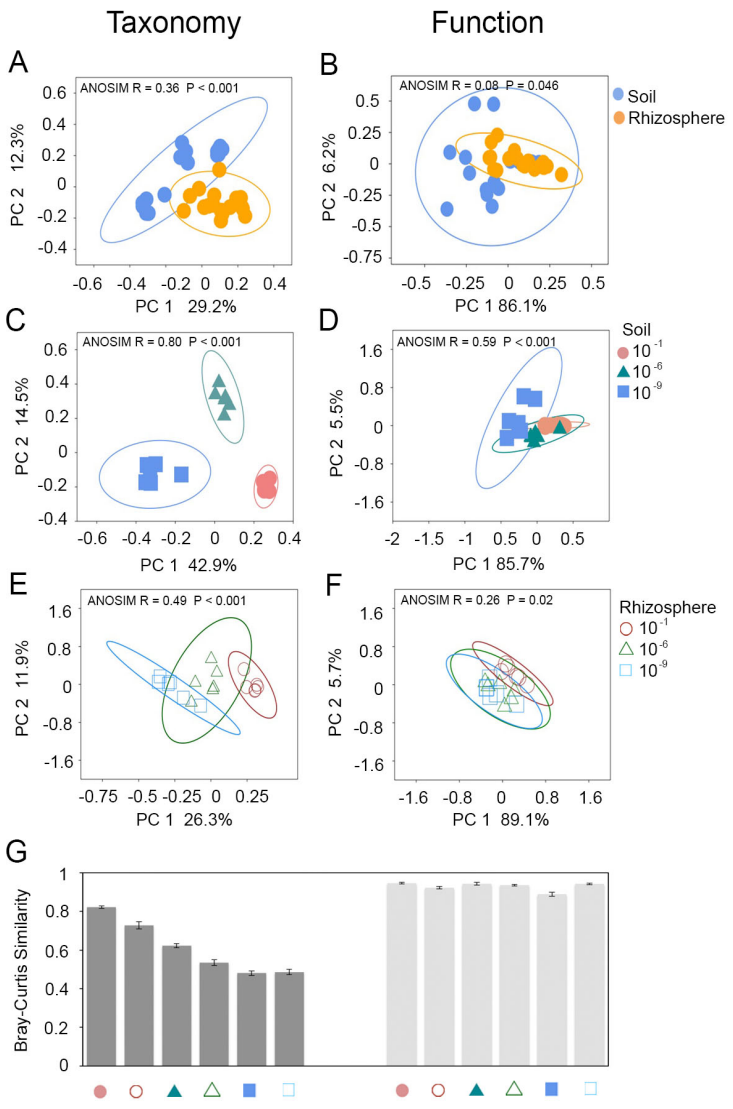

Figure 4.2. Principal Coordinate Analysis of the soil and rhizosphere bacteria community compositions and functional traits. (A) Variation between samples of soil and rhizosphere based on Bray-Curtis similarity for taxonomical data and (B) functional traits using relative abundances based on FOAM 'level 1'. Variation between dilutions of soil samples based on Bray-Curtis similarity for taxonomical data (C) and functional traits (D). Variation between dilutions of rhizosphere samples based on Bray-Curtis similarity for taxonomical data (E) and functional traits (F). Similarity values (analysis of similarity) are shown in the upper left of each plot. Similarities between replicates of each dilution are shown in (G); dark grey bars represent taxonomical data, light grey bars refer to functional traits. The error bars show standard errors of six replicates.

There were marked differences in the network analysis of the soil and rhizosphere samples for all three dilutions (Fig. 4.3A and B). In general, the number of correlations in the rhizosphere was larger than in the soil (Table 4.3), and the number of positive correlations was higher than negative ones for both soil and rhizosphere samples. Between-ness Centrality (BC) of the rhizosphere community networks was much stronger than that of the soil communities, decreasing gradually upon dilution (Fig. 4.3C). In the 10^{-9} diluted samples of the rhizosphere communities, no potential keystone species were obtained (Fig. 4.3C).

Table 4.3. Number of network correlations as inferred by sparCC.

Treatment	Number of nodes	Total number of significant correlations	Number of significant positive correlations	Number of significant negative correlations
Soil 10^{-1}	52	54	40	14
Soil 10^{-6}	59	84	54	30
Soil 10^{-9}	36	38	28	10
Rhizosphere 10^{-1}	119	471	309	162
Rhizosphere 10^{-6}	63	100	59	41
Rhizosphere 10^{-9}	73	102	69	33

4.3.3. Effects of dilution, soil and plant on the functional potential of the bacterial community

The functional profiles of rhizosphere samples were separated from the incubated soil samples based on Bray-Curtis similarity matrices (Fig. 4.2B, R= 0.08, $P = 0.046$). The PCoA plot of functional profiles of the different dilutions of rhizosphere samples showed a higher similarity than those of soil samples (Fig. 4.2D and F; Soil: R = 0.59, $P = 0.0001$; Rhizosphere: R = 0.25, $P = 0.02$). The functional profiles of the soil samples differed significantly among the dilutions, but in rhizosphere the only significant difference in the functional profiles was between the undiluted (10^{-1}) and the most diluted samples (10^{-9}).

The functional profiles of the soil and rhizosphere communities overlapped more as compared to the species community structures (Fig. 4.2). Similarly, the functional genes of all three dilutions of both soil and rhizosphere samples were more strongly centered in the ternary plot than were OTUs (Fig. 4.4C and D). To compare the similarity among replicate samples of the six groups, we calculated the mean values of Bray-Curtis similarity for both the

taxonomic and functional data. Replicates of functional data within each dilution were highly similar (light gray bars in Fig. 4.2G), whereas the taxonomic similarity decreased upon dilutions for both soil and rhizosphere samples (dark gray bars in Fig. 4.2G).

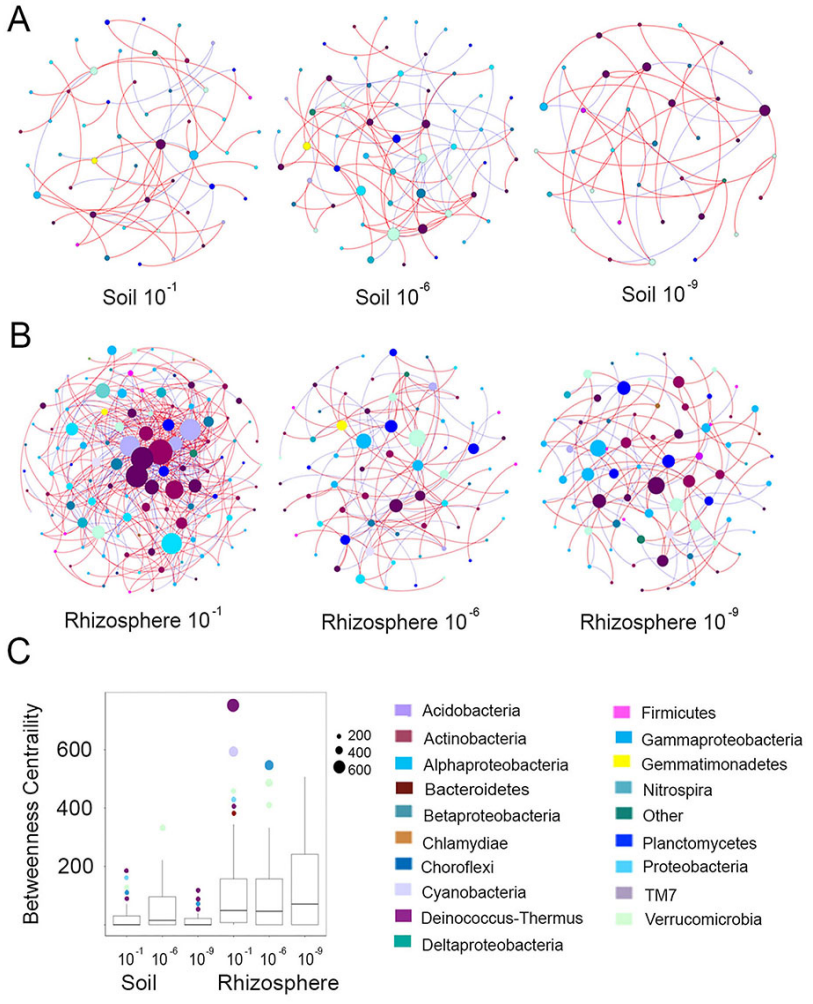

Figure 4.3. Co-occurrence patterns of bacteria in soil and rhizosphere. Correlations were presented in the soil samples (A) and in the rhizosphere samples of each dilution (B). Nodes indicate taxonomic affiliation at genus level. Red lines indicate positive correlations, and blue lines indicate negative correlations. The color of each node indicates the phylum shown below of the figures. The size of each node is proportional to the Betweenness Centrality (C). The box-and-whiskers graphics show the median of betweenness centrality as a line, the 25th and 27th percentiles of the data as the top and bottom of the box, and outlier dots to indicate the most extreme data point within 1.5*(75th to 25th percentile) of the median. The size of outlier data points corresponds to the value of the Betweenness Centrality.

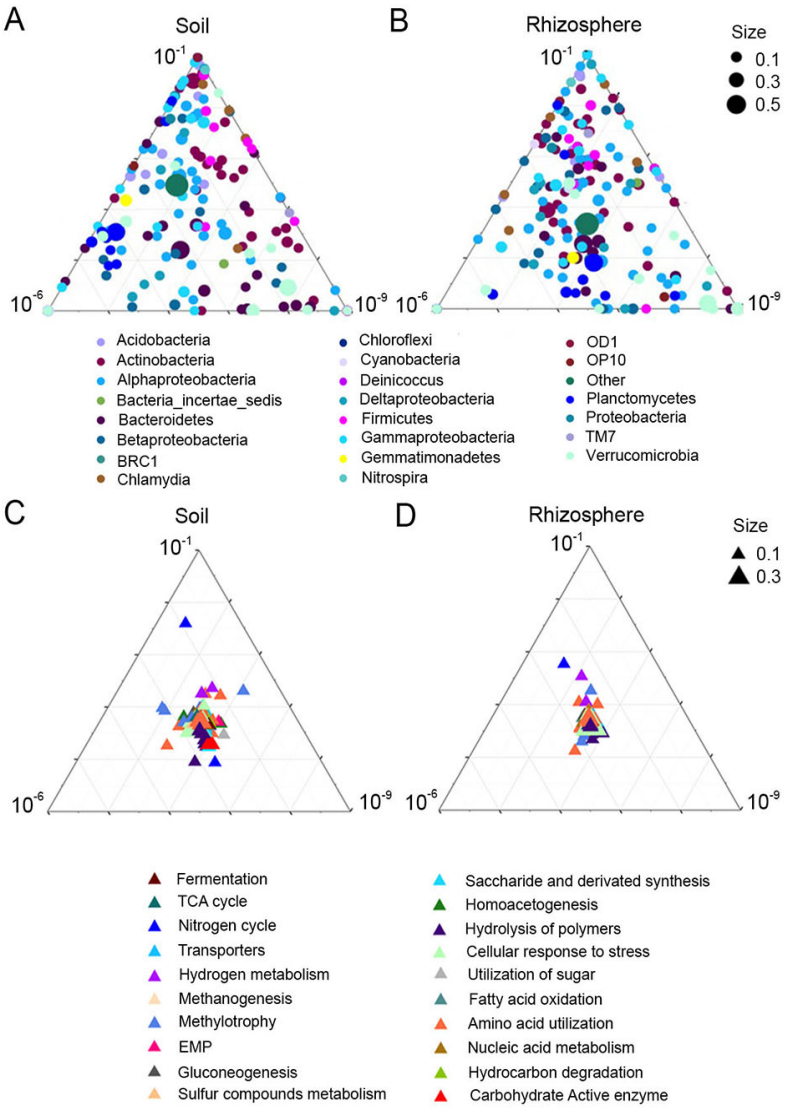

Figure 4.4. Distribution of species and functional traits in each dilution of the soil and rhizosphere samples. Ternary plots of OTUs associated with each dilution in soil (A) and rhizosphere (B) and of functional cores at FOAM 'level 1' associated with each dilution in soil (C) and rhizosphere (D). The position of each point is determined by the contribution of the indicated compartments to the total relative abundance. The size of the dots represents its relative abundance (weighted average). Colors indicate phyla (A and B) and functions at 'level 1' (C and D).

A higher number of significant differences in the functional traits of soil and rhizosphere were observed in the diluted communities than in the undiluted 10^{-1} communities (Fig. 4.5). One of the most abundant types of genes, the transporter genes, was significantly over-represented in the rhizosphere of all samples. This was also observed for the functions related to Embden Meyerhof-Parnas (EMP) pathway and hydrogen metabolism in the rhizosphere of at least two dilutions. By contrast, the core functions related to cellular response to stress and carbohydrate active enzymes were more abundant in the soil than in the rhizosphere.

To further investigate differences in the functional traits of the soil and rhizosphere communities, we binned species within selected functions and then compared the species composition of the soil and the rhizosphere. The functions that were selected differed significantly ($P < 0.05$) in soil and rhizosphere samples in at least two dilutions and these differences were in the same direction for all three dilutions. When testing the functions that were more abundant in the rhizosphere than in the soil, e.g. 'transporters', 'EMP pathway' and 'hydrogen metabolism', we found that rhizosphere communities were clustered and significantly ($P < 0.05$) separated from soil communities (Fig. 4.5B). However, when testing functions that were more abundant in the soil than in the rhizosphere, e.g. cellular response to stress and carbohydrate active enzymes, we observed that soil and rhizosphere communities were not significantly separated (Fig. 4.5B). Although we should be cautious with the interpretation of these results (the analysis is based on only 5 groups of functional traits), this seems to suggest that selection in the rhizosphere is for, rather than against species with particular functional traits.

As an illustration of the changes in the composition of the communities involved in these functions in soil and rhizosphere, we identified the species as detected by metagenomic shotgun data analysis that were involved in the 'transporters' function which differed in abundance between soil and rhizosphere samples. STAMP analysis showed that *Phyllobacteriaceae*, *Rhizobiaceae*, unclassified *Rhizobiales* and *Micrococcaceae* were the major families based on PC1 score (with abundance above 1%) responsible for the PCA separations in the rhizosphere (Fig. 4.5C). In contrast, *Caulobacteraceae*, unclassified *Bacteroidetes,* and, surprisingly, *Pseudomonadaceae* were over-represented in the soil.

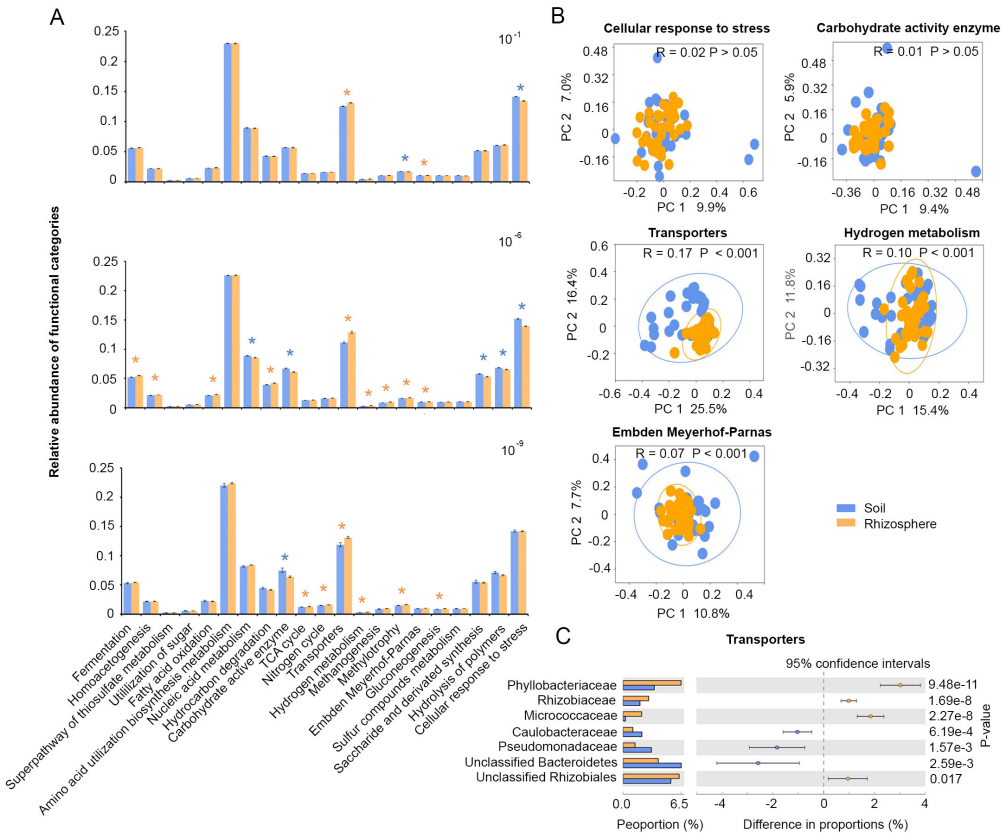

Figure 4.5. Profiles of soil and rhizosphere bacterial functional traits. (A) The relative abundance of groups of functional genes in soil and rhizosphere for three dilutions. Relative abundance of functional genes (FOAM 'level 1') based on normalized shotgun metagenomics data of dilutions of 10^{-1}, 10^{-6} and 10^{-9}. The percentage of the total sequence reads in samples from soil and rhizosphere is presented for each dilution. The error bars show standard errors of six replicates and orange asterisks (*) indicate categories that are more abundant in rhizosphere samples ($P < 0.05$) and blue asterisks (*) indicate categories that are more abundant in soil samples ($P < 0.05$). (B) PCoA plots of species with particular functional genes that were more abundant in the soil than in the rhizosphere (cellular response to stress and carbohydrate activity enzymes) and plots of species with particular functional genes that were more abundant in the rhizosphere than in the soil (transporter genes, Embden Meyerhof-Parnas pathway and hydrogen metabolism). Similarity values are shown in the upper right corner of each plot. The circles represent the clustering of the soil and rhizosphere samples, respectively. (C) Differences in abundance of families with transporter genes between soil and rhizosphere samples (Welch's t-test; $P < 0.05$).

4.4. Discussion

There was a clear separation between soil and rhizosphere samples on the basis of species composition (Fig. 4.2A). A selective change in the microbial community structure of the rhizosphere has also been reported in many other studies (Duineveld et al 1998, Mendes et al 2011, Mendes et al 2014) and plant hosts (Ofek-Lalzar et al 2014, Bulgarelli et al 2015), and soil characteristics (Kuramae et al 2012) may contribute to this. The number of species detected in the rhizosphere was, however, larger than in the bulk soil. Considering that we used sterile plants, it is fair to assume that the plants did not add a substantial inoculum to the community. Presumably, the depth of sequencing is still not sufficient to encompass the entire microbial community in suspensions and soil, although Good's estimator of coverage was always above 0.97.

There was more similarity between the different dilutions of the rhizosphere samples than between different dilutions of soil samples. This shows that convergence took place in the rhizosphere as a direct or indirect selective effect of the roots. This is especially true for the functional traits (Fig. 4.2).

Our results, regarding both species composition and functional traits, clearly indicated that the plant exerts selection on the microbial community in the rhizosphere based on particular functional traits, which may occur directly or through changes in abiotic environmental factors. The enrichment processes in the rhizosphere selected microbes with specific functional genes in particular related to transporters, EMP pathway and hydrogen metabolism. These three functional cores that were over-represented in the rhizosphere suggest that the rhizosphere selects specific species based on functional traits. These functions appeared to be relevant for interactions with the plant. Some of these features have also been shown by others to be important in rhizosphere communities (Mendes et al 2014, Ofek-Lalzar et al 2014, Bulgarelli et al 2015). Consistently with our study, transporter systems were found to be of great importance in the rhizosphere. This was not reported on EMP pathway and hydrogen metabolism.

A clear separation between soil and rhizosphere samples was found for species with particular functional traits only if these were over-represented in the rhizosphere samples (Fig. 4.5B). The latter suggests that the above mentioned rhizosphere selection process across species was not random. As an

example we showed that few specific species containing the "transporters" functions were selected in the rhizosphere. The species found belonged to the families of *Phyllobacteriaceae*, *Rhizobiaceae*, unclassified *Rhizobiales* and *Micrococcaceae*, in particular the genus *Arthrobacter*. These species have been reported earlier as beneficial to plants (Sanguin et al 2009, Hayat et al 2010). Remarkably, 'transports' genes of species belonging to the family of the *Pseudomonadaceae,* which are considered generally as typical rhizosphere organisms (Mendes et al 2011), were found to be more abundant in soil. This may question the role of this family in plant-microbe interactions. We only focused on species involved in transport functions as an illustration of the details of the taxonomic analysis that is possible on the basis of the metagenomics shotgun data. However, these analyses of the composition of species community involved in the functional traits must be taken with caution: in our experience usually only 25-30% of the reads of the assembled shotgun data can be annotated. Thus, in our opinion, the used approach does not allow for more detailed considerations, because of the weak coverage of the sequence data.

Nevertheless the conclusion is justified that the core functional genes selected in the rhizosphere are not restricted to one particular taxonomic group. This is consistent with a report on the *Ulva australis* (marine alga) that showed that they selected functional genes, rather than taxonomic relatedness (Burke et al 2011). If, indeed, the selection process in the rhizosphere is also based on functional traits, and these specific functional traits are not randomly distributed over all bacterial phyla detected here, this may be an explanation for the variation in the taxonomic diversity of the different phyla as presented in Table 4.2.

The network analysis revealed many more correlations and potential keystone species in the rhizosphere than in the soil (Fig. 4.3). This indicates that the network architecture was more stable and had more complex connections in the rhizosphere than in the soil. This is what we expected given the stronger selection observed on the bacterial community in the rhizosphere than in the soil. We based our network analysis on 16S rRNA amplicon data and not on the binned shotgun data because of the above mentioned low annotation rate of the sequences.

In conclusion, we have shown here that the rhizosphere exerts selection on the microbial community also based on particular functional traits. However, to what extent this selection is controlled by the plants or is caused by indirect factors remains to be investigated. At this point, the categorization of the functional genes is too broad to relate these genes to potential effects on plant fitness. We found that the relative abundance of some particular functional genes in the rhizosphere was generally higher than in soil, suggesting that the rhizosphere selects for these functional traits rather than against them. The case in which the relative abundance was clearly higher in soil than in the rhizosphere was for functional traits related to cellular response to stress. This may indicate that the environment in the rhizosphere is less stressful for the bacterial community. On the other hand, the relative abundance of functional transporter genes was significantly higher in the rhizosphere than in soil, clearly showing that selective processes operated on these genes in the rhizosphere. Although the mechanisms and consequences of the functional selection in the rhizosphere for plant fitness remain unclear, the present results add valuable information to better understand the highly complex processes of microbial community assemblage in both soil and rhizosphere.

4.5. References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS *et al* (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* **19:** 455-477.

Bastian MHS JM (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*: San Jose, CA, USA.

Berendsen RL, Pieterse CMJ, Bakker PAHM (2012). The rhizosphere microbiome and plant health. *Trends in Plant Science* **17:** 478-486.

Bergmann GT, Bates ST, Eilers KG, Lauber CL, Caporaso JG, Walters WA *et al* (2011). The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biology & Biochemistry* **43:** 1450-1455.

Bulgarelli D, Garrido-Oter R, Munch PC, Weiman A, Droge J, Pan Y *et al* (2015). Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley. *Cell Host & Microbe* **17:** 392-403.

Burke C, Steinberg P, Rusch D, Kjelleberg S, Thomas T (2011). Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 14288-14293.

Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW, Almond REA *et al* (2010). Global Biodiversity: Indicators of Recent Declines. *Science* **328:** 1164-1168.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Meth* **7:** 335-336.

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al* (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* **37:** D141-D145.

Dixon P (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14:** 927-930.

Duineveld BM, Rosado AS, van Elsas JD, van Veen JA (1998). Analysis of the dynamics of bacterial communities in the rhizosphere of the chrysanthemum via denaturing gradient gel electrophoresis and substrate utilization patterns. *Appl Environ Microbiol* **64:** 4950-4957.

Edgar RC (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26:** 2460-2461.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27:** 2194-2200.

Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F *et al* (2015). HMMER web server: 2015 update. *Nucleic Acids Res* **43:** W30-38.

Franklin RB, Garland JL, Bolster CH, Mills AL (2001). Impact of dilution on microbial community structure and functional potential: Comparison of numerical simulations and batch culture experiments. *Applied and Environmental Microbiology* **67:** 702-712.

Franklin RB, Mills AL (2006). Structural and functional responses of a sewage microbial community to dilution-induced reductions in diversity. *Microbial Ecology* **52:** 280-288.

Friedman J, Alm EJ (2012). Inferring Correlation Networks from Genomic Survey Data. *Plos Computational Biology* **8**.

Garland JL, Lehman RM (1999). Dilution/extinction of community phenotypic characters to estimate relative structural diversity in mixed communities. *Fems Microbiology Ecology* **30:** 333-343.

Good IJ (1953). The Population Frequencies of Species and the Estimation of Population Parameters. *Biometrika* **40:** 237-264.

Hammer et al HD, Ryan PD. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4: 9. pp.**

Hayat R, Ali S, Amara U, Khalid R, Ahmed I (2010). Soil beneficial bacteria and their role in plant growth promotion: a review. *Annals of Microbiology* **60:** 579-598.

Hol WHG, de Boer W, Termorshuizen AJ, Meyer KM, Schneider JHM, van Dam NM *et al* (2010). Reduction of rare soil microbes modifies plant-herbivore interactions. *Ecology Letters* **13:** 292-301.

Hughes JB, Hellmann JJ (2005). The application of rarefaction techniques to molecular inventories of microbial diversity. *Environmental Microbiology* **397:** 292-308.

Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *Bmc Bioinformatics* **11**.

Imelfort M, Woodcroft, B., Parks, D. (2015). Ecogenomics/BamM [WWW Document]. GitHub. URL https://github.com/Ecogenomics/BamM (accessed 7.21.15).

Joosten L, Mulder PPJ, Klinkhamer PGL, van Veen JA (2009). Soil-borne microorganisms and soil-type affect pyrrolizidine alkaloids in Jacobaea vulgaris. *Plant and Soil* **325:** 133-143.

Joshi NA, Fass JN (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. Available at https://github.com/najoshi/sickle.

Koster J, Rahmann S (2012). Snakemake-a scalable bioinformatics workflow engine. Bioinformatics **28:** 2520-2522.

Kuramae EE, Yergeau E, Wong LC, Pijl AS, van Veen JA, Kowalchuk GA (2012). Soil characteristics more strongly influence soil bacterial communities than land-use type. *Fems Microbiology Ecology* **79:** 12-24.

Langenheder S, Szekely AJ (2011). Species sorting and neutral processes are both important during the initial assembly of bacterial communities. *ISMEJ* **5:** 1086-1094.

Li H, Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25:** 1754-1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:** 2078-2079.

Liao Y, Smyth GK, Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30:** 923-930.

Magurran AE, Henderson PA (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature* **422:** 714-716.

Matos A, Kerkhof L, Garland JL (2005). Effects of microbial community diversity on the survival of Pseudomonas aeruginosa in the wheat rhizosphere. *Microbial Ecology* **49:** 257-264.

McGill BJ, Etienne RS, Gray JS, Alonso D, Anderson MJ, Benecha HK *et al* (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* **10:** 995-1015.

McKinney W (2015). Python Data Analysis Library — pandas: Python Data Analysis Library [WWW Document]. URL http://pandas.pydata.org/ (accessed 7.22.15).

Mendes LW, Kuramae EE, Navarrete AA, van Veen JA, Tsai SM (2014). Taxonomical and functional microbial community selection in soybean rhizosphere. *ISMEJ* **8:** 1577-1587.

Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al* (2011). Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **332:** 1097-1100.

Ofek-Lalzar M, Sela N, Goldman-Voronov M, Green SJ, Hadar Y, Minz D (2014). Niche and host-associated functional signatures of the root surface microbiome. *Nature Communications* **5**.

Philippot L, Spor A, Henault C, Bru D, Bizouard F, Jones CM *et al* (2013). Loss in microbial diversity affects nitrogen cycling in soil. *ISMEJ* **7:** 1609-1619.

Prestat E, David MM, Hultman J, Tas N, Lamendella R, Dvornik J *et al* (2014). FOAM (Functional Ontology Assignments for Metagenomes): a Hidden Markov Model (HMM) database with environmental focus. *Nucleic Acids Res* **42:** e145.

Reeder J, Knight R (2010). Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nature Methods* **7:** 668-669.

Sala OE, Chapin FS, Armesto JJ, Berlow E, Bloomfield J, Dirzo R *et al* (2000). Biodiversity - Global biodiversity scenarios for the year 2100. *Science* **287:** 1770-1774.

Salonius PO (1981). Metabolic Capabilities of Forest Soil Microbial-Populations with Reduced Species-Diversity. *Soil Biology & Biochemistry* **13:** 1-&.

Sanguin H, Sarniguet A, Gazengel K, Moenne-Loccoz Y, Grundmann GL (2009). Rhizosphere bacterial communities associated with disease suppressiveness stages of take-all decline in wheat monoculture. *New Phytologist* **184:** 694-707.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology* **75:** 7537-7541.

Stegen JC, Lin XJ, Konopka AE, Fredrickson JK (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISMEJ* **6:** 1653-1664.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ *et al* (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* **28:** 511-U174.

Verginer M, Siegmund B, Cardinale M, Muller H, Choi Y, Miguez CB *et al* (2010). Monitoring the plant epiphyte Methylobacterium extorquens DSM 21961 by real-time PCR and its influence on the strawberry flavor. *Fems Microbiology Ecology* **74:** 136-145.

Vivant AL, Garmyn D, Maron PA, Nowak V, Piveteau P (2013). Microbial Diversity and Structure Are Drivers of the Biological Barrier Effect against Listeria monocytogenes in Soil. *Plos One* **8**.

Wertz S, Degrange V, Prosser JI, Poly F, Commeaux C, Freitag T et al (2006). Maintenance of soil functioning following erosion of microbial diversity. *Environ Microbiol* 8: 2162-2169.

Worm B, Duffy JE (2003). Biodiversity, productivity and stability in real food webs. *Trends in Ecology & Evolution* **18:** 628-632.

Chapter 5

Functional genes rather than taxonomical composition of the rhizosphere bacterial community determine plant biomass in *Jacobaea vulgaris*

Yan Yan, Eiko E. Kuramae, Mattias de Hollander, Johannes A. van Veen, Peter G.L. Klinkhamer

Manuscript

Abstract

The relationship between plants and their surrounding microbiota belowground is complex and has been the focus of much research. In reality, the functionality of the microorganisms that are involved in plant-microbe interactions is still not well understood. We used 16S rRNA gene sequencing to determine the bacterial composition and a shotgun metagenomics approach to determine the functional traits of the rhizosphere microbiome. We examined the effects of the taxonomical composition and the functional traits of the bacterial community in the rhizosphere on plant biomass production. Both were significantly different among soils inoculated with different dilutions of the original bacterial community. Plant biomass production was on average the lowest and showed the highest variation after inoculation of the undiluted communities. A combination of unsupervised multivariate statistics and partial correlations showed that *Arthrobacter* was the taxonomical group that was most strongly related to plant biomass and that 'transporters' genes were the functional genes most strongly related to plant biomass. Both were positively correlated to plant biomass and positively correlated with each other. Specifically, the 'monosaccharide transporters' genes significantly positively correlated to plant biomass when all three dilutions samples were taken together, and this group of genes increased significantly upon dilutions in the rhizosphere. The frequency of 'transporters' genes was higher in *Arthrobacter* than in other components of the bacterial community. Partial correlation indicated that after taking the frequency of 'transporters' genes into account the correlation between *Arthrobacter* and plant biomass was no longer significant while after taking the frequency of *Arthrobacter* into account the correlation between 'transporters' genes and plant biomass was still highly significant. Although these results should be considered with caution this seems to suggest that functional genes rather than the taxonomical composition of the bacterial community of the rhizosphere determine plant biomass production.

Keywords

Rhizosphere metagenome | Functional traits | Unsupervised multivariate analyses | Plant biomass

5.1. Introduction

The microbiome of the rhizosphere plays critical roles in the functioning of terrestrial ecosystems (Philippot et al 2013). The rhizomicrobiome drives and responds to the specificity of its environment, including host plant characteristics (Haichar et al 2008, Bulgarelli et al 2012), and factors such as pH, salinity, moisture and the availability of nutrients (Fierer and Jackson 2006, Logue and Lindstrom 2010, Nemergut et al 2010, Brockett et al 2012). The rhizomicrobiome plays a key role in plant development and the productivity of the aboveground vegetation (van der Heijden et al 1998, Wagg et al 2011).

Despite the general acceptance that plant roots select specific microbial species which directly or indirectly influence host plant physiology and development (Mendes et al 2011), the extent to which functional traits linked with the rhizosphere microbiome determine colonization and impact on the host plant remains largely unknown. Therefore, characterization of the functional traits of the rhizosphere microbiome is crucial for understanding the effect of soil-borne microbes on plant development.

Advanced shotgun metagenomics approaches offer promising tools to target the microbial genes related to host plant-microbe interactions and so the associated functions. Current studies using this approach that focus on describing the microbiome of humans or other mammal host revealed that the microbiome composition and functions are determinative for the physiology of the host (Turnbaugh et al 2006, Tremaroli and Backhed 2012). Transcription analyses of bacterial genes in the rhizosphere have mostly been performed on single rhizobacterial strains (Mark et al 2005, Matilla et al 2007, Dennis et al 2010). However, because of the complexity of the rhizomicrobiome and the inability to culture many microorganisms, comprehensive, overall, pictures of the microbial community and its functionality related to its link with host plant productivity in natural ecosystems are scarce (Ofek-Lalzar et al 2014, Bulgarelli et al 2015). Thus, in order to improve our understanding of the mechanisms of plant-microbe interactions, we need to characterize better the fundamental ecological processes that underlie the composition and the functionality of the rhizomicrobiome.

The major aim of this study was to acquire better understanding of the relationship between the rhizosphere microbiome and plant growth both at the

level of the taxonomical composition and at the level of the functional genes of the bacterial community. To establish differences in the microbial communities and differences in plant growth, serial dilutions of a soil suspension were prepared and the obtained inocula were, subsequently, re-inoculated into the original soil previously sterilized by γ-irradiation. After an incubation period, plants were potted in the soil samples. We used *Jacobaea vulgaris*, one of the most common weeds in the Netherlands. We used 16S rRNA gene sequencing to assess the composition of the bacterial community in the rhizosphere and a total DNA shotgun metagenomics approach to assess the potential microbiome functionality. In Chapter 4 we already showed that the selection of rhizosphere microbial communities from soil communities was strongly based on the functional traits of the selected microbes. So, here, we hypothesized that selection on the basis of particular functional traits will also have a strong impact on plant growth. We addressed three basic questions: 1) Is plant growth related to the taxonomical composition of bacterial communities in the rhizosphere? 2) Is plant growth related to the frequency of particular functional genes in the rhizosphere? 3) Is the taxonomical composition related to the frequency of particular functional genes and if so which of the two is most strongly related to plant growth?

5.2. Materials and methods

5.2.1. Soil sampling and plant selection

Thirty liters of soil were collected at a depth of 15 cm from a dune soil in Meijendel, The Netherlands. The soil had a sandy texture, an organic matter content of 9.1%, pH of 7.4 and the ammonium, nitrate and phosphorus content of 30.4 mg/kg, 2.2 mg/kg and 15.2 mg/kg respectively. The soil was sieved and homogenized and stored in 500 g aliquots in plastic bags. One bag of soil was kept separately to prepare the inoculum. The soil was sterilized by γ-irradiation (> 25 kGray, Isotron, Ede, the Netherlands). Sterility was tested by spreading 0.5 g of the soil from the inoculum-bag onto TSA and PDA media. No bacterial and fungal growth was observed on agar plates after 6 days for 6 replicates. A subsample of the fresh soil was used to determine soil moisture (24 h, 105 ºC). For the dilution treatments, a 10 % suspension of untreated soil in sterilized water (10^{-1}) was sequentially diluted to obtain further dilutions of 10^{-6} and 10^{-9}

and these were added to the sterilized soil. The 10^{-1} suspension was considered to be the undiluted treatment.

Jacobaea vulgaris was used as study plant species. Seeds were collected in Meijendel (52°9'N, 4°22'E), The Netherlands. One seed was propagated by tissue culture. This genotype showed a strong negative feedback in the inoculated soil compared to growth in sterile soil in a previous study (Joosten et al 2009). Since tissue culture plants are more or less "sterile", it was reasonable to use this "clean" plant for the experiments. After 8 weeks of incubation of the inoculated soils, at the moment that the regrown microbial communities reached similar abundances (Chapter 2), tissue culture plants were potted in 0.5 L pots containing the incubated soil. Samples were taken from the bulk soil at the moment of planting. Plants were grown randomly distributed in a climate room (relative humidity 70%, light 16h at 20 ºC, dark 8h at 20 ºC). Sterile demineralized water was given every two days with additions of 10 ml nutrient solution (Steiner 1968) once every two weeks, in order to avoid nutrient limitation to plant growth. After 6 weeks of plant growth, plants were harvested and gently shaken to remove the loosely adhering soil after which rhizosphere soil samples were collected by removing the remnant soil with a fine sterile brush. Soil samples were stored at -20 ºC for further analysis. Harvested plant parts (shoots and roots) were freeze-dried at -80 ºC for one week until constant weight. The design of the experiment included 3 dilutions, with 6 replicates each and duplicate samples per replicate for both the incubated soil and rhizosphere samples. Given that during plant growth the soil was only isolated by a layer of tin foil from the atmosphere, there is a possibility that this could constitute an unknown source of bacteria. However, we assumed that this would not have a major effect on our results as we know that the bulk soil had a full grown community of over 10^9 cells per gram of soil after the 8-week pre-incubation period in closed bags following inoculation of the (un-) diluted suspensions as found in Chapter 2.

5.2.2. Amplicon sequence analysis

The raw data was processed using the QIIME v.1.6.0 pipeline (Caporaso et al 2010). Low quality sequences below 150 bp in length or with an average quality score below 25 were removed. After denoising the sequences using

Denoiser 0.91 (Reeder and Knight 2010), and testing for chimeras using USEARCH (Edgar et al 2011), Operational Taxonomic Units (OTUs) were identified using the UCLUST 1.2.21 algorithm (Edgar 2010) with a phylotype defined at the 97% sequence similarity level. The resulting OTUs were aligned against the Ribosomal Database Project database (Cole et al 2009).

5.2.3. Metagenomics library preparation for DNA shotgun sequencing

Shotgun metagenomic analyses were conducted on the soil DNA extracts (according to the supplier's manual (MO BIO Laboratories, Carlsbad, CA, USA) following the Illumina Pair-End Prep kit protocol with sequencing performed using 2×300 bp sequencing run on the Illumina Miseq2000 (Macrogen Inc. Company, South Korea).

Paired end reads were trimmed using Sickle (Joshi and Fass 2011) with a minimum PHRED score of 30 and at least 150 bp in length. Subsequently a co-assembly of all data was made with Spades 3.1.1 (Bankevich et al 2012) at different k-mer length of 31,91,101 and 121. Following the final assembly genes are predicted using Prodigal 2.61 (Hyatt et al 2010) and converted from GFF (General Feature Format) to GTF (General Transfer Format) using cufflinks 2.1.1 (Trapnell et al 2010). Per sample, reads were mapped to contigs using BamM 1.4.1 (Imelfort 2015), which uses BWA 0.7.12-r1039 (Li and Durbin 2009) and samtools 1.2 (Li et al 2009). Next the number of reads per sample mapping to genes was calculated using featureCounts (Liao et al 2014). To annotate the set of genes hmmsearch 3.0 (Finn et al 2015) was used to screen the FOAM (Prestat et al 2014) set of Hidden Markov Models (release 1.0). Scripts provided by FOAM were used to select the best hit to the database. For each gene the best KO hits were added to the count matrix of featureCounts as a single column. Thereafter, the KO column was aggregated using the Python Pandas library (McKinney 2015). Hits to multiple KO terms were split. Finally, for each FOAM level a count matrix was made. The whole analysis has been implemented in a Snakemake workflow (Koster and Rahmann 2012).

5.2.4. Data analysis

Alpha diversity calculations were performed based on the rarefied OTU table to compare the diversity among samples at a given level of sampling effort (Hughes and Hellmann 2005). The OTU table was rarefied to 1,535 reads by "single rarefaction" QIIME script since this number was the lowest number of reads for all samples. Four undiluted 10^{-1} samples were filtered out because of the very low number of reads. The average sequence reads from 3 sterilized controls were used as a baseline that was subtracted from the reads of all samples.

All statistical analyses were conducted using R and the vegan package (Dixon 2003). To assess whether manipulation of the bacterial community could explain changes in total plant biomass, ANOVA (False Discovery Rate-corrected) was determined across dilution groups. Data was transformed to fit normal distributions when needed. Unsupervised Principal Component Analysis (PCA) was applied by PAST (Hammer et al 2001). PCAs were performed to visualize the different dilution effects on both taxonomical profiles and functional traits based on normalized functional data. Each broad functional category could be divided in in a subset of functions based on the FOAM dataset. The weight of each taxonomical unit and each functional trait was assigned on the PC score, respectively. In this way the important species functional traits for the PCA separation were distinguished from all other functions on the basis of PC score. The unsupervised analyses were followed up by correlation analyses of the selected potentially important taxonomical units and functional traits with plant biomass. We then used partial correlations to identify the most important taxonomical units and functional traits that were related to plant biomass. As a last step we used again partial correlations with plant biomass to identify whether the taxonomical composition or functional traits were the most important to explain differences in plant biomass.

Network analysis was conducted based on the correlations between the selected functional traits and plant biomass of the undiluted 10^{-1} rhizosphere samples. Significant correlations were identified based on P-values < 0.05, this corresponds to correlation coefficients > 0.5 or < -0.5. The resulting correlation matrix was translated into an association network using Cytoscape 3.2.1 (Shannon et al 2003).

A heatmap was created using the relative abundance of the selected functional traits and classified by R package. The distance used were Pearson correlation for clustering the genes. Partial correlations were calculated by R and the 'ppcor' package.

5.3. Results

5.3.1. Effect of bacterial community composition on plant biomass

Clearly, as was already demonstrated in Chapter 4, dilution and rhizosphere selection led to changes in diversity and structure of the bacterial communities (Fig. 5.1A). The taxonomical profile of rhizosphere samples showed a significant separation amongst three dilutions (Fig. 5.1B; ANOSIM, R = 0.49, $P < 0.01$) with PC1 and PC2 explained 26.3 % and 11.9 % of the observed variation, respectively.

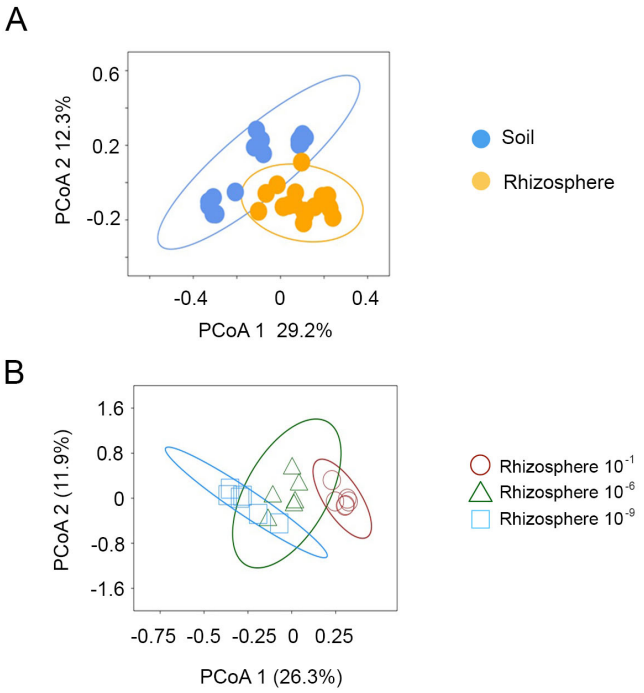

Figure 5.1. PCoA of Bray-curtis similarity matrix among samples using taxonomic profiles based on the relative abundance of OTUs. (A) Variation between samples of soil and rhizosphere. (B) Variation between dilutions of rhizosphere samples.

The plants grew significantly less well in the undiluted 10^{-1} rhizosphere samples compared to the diluted samples (Fig. 5.2). Variation in plant biomass production among replicated samples differed for the different dilutions. In the undiluted 10^{-1} rhizosphere samples, the largest variation in plant biomass amongst replicates was observed.

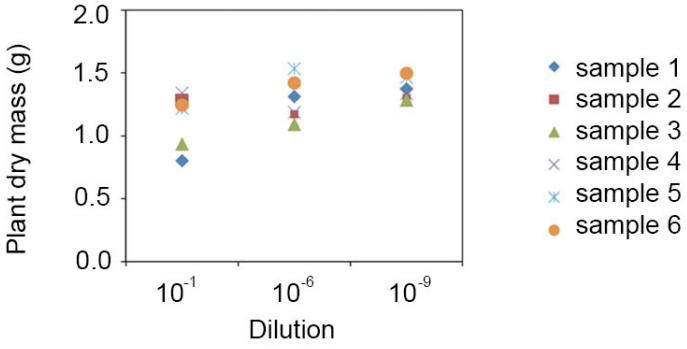

Figure 5.2. Effect of soil microbial communities on plant dry biomass (Mean dry weight, n=12 per dilution).

To determine the species that were potentially responsible for the differences in plant biomass production, we determined Spearman's rank correlation between the PC1 score of the rhizosphere taxonomic profile and plant biomass of the undiluted 10^{-1} samples. Interestingly, the PC1 and PC2 scores significantly correlated with plant biomass in soils inoculated with undiluted 10^{-1} samples (PC1: n = 8, R = -0.91, $P < 0.001$; PC2: R = 0.82, $P < 0.01$). To pre-select OTUs, we zoomed in on PC1 and PC2 of the taxonomical profile and selected species with scores < -0.3 and > 0.3. This resulted in two species from PC1 and three species from PC2 (Fig. 5.3). One group of OTUs (*Arthrobacter*) overlapped so this resulted in four species in total (Fig. 5.4). *Arthrobacter* was negatively correlated with PC1 and positively with PC2, and as expected it showed a positive correlation with plant biomass (Fig. 5.4; n = 8, R = 0.87, $P < 0.01$). *Planctomycetales* was positively correlated with PC2 and as expected it was also positively related to plant biomass (n = 8, R = 0.79, $P < 0.01$). *Verrucomicrobia* and *Chitinophagaceae* that positively correlated to PC1 and PC2, respectively, were not significantly correlated to plant biomass, although the trends were in the direction as expected on basis of their PC scores (Fig. 5.4B).

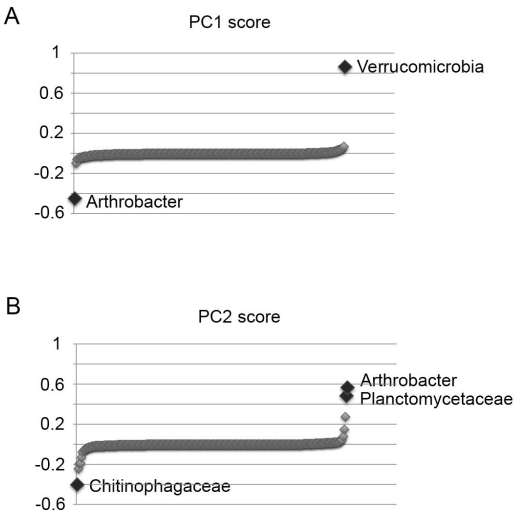

Figure 5.3. Loading plot of Principle Component Analysis (PC1) of the taxonomical profiles with red dots as important factors and gray dots as not important. Species names are indicated near each red dot.

To identify the potentially most important taxonomical unit, *i.e. Arthrobacter* or *Planctomycetaceae* we calculated the partial correlations with plant biomass. After taking *Arthrobacter* into account *Planctomycetaceae* was no longer correlated with plant biomass while after taking *Planctomycetaceae* into account *Arthrobacter* still significantly positively correlated to plant biomass ($n = 8$, $R_p = 0.80$, $P < 0.05$; Table 5.1). Interestingly, we found in our previous paper (Chapter 4) that, indeed, *Arthrobacter* occurred at higher frequency in the rhizosphere compared to bulk soil.

Table 5.1. Partial correlation matrix between the two main species, i.e. *Arthrobacter* and *Planctomycetaceae,* controlling plant biomass.

Variables studied	Arthrobacter	*Planctomycetaceae*	Plant biomass 10^{-1}
Arthrobacter	1.00	- 0.12 (0.79)	0.80^{*} (0.03)
Planctomycetaceae		1.00	0.52 (0.23)
Plant biomass 10^{-1}			1.00

Values indicate partial correlation coefficients (*P*-value) between two species; $*P < 0.05$.

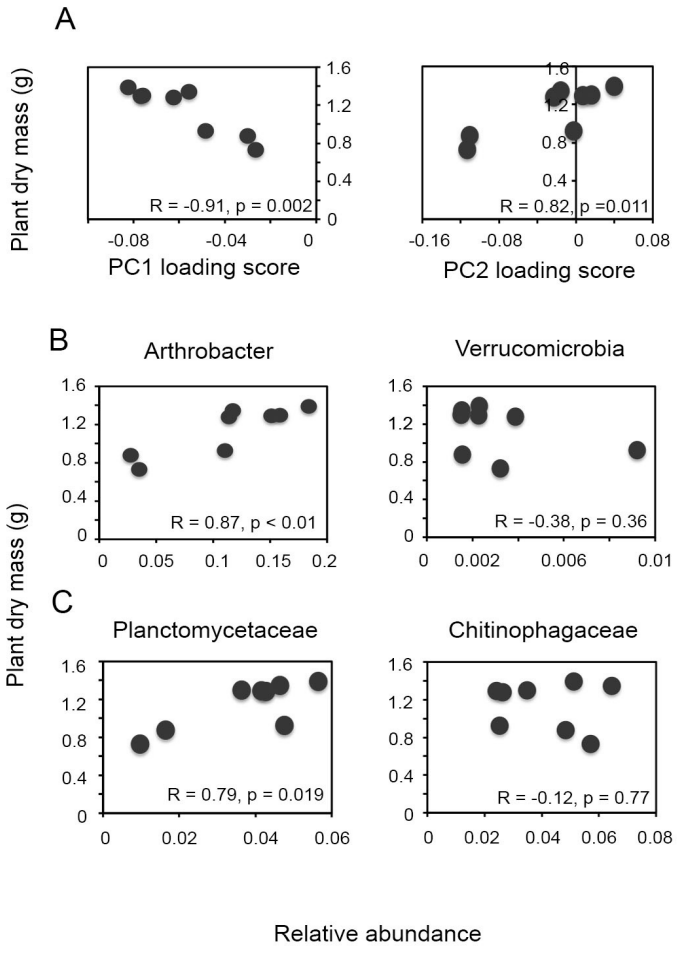

Figure 5.4. Interactions between species and their host plant. (A) Correlations between PC1 and PC2 of the taxonomical profiles in the rhizosphere and the plant biomass of the undiluted 10^{-1} samples; (B) Correlations between plant biomass of the undiluted 10^{-1} samples and the species that were selected from the PC1 score; (C) Correlations between plant biomass of the undiluted 10^{-1} samples and the species that were selected from the PC2 score.

5.3.2. Differences in functional traits among the bacterial communities

To further assess the functional traits responsible for the discrimination amongst the dilutions of the rhizosphere samples, an unsupervised multivariate data analyses, Principal Component Analysis (PCA), was performed on the bacterial functional profile (Fig. 5.5). A PCA of the functional profile based on FOAM Dataset 'level 1' of the rhizosphere samples showed a significant separation amongst three dilutions (ANOSIM, R = 0.41; $P < 0.001$) with PC1 and PC2 explaining 60 % and 18.7 % of the observed variation, respectively.

Based on the PC1 score, the functional traits with scores < -0.3 and > 0.3 were selected. As a result, seven out of twenty-one functional traits ('level 1') were identified as significantly influenced by dilutions (Fig. 5.6).

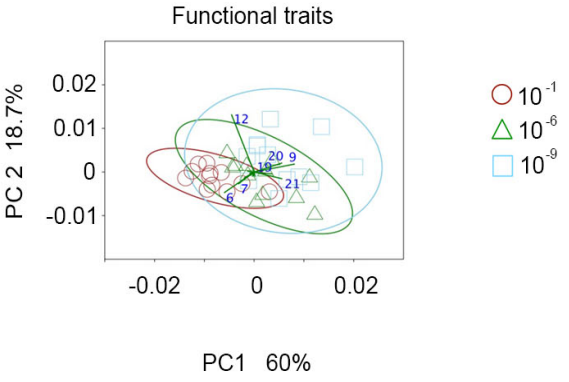

Figure 5.5. Principle component analysis of functional traits of the rhizosphere samples (FOAM 'level 1'). The functions responsible for the PCA separation were indicated in the biplot. 6: amino acid utilization biosynthesis metabolism; 7: nucleic acid metabolism; 9: carbohydrate active enzyme; 12: transporters; 19: saccharide and derivate synthesis; 20: hydrolysis of polymers; 21: cellular response to stress.

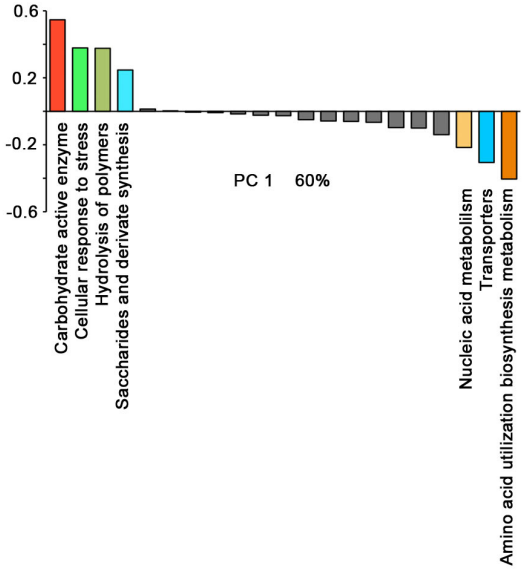

Figure 5.6. Loading plot of Principle Component Analysis (PC1) of functional traits of the rhizosphere samples (FOAM 'level 1') with colored bars as the important factors and the gray bars as not important.

5.3.3. Functional traits related to plant biomass

In order to predict the effect of the first two principal components of the functional profile on plant biomass, Spearman's rank correlations between the first two principal components and plant biomass of each dilution was performed (Table 5.2). For the undiluted 10^{-1} samples, plant biomass was significantly, negatively, correlated with PC1 ($n = 12$, $R = -0.69$, $P < 0.01$) and positively correlated with PC2 ($n = 12$, $R = 0.86$, $P < 0.001$). The trends were the same for the diluted samples except for PC2 of the 10^{-9} dilution (Table 5.2). Based on the biplot (Fig. 5.5), the important functions that were responsible for the differences in the PCA separation, *i.e.* 'level 1': 'amino acid utilization biosynthesis metabolism', 'nucleic acid metabolism', 'carbohydrate active enzyme', 'transporters', 'saccharide and derivate synthesis', 'hydrolysis of polymers' and 'cellular response to stress', were not correlated to plant biomass when all three dilutions are taken together. Thus, in order to examine which functional traits contributed to plant biomass, we focused on the undiluted, 10^{-1}, samples where we observed the largest differences in the plant biomass production. Overall, five out of seven functional traits (at 'level 1') were, on basis of relative gene abundances, significantly correlated with plant biomass (Fig. 5.7).

Table 5.2. Spearman's correlation coefficients between PC1 loading of functional traits and plant dry biomass of each dilution.

Dilution	PC loading	Spearman's coefficient	P value
10^{-1}	PC 1	- 0.69	0.013 **
10^{-6}	PC 1	- 0.33	0.291
10^{-9}	PC 1	- 0.48	0.118
10^{-1}	PC 2	0.86	0.001***
10^{-6}	PC 2	0.52	0.079
10^{-9}	PC 2	- 0.28	0.381

To pre-select functional traits at deeper levels (e.g. 'level 2' or 'level 3'), first we tested the five functional traits ('level 1'), and only for the four ones for which we found a significant correlation with plant biomass we zoomed in at deeper levels of particular functional traits. Correlations between each potential

functional trait and plant biomass were determined to generate a correlation network. For three of the functional traits, we zoomed in at 'level 2', for two others we could zoom in at 'level 3' (Fig. 5.8). The results of the network analysis indicated twelve functional traits belonging to four broad functional categories that were significantly correlated with plant biomass (Fig. 5.9). For one category, i.e. 'carbohydrate active enzymes', no deeper level function was correlated with plant biomass. Functional traits related to 'transporters' and 'nucleic acid mechanism' showed positive correlations with plant biomass. In contrast, functional traits that were related to 'cellular response to stress' and 'saccharide and derivate synthesis' were negatively correlated with plant biomass.

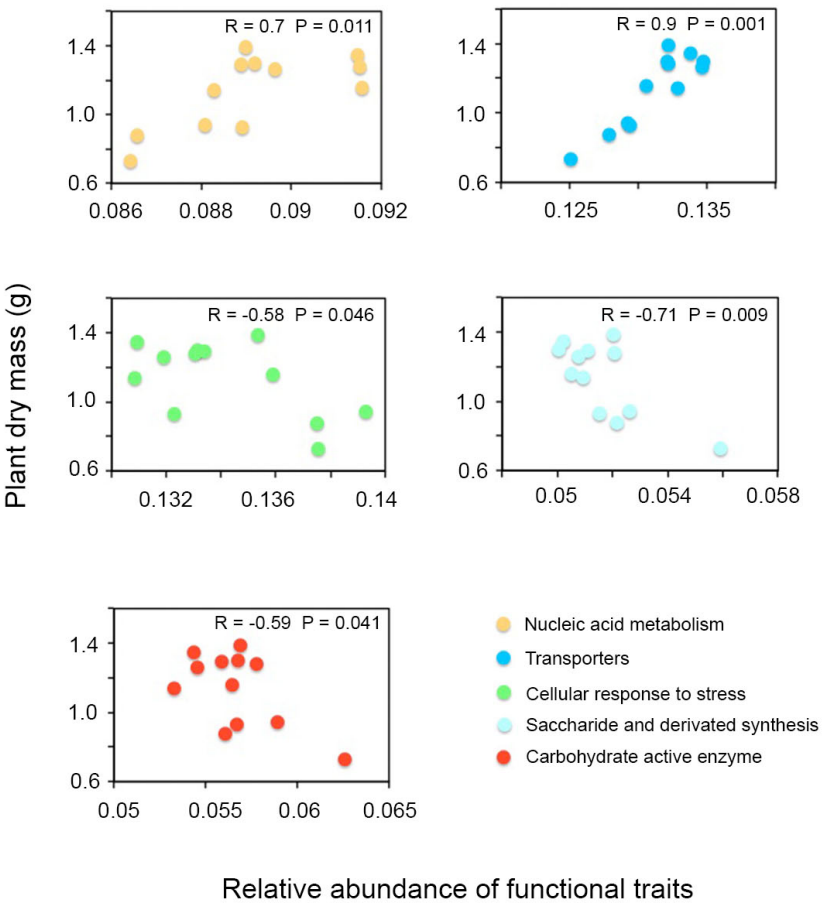

Figure 5.7. Correlations between the functional traits of the rhizosphere samples and the plant biomass of the undiluted 10^{-1} samples. The colour of each dot indicates the functional categories of 'level 1' in the FOAM dataset to which the traits belong.

To further determine the above selected twelve functional genes (e.g. 'level 2' or 'level 3') that could determine the differences in plant biomass when all three dilutions were taken together, Spearman correlations were performed between these functional genes and plant biomass, respectively. 'Monosaccharide transporters' genes significantly positive correlated to plant biomass (n = 36, R = 0.48, $P < 0.01$).

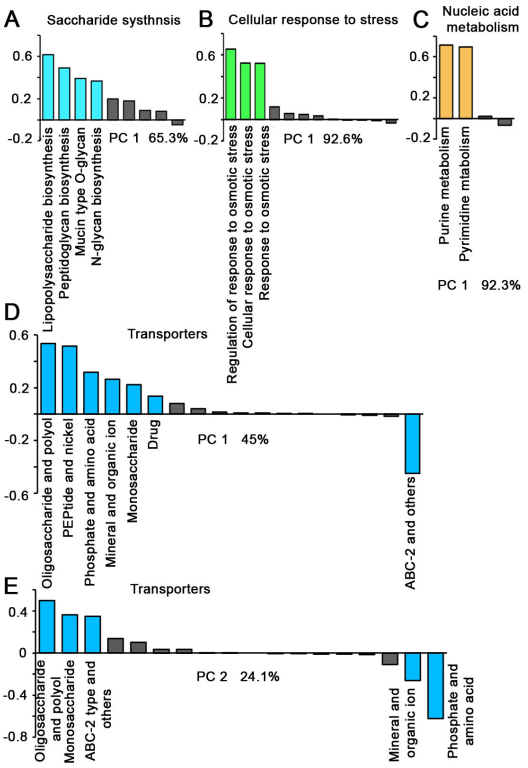

Figure 5.8. Loading plot of principle component analysis (PC1) of functional traits of the rhizosphere samples (FOAM 'level 1') with colored bars as important factors.

5.3.4. Abundance of predicted functional traits in the rhizosphere and in the soil

We visualized the relative abundances of these twelve selected functional traits in the rhizosphere and in the bulk soil of the three dilutions in a heatmap (Fig. 5.10). The functional traits, which were positively correlated with plant biomass of the undiluted 10^{-1} samples, *i.e.* genes related to 'transporters' and 'nucleic acid metabolism', clustered together and were over-represented in the rhizosphere compared to the soil samples (Fig. 5.9). This overrepresentation

was strongest for the 'transporters' genes. Therefore we analyzed this category in more detail. Permanova test yielded significant results for the interaction between dilutions and the presence of plants for the 'transporters' genes (F = 7.97, $P < 0.0001$). Visual inspection of the heatmap clearly showed effects of both dilutions and the presence of plants (Fig. 5.10). The strongest differences that were consistent with the Permanova test were between the soil samples and rhizosphere samples of 10^{-6} dilution. Furthermore, Permanova tests showed that both dilutions and the presence of plants had a significant influence on the relative abundance of 'transporters' genes ('level 1'), respectively (Table 5.3; F = 14.98, $P < 0.001$; F = 15.13, $P < 0.001$). The relative abundance of 'ABC transporters' genes in the rhizosphere, involved in the uptake of monosaccharides, oligosaccharide and other compounds, was significantly affected by dilutions (F = 14.98, $P < 0.001$; F = 10.59, $P < 0.001$; F = 11.72, $P < 0.001$, for the three 'transporters' genes, respectively), and was higher in the rhizosphere than in the soil (F = 15.13, $P < 0.001$; F = 12.87, $P < 0.001$; F = 21.77, $P < 0.001$). Both dilutions and the presence of plants increased the relative abundance of 'drug transporters', respectively (F = 17.97, $P < 0.001$; F = 18.72, $P < 0.001$).

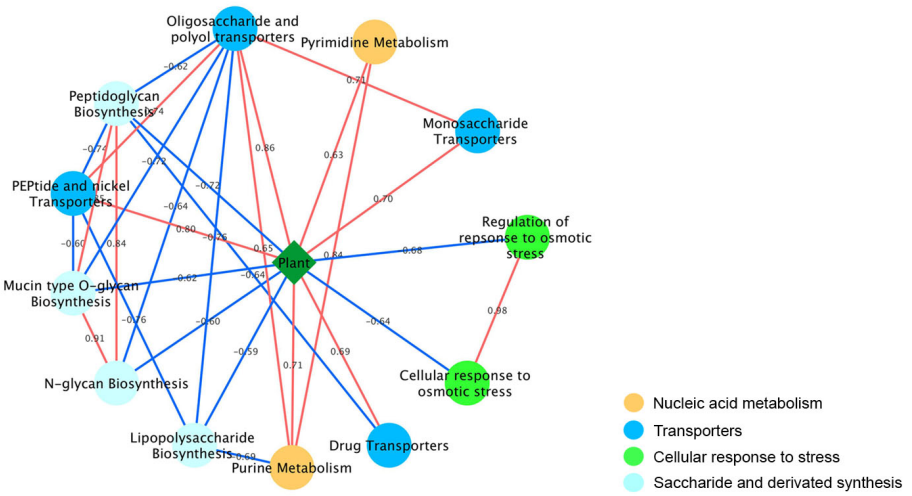

Figure 5.9. Network correlations of twelve functional traits (FOAM 'level 2' and 'level 3') with the plant biomass of the undiluted 10^{-1} samples. Red lines indicate positive correlations, blue lines indicate negative correlations. The colour of each node indicates the functional categories of level 1 to which the traits belong. $P < 0.05$, R > 0.05 or R < -0.05.

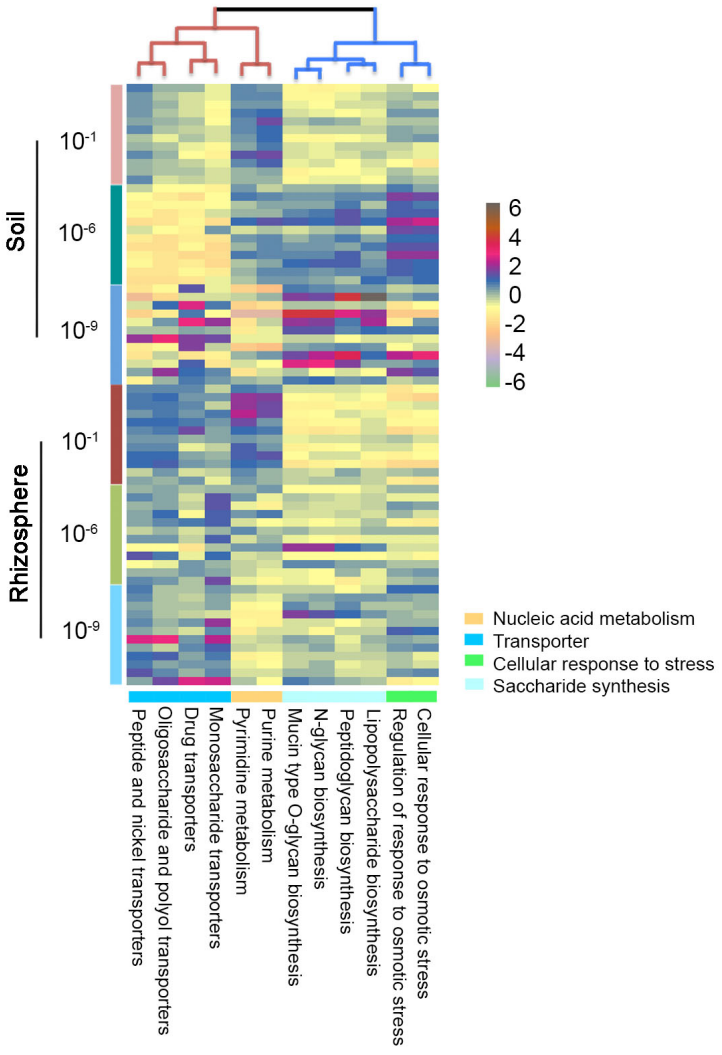

Figure 5.10. Heatmap of relative abundance of twelve functional traits in each dilution of the soil and the rhizosphere. Red lines cluster functional traits positively correlated with plant biomass of undiluted 10^{-1} samples, blue lines cluster functional traits negatively correlated with plant biomass of undiluted 10^{-1} samples. The colour at the bottom of the heatmap profile indicates the functional categories of level 1 to which the traits belong.

Table 5.3. Two-way Permanova using Bray-Curtis similarity showing the effects of dilution and plant presence on functional traits.

Functions			Factors	Sum of sqrs	df	Mean Square	*F*	p
Level 1	Level 2	Level 3						
Transporters			Plant presence	0.019	1	0.019	15.13	< 0.001
			Dilution	0.038	2	0.019	14.98	< 0.001
			Interaction	0.020	2	0.010	7.97	< 0.001
Transporters	ABC transporters	Monosaccharide transporters	Plant presence	0.019	1	0.019	15.13	< 0.001
			Dilution	0.038	2	0.019	14.98	< 0.001
			Interaction	0.020	2	0.010	7.97	< 0.001
Transporters	ABC transporters	Oligosaccharide and polyol transporters	Plant presence	0.002	1	0.002	12.87	< 0.001
			Dilution	0.003	2	0.002	10.59	< 0.001
			Interaction	0.002	2	0.001	5.27	< 0.001
Transporters	ABC transporters	Peptide and nickel transporters	Plant presence	0.002	1	0.002	21.77	< 0.001
			Dilution	0.003	2	0.001	11.72	< 0.001
			Interaction	0.002	2	0.001	8.84	< 0.001
Transporters	Major Facilitator Superfamily	Drug transporters	Plant presence	0.001	1	0.001	17.97	< 0.001
			Dilution	0.002	2	0.001	18.72	< 0.001
			Interaction	0.001	2	0.001	13.28	< 0.001
Saccharides and derivate synthesis			Plant presence	0.001	1	0.001	0.75	0.52
			Dilution	0.005	2	0.002	6.08	< 0.001
			Interaction	0.001	2	0.001	1.23	0.30
Cellular response to stress			Plant presence	0.001	1	0.001	1.022	0.317
			Dilution	0.021	2	0.010	44.84	< 0.001
			Interaction	0.002	2	0.001	4.62	0.012

Functional traits related to 'saccharides and derivate synthesis', which were negatively correlated with plant biomass of the undiluted 10^{-1} samples, were over-represented in the bulk soil compared to the rhizosphere samples (Fig. 5.10). This is also true for the genes of 'cellular response to stress'. Permanova test for the effects of dilutions and the presence of plants on the relative abundance of 'cellular response to stress' related genes resulted in a significant interaction ($F = 4.62$, $P = 0.01$). Furthermore, Permanova test for the effect of dilutions and the presence of plants on the relative abundance of 'saccharides and derivate synthesis' resulted in significant results for dilutions (Table 5.1; $F = 6.08$, $P < 0.001$), but not for the presence of plants. However, dilutions also had a significant influence on the relative abundance of 'cellular response to stress' related genes ($F = 44.84$, $P < 0.001$), but not for the presence of plant.

The over representation in the rhizosphere was the strongest for the 'transporters' genes, the group of genes that also had the highest correlation with plant biomass ($n = 12$, $R = 0.90$, $p < 0.001$). To further analyze which of the four groups of functional genes ('level 1') potentially was the most important one to explain variation in plant biomass we calculated partial correlations with plant biomass (Table 5.4). After taking the 'transporters' genes into account none of the other groups of functional genes was significantly correlated with plant biomass while after taking the other groups into account in each case 'transporters' genes were significantly correlated with plant biomass.

Table 5.4. Partial correlation matrix between 'transporters' and other functional traits controlling plant biomass production.

Functional traits studied	Nucleic acid metabolism	Transporters	Plant biomass 10^{-1}
Nucleic acid metabolism	1.00	0.29 (0.530)	0.15 (0.750)
Transporters		1.00	0.84*(0.020)
Plant biomass 10^{-1}			1.00
	Carbohydrate Active enzyme	Transporters	Plant biomass 10^{-1}
Carbohydrate Active enzyme	1.00	-0.46 (0.300)	0.17 (0.710)
Transporters		1.00	0.90**(0.006)
Plant biomass 10^{-1}			1.00
	Saccharide and derivate synthesis	Transporters	Plant biomass 10^{-1}
Saccharides and derivate synthesis	1.00	-0.82* (0.020)	0.58 (0.170)
Transporters		1.00	0.91**(0.004)
Plant biomass 10^{-1}			1.00
	Cellular response to stress	Transporters	Plant biomass 10^{-1}
Cellular response to stress	1.00	-0.62 (0.140)	0.36 (0.43)
Transporters		1.00	0.91**(0.005)
Plant biomass 10^{-1}			1.00

Values in the table indicate partial correlation coefficients (*P*-value) between two functional traits; *$P < 0.05$, **$P < 0.01$.

5.3.5. A combined analysis of the effects of taxonomical composition and functional traits on plant biomass.

We first analyzed the relative frequency of the twelve selected functional genes for *Arthrobacter* and the remainder of the bacterial community (Fig. 5.11). A heatmap revealed that functional traits belonging to 'transporters', 'nucleic acid

metabolism', and most genes of 'saccharides and derivate synthesis' clustered together. Specifically, genes of 'purine metabolism', 'pyrimidine metabolism', 'drug transporters', 'monosaccharide transporters' and 'peptidoglycan biosynthesis' were significantly enriched in the community with *Arthrobacter*, while 'cellular response to osmotic stress', 'regulation of response to osmotic stress' and 'lipopolysaccharide biosynthesis' were clustered together and were significantly enriched in bacterial community without *Arthrobacter*.

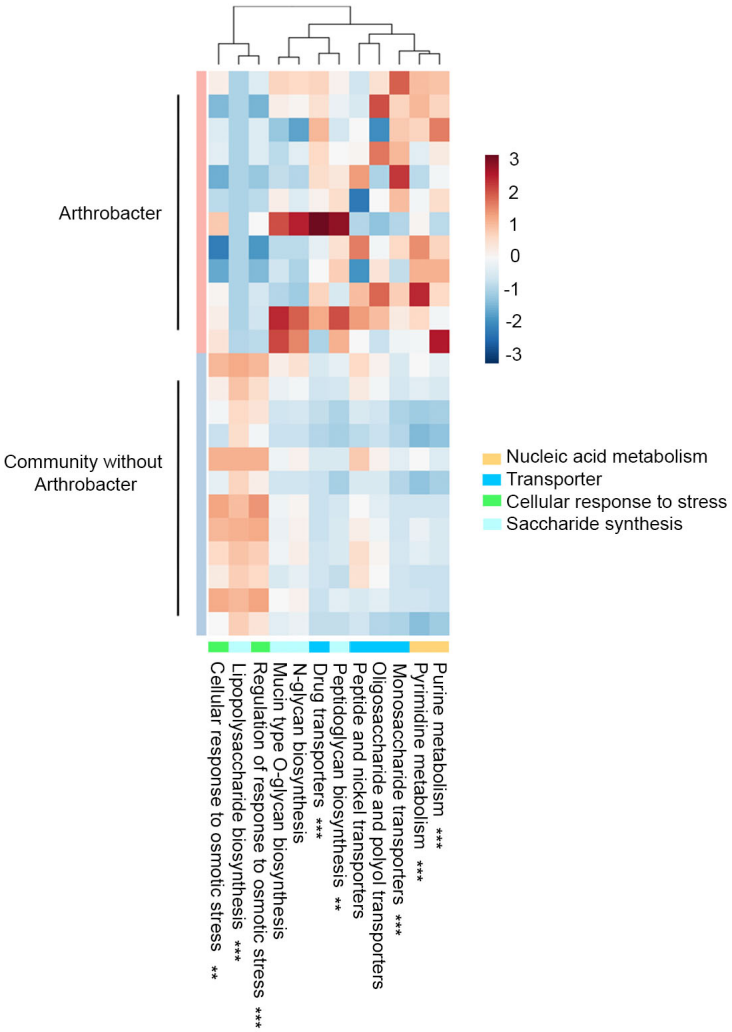

Figure 5.11. Heatmap of relative abundance of the twelve functional traits of the undiluted 10^{-1} rhizosphere samples for communities with and without *Arthrobacter*. The colour at the bottom of heatmap profile indicates the functional traits at 'level 1' (ANOVA: ** $P < 0.01$, *** $P < 0.001$).

Because the frequency of 'transporters' genes is higher in the community with *Arthrobacter*, the correlation of 'transporters' genes with plant biomass may be indirect through *Arthrobacter* or vice versa. We therefore calculated partial correlations. After taking the frequency of 'transporters' genes into account, the correlation of *Arthrobacter* with plant biomass is no longer significant. While after taking the effect of *Arthrobacter* into account, the correlation of 'transporters' genes with plant biomass is still significant (Table 5.5; n = 12, R_p = 0.79, P = 0.036). This suggests the 'transporters' genes are more important to explain differences in plant biomass than *Arthrobacter*.

Table 5.5. Partial correlation matrix between *Arthrobacter* and transporters controlling plant biomass production.

Variables studied	Transporters	Arthrobacter	Plant biomass 10^{-1}
Transporters	1.00	0.05 (0.916)	0.79* (0.036)
Arthrobacter		1.00	0.50 (0.258)
Plant biomass 10^{-1}			1.00

Values in the table indicate partial correlation coefficients (P-value) within Arthrobacter and transporters; *P < 0.05.

5.4. Discussion

In Chapter 2, we demonstrated that the dilution procedure changes the diversity and structure of the bacterial community after regrown in the soil and in Chapter 4 that further selection proceeds in the rhizosphere, largely on the basis of functional traits. This may imply that functional traits that are selected in the rhizosphere may also have a strong influence on plant growth. As we already showed in Chapter 4 that the selection of functional traits is not randomly associated with taxonomic selection, the functional selection as observed in the rhizosphere is, of course, intimately associated with selection of bacterial species. Therefore, here, we also assessed the taxonomic relationship between plant biomass and the bacterial community composition.

Indeed, we demonstrated that manipulation of the bacterial community by the dilution approach, affected plant biomass production. Plants gained the lowest biomass in soils inoculated with the lower dilutions, i.e. the more diverse rhizosphere communities. Recent studies on 'plant-soil feedback' have shown that rhizomicrobiome could directly or indirectly influence the composition and productivity (i.e. biomass) of plant communities (van der Heijen 2008, Joosten et al 2009, van Elsas et al 2012). Moreover, reduction of abundant and/or rare species by manipulation of microbial community could promote plant growth (Hol et al 2010). Hence, microbial community composition belowground has been identified as predictor of fitness of the aboveground vegetation (van der Heijden et al 2008, Lau and Lennon 2011, Wagg et al 2011). In our study, we detected two OTUs (Fig. 5.4), which were actually significantly related to plant biomass and, thus, being potential candidates to explain the observed differences by unsupervised multivariate analysis. The first taxonomical unit was *Arthrobacter* that is known to promote plant growth (Dimkpa et al 2009). *Arthrobacter* is typically found in soil and several species of *Arthrobacter* have been described as plant growth promoter (Gusain et al 2015, Ullah and Bano 2015). Analysis of the wheat rhizosphere using 16S rRNA gene sequencing revealed that *Arthrobacter* belonged to the group of rhizobacteria (Tahir et al 2015). The other group of bacteria that also showed a positive correlation with plant biomass was *Planctomycetaceae.* This genus is known to include typical rhizosphere species (Tesfaye et al 2003), but from literature it is not known if it includes growth-promoting species. Partial correlation analysis identified *Arthrobacter* as the most important one to explain differences in plant biomass (Table 5.1). Identifying genera that promote or inhibit plant growth gives us little information on the mechanisms causing these effects. It is therefore also of great interest to study, in addition to the taxonomic composition, the relationship between plant growth and the functional genes of the bacteria from the rhizosphere community.

This study showed the power of the metagenomics approach in combination with an unsupervised Principal Component Analysis to predict plant biomass production in relation to the functional traits of the rhizomicrobiome. Interestingly, as was shown in the heatmap and Permanova test, the functional traits that were positively correlated with plant biomass were over-represented in the rhizosphere compared to the bulk soil, which suggests that plants selected beneficial bacterial activities surrounding their roots. This

particular bacterial functionality may lead to plant growth promotion. In contrast, the functional traits that were negatively correlated with plant biomass were more abundant in the soil than in the rhizosphere, suggesting plants selected against such functions leading to over-representation in the bulk soil compared to the rhizosphere.

More precisely, a group of 'ABC-type transporters' of peptides, oligosaccharides and drugs and the uptake and release of many different compounds were over-represented in the rhizosphere compared to the bulk soil. This observation is consistent with the fact that numerous genes for 'membrane transporters' systems have been reported as enriched in the rhizosphere (Mendes et al 2014). Another group of functional genes that was over-represented in the rhizosphere compared to the bulk soil is linked to 'nucleic acid metabolism'. Given that the category 'nucleic acid metabolism' involves several interconnected pathways, and may be indicative of cellular growth processes, this suggests higher bacterial growth and activity in the rhizosphere than in soil. This may presumably result in increased plant biomass production for instance by protection against pathogens or by increasing nutrient acquisition for the host.

In this study we observed not only positive but also negative correlations between functional traits of the rhizomicrobiome and plant biomass. As mentioned above, in contrast to the functional genes that were positively related to plant biomass, the ones that were negatively correlated with plant biomass were over-represented in the soil compared to the rhizosphere. This would suggest that plants selected against such genes in the rhizosphere. If, for example, plants create a less stressful environment for bacteria by rhizodeposition this would cause a less stressful environment in the rhizosphere compared to the soil and consequently to an under representation of these genes in the rhizosphere as compared to the soil. At the same time, if plants are growing well, they would produce more roots and more developed rhizosphere and thus to a less stressful environment for the bacteria, leading to a negative correlation between plant biomass and the density of stress genes. We should be careful however with such an interpretation because one of the negative correlations was between plant biomass and a group of osmotic stress genes, i.e. 'cellular response to osmotic stress' and 'regulation of response to osmotic stress'. As larger plants take up more water this could create a more stress full

environment to the microbes and so one could expect a positive relationship between plant biomass and osmotic stress genes Because plants were watered every two days during plant growth we may have created an environment that did not have moisture stress, nor for plants and nor for microbes

We also found a negative relationship between genes related to 'saccharides and derivate synthesis'. It could be that this negative correlation is due to the fact that the plants that grow better provide more carbohydrates and saccharides to the microbes so that biosynthesis of these products is repressed in their rhizosphere. If the above reasoning is correct we would also expect that the rhizosphere provide an environment where these genes are repressed compared to the bulk soil. However, for this group of genes we did not find an under- or over representation in the rhizosphere compared to the bulk soil.

Partial correlation analyses identified 'transporters' genes as the most important ones to potentially explain the observed differences in plant biomass. Likewise we identified *Arthrobacter* as the most important taxonomical unit in this respect. Well-known activities of *Arthrobacter* are degradation of pollutants in the rhizosphere (Khan et al 2009), production of auxin that might stimulate nutrient uptake (Tsavkelova et al 2006) and production of indole-3-acetic acid (IAA) (Sziderics et al 2007). Because the frequency of 'transporters' genes was relatively high in *Arthrobacter* compared to the rest of the bacterial community we used partial correlation to test for the relative importance of the two for plant growth. This analysis suggests that the frequency of 'transporters' genes is the most important factor and that plants select for favorable functions rather than species to benefit their growth. Thus, the high abundance of 'transporters' genes in *Arthrobacter* may lead to the high abundance of *Arthrobacter* in the rhizosphere. Yet, we should treat these results with caution because our analyses may be biased by the fact that the two methods may have different sensitivity. Furthermore, the genes of 'monosaccharide transporters' ('level 3') significantly correlated positively to plant biomass when all three dilutions were taken together, and this group of genes increased significantly upon dilutions in the rhizosphere samples, indicating that this group of genes most likely explained the plant biomass differences in dilutions. Because the 'transporters' genes were enriched in the rhizosphere compared to the soil we can conclude that plants positively affect bacterial species with such genes and this may in part explain the positive correlation with plant growth. At the same

time bacterial species with these 'transporters' genes may stimulate plant growth, which would contribute to the positive correlation. Whether either one of the two or both explanations are true cannot be concluded from our experiments with certainty. Additional experiments are needed that use the microbial communities from soil of pots with different plant growth to inoculate sterile soil again and measure plant growth for a second and following generations in order to achieve maximum enrichment of the most responsible genes. Clearly, it is not very likely that there is a single function that determines the variation in plant biomass. The problem in the analyses of this type of studies is the fact that there are many potential factors and a limited number of replicates. We therefore used an unsupervised method to make an unbiased pre-selection of potential taxonomical units and groups of functional genes. This improves the statistical power but comes at the cost of many important factors going unnoticed.

Nevertheless, our study provides a comprehensive framework for understanding the mechanism of plant-microbe interactions. The latter is not merely of scientific interest but is also useful for the development of sustainable crop production systems, e.g. by application of beneficial soil microbial communities or species to optimize crop yields.

5.5. References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS *et al* (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* **19:** 455-477.

Brockett BFT, Prescott CE, Grayston SJ (2012). Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. *Soil Biology & Biochemistry* **44:** 9-20.

Bulgarelli D, Rott M, Schlaeppi K, van Themaat EVL, Ahmadinejad N, Assenza F *et al* (2012). Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* **488:** 91-95.

Bulgarelli D, Garrido-Oter R, Munch PC, Weiman A, Droge J, Pan Y *et al* (2015). Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley. *Cell Host & Microbe* **17:** 392-403.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Meth* **7:** 335-336.

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al* (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* **37:** D141-D145.

Dennis PG, Miller AJ, Hirsch PR (2010). Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *Fems Microbiology Ecology* **72:** 313-327.

Dimkpa C, Weinand T, Asch F (2009). Plant-rhizobacteria interactions alleviate abiotic stress conditions. *Plant Cell Environ* **32:** 1682-1694.

Dixon P (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14:** 927-930.

Edgar RC (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26:** 2460-2461.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27:** 2194-2200.

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

Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F *et al* (2015). HMMER web server: 2015 update. *Nucleic Acids Res* **43:** W30-38.

Gusain YS, Kamal R, Mehta CM, Singh US, Sharma AK (2015). Phosphate solubilizing and indole-3-acetic acid producing bacteria from the soil of Garhwal Himalaya aimed to improve the growth of rice. *Journal of Environmental Biology* **36:** 301-307.

Haichar FE, Marol C, Berge O, Rangel-Castro JI, Prosser JI, Balesdent J *et al* (2008). Plant host habitat and root exudates shape soil bacterial community structure. *ISME J* **2:** 1221-1230.

Hammer et al HD, Ryan PD. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4: 9. pp.**

Hol WHG, de Boer W, Termorshuizen AJ, Meyer KM, Schneider JHM, van Dam NM *et al* (2010). Reduction of rare soil microbes modifies plant-herbivore interactions. *Ecology Letters* **13:** 292-301.

Hughes JB, Hellmann JJ (2005). The application of rarefaction techniques to molecular inventories of microbial diversity. *Environmental Microbiology* **397:** 292-308.

Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *Bmc Bioinformatics* **11**.

Imelfort M, Woodcroft, B., Parks, D. (2015). Ecogenomics/BamM [WWW Document]. GitHub. URL https://github.com/Ecogenomics/BamM (accessed 7.21.15).

Joosten L, Mulder PPJ, Klinkhamer PGL, van Veen JA (2009). Soil-borne microorganisms and soil-type affect pyrrolizidine alkaloids in Jacobaea vulgaris. *Plant and Soil* **325:** 133-143.

Joshi NA, Fass JN (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. Available at https://github.com/najoshi/sickle.

Khan MS, Zaidi A, Wani PA, Oves M (2009). Role of plant growth promoting rhizobacteria in the remediation of metal contaminated soils. *Environmental Chemistry Letters* **7:** 1-19.

Koster J, Rahmann S (2012). Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* **28:** 2520-2522.

Lau JA, Lennon JT (2011). Evolutionary ecology of plant-microbe interactions: soil microbial structure alters selection on plant traits. *New Phytologist* **192:** 215-224.

Li H, Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25:** 1754-1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:** 2078-2079.

Liao Y, Smyth GK, Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30:** 923-930.

Logue JB, Lindstrom ES (2010). Species sorting affects bacterioplankton community composition as determined by 16S rDNA and 16S rRNA fingerprints. *ISME J* **4:** 729-738.

Mark GL, Dow JM, Kiely PD, Higgins H, Haynes J, Baysse C *et al* (2005). Transcriptome profiling of bacterial responses to root exudates identifies genes involved in microbe-plant interactions. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 17454-17459.

Matilla MA, Espinosa-Urgel M, Rodriguez-Herva JJ, Ramos JL, Ramos-Gonzalez MI (2007). Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. *Genome Biology* **8**.

McKinney W (2015). Python Data Analysis Library — pandas: Python Data Analysis Library [WWW Document]. URL http://pandas.pydata.org/ (accessed 7.22.15).

Mendes LW, Kuramae EE, Navarrete AA, van Veen JA, Tsai SM (2014). Taxonomical and functional microbial community selection in soybean rhizosphere. *ISME J* **8:** 1577-1587.

Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM *et al* (2011). Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* **332:** 1097-1100.

Nemergut DR, Cleveland CC, Wieder WR, Washenberger CL, Townsend AR (2010). Plot-scale manipulations of organic matter inputs to soils correlate with shifts in microbial community composition in a lowland tropical rain forest. *Soil Biology & Biochemistry* **42:** 2153-2160.

Ofek-Lalzar M, Sela N, Goldman-Voronov M, Green SJ, Hadar Y, Minz D (2014). Niche and host-associated functional signatures of the root surface microbiome. *Nature Communications* **5**.

Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH (2013). Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews Microbiology* **11:** 789-799.

Prestat E, David MM, Hultman J, Tas N, Lamendella R, Dvornik J *et al* (2014). FOAM (Functional Ontology Assignments for Metagenomes): a Hidden Markov Model (HMM) database with environmental focus. *Nucleic Acids Res* **42:** e145.

Reeder J, Knight R (2010). Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nature Methods* **7:** 668-669.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D *et al* (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research* **13:** 2498-2504.

Steiner A (1968). Soilless culture. Proceedings of the 6th colloquium of the International Potash Institute, Florence,Italy. International Potash Institute, Bern, Switzerland, pp 324-341.

Sziderics AH, Rasche F, Trognitz F, Sessitsch A, Wilhelm E (2007). Bacterial endophytes contribute to abiotic stress adaptation in pepper plants (Capsicum annuum L.). *Canadian Journal of Microbiology* **53:** 1195-1202.

Tahir M, Mirza MS, Hameed S, Dimitrov MR, Smidt H (2015). Cultivation-Based and Molecular Assessment of Bacterial Diversity in the Rhizosheath of Wheat under Different Crop Rotations. *Plos One* **10**.

Tesfaye M, Dufault NS, Dornbusch MR, Allan DL, Vance CP, Samac DA (2003). Influence of enhanced malate dehydrogenase expression by alfalfa on diversity of rhizobacteria and soil nutrient availability. *Soil Biology & Biochemistry* **35:** 1103-1113.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ *et al* (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* **28:** 511-U174.

Tremaroli V, Backhed F (2012). Functional interactions between the gut microbiota and host metabolism. *Nature* **489:** 242-249.

Tsavkelova EA, Klimova SY, Cherdyntseva TA, Netrusov AI (2006). Microbial producers of plant growth Stimulators and their practical use: A review. *Applied Biochemistry and Microbiology* **42:** 117-126.

Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444:** 1027-1031.

Ullah S, Bano A (2015). Isolation of plant-growth-promoting rhizobacteria from rhizospheric soil of halophytes and their impact on maize (Zea mays L.) under induced soil salinity. *Canadian Journal of Microbiology* **61:** 307-313.

van der Heijden MGA, Bardgett RD, van Straalen NM (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters* **11:** 296-310.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T *et al* (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396:** 69-72.

van der Heijen MGA (2008). The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems (vol 11, pg 296, 2008). *Ecology Letters* **11:** 651-651.

van Elsas JD, Chiurazzi M, Mallon CA, Elhottova D, Kristufek V, Salles JF (2012). Microbial diversity determines the invasion of soil by a bacterial pathogen. *Proceedings of the National Academy of Sciences of the United States of America* **109:** 1159-1164.

Wagg C, Jansa J, Schmid B, van der Heijden MGA (2011). Belowground biodiversity effects of plant symbionts support aboveground productivity. *Ecology Letters* **14:** 1001-1009.

Chapter 6

General discussion

The main goal of the research described in this thesis was to obtain a better understanding of the assemblage and diversity of bacterial communities in soil and rhizosphere as well as of their functionality. To reach this, I manipulated the community diversity by use of the so-called dilution approach focusing on both the effects of soil and plant on the community assemblage. In order to study the microbial community diversity and functionality, I applied a combined approach of next generation amplicon and shotgun metagenome sequencing followed by advanced bioinformatics and statistical analyses. Here, I will first discuss the methodology of studying microbial diversity in soil, which could be used as a general approach in further studies to analyze microbial diversity experimentally. Secondly, I will discuss the importance of the impact of soil and the relevant physicochemical soil characteristics on the structuring of microbial communities in soil. Thirdly, I will concentrate on the microbial community assemblage processes operating in soil and rhizosphere at both taxonomic and functional levels. In the fourth section regarding plant-microbe interactions, I will discuss the feedback of soil-borne bacteria and plants, and the functional traits that determinate the relationship between the bacterial community and plant growth. Finally, I will discuss ideas and directions for future research on soil microbial diversity.

6.1 Methodology: assessing the microbial community diversity in terrestrial ecosystems

Previously, many studies on the creation and functionality of biodiversity have focused on macro-organisms and much less on microorganisms (Bever 1994, Shanmugam et al 2011, Tilman et al 1997) despite the increasingly recognized importance of microbial diversity in terrestrial and other natural ecosystems (Thiele-Bruhn et al 2012, Wall et al 2015). One of the major hurdles in these studies is the lack of sound approaches to manipulate experimentally microbial biodiversity. One of the main approaches applied to microbial biodiversity and assemblage studies is the so-called dilution approach, which is used here. Until now, the studies performed to assess microbial biodiversity based on this approach have often been restricted by low-resolution based analytical methodologies (Griffiths et al 2001, Mandeel et al 2005, Nielsen et al 2015, Prakamhang et al 2015, Wall and Six 2015). So, they failed to comprise the total microbial community profiles, while there are sufficient arguments that it is of utmost importance to study the diversity and functionality of the total microbiome in terrestrial ecosystems (Berendsen et al 2012, Chaparro et al 2014).

The estimates of bacterial diversity based on the results obtained with the dilution approach revealed that the bacterial community diversity was reduced significantly at species or OTU level by dilution of a soil suspension (Chapters 2 and 3). Previous studies claimed that by dilution particularly rare species were removed from soil suspensions and that therefore the abundant ones would dominate the microbial community formed after incubation of the diluted suspensions in soil (Franklin and Mills 2006, Garland and Lehman 1999). As the role of rare species in ecosystem functioning is a hot topic in ecology (Gaston 2012, Pedros-Alio 2012), I was especially interested in the possibilities provided by the dilution approach to, indeed, separate abundant and rare species. The results of my studies, however, showed that unique species were present in all dilutions including the most diluted suspensions. Probably certain species are suppressed for the sequencing assessments in the less diluted suspensions and only showed up in the more diluted, less diverse, suspensions. Thus, the conclusion can be drawn that the common presumption underlying the dilution approach that rare species would be out diluted, is not correct. Thus, the dilution approach does not allow for the separations of rare and

abundant species, and, so, is not the appropriate approach to study the importance of rare and/or abundant microorganisms in ecosystems.

It should be pointed out however, that it is hard to formulate a clear-cut definition of 'abundant' and 'rare' organisms (Fuhrman 2009). Rare species are often described as organisms occurring in the relative abundance range of approximately 0.1% (Postma-Blaauw et al 2005) to 0.01% (Qin et al 2010) of the total community. However, organisms that do occur in one environment in that abundance range may become common, even dominant, when the environment changes (Kulmatiski et al 2008). So, they may be regarded as a seed bank of diversity and functionality when local conditions change through natural or anthropogenic causes (Fuhrman 2009). Yet, certain studies indicated that it would be the abundant species that mainly perform most of the functions in marine ecosystems (Cottrell and Kirchman 2003). Similarly, studies based on advanced sequencing approaches indicated that the abundant members of the community are primarily responsible for most major biogeochemical processes such as the nutrient cycling (Pedros-Alio 2006). So, in conclusion: our understanding of the functional importance of different groups, *i.e.* rare/abundant species, in natural ecosystems is limited mainly by limitations to the possibilities provided by the currently available methodological approaches to provide a comprehensive assessment of the microbial community diversity at the phylum (van de Voorde et al 2012) and/or the OTU level (Bulgarelli et al 2015).

Recently, new low-cost, high-throughput sequencing approaches have greatly improved the understanding of the huge diversity of microbial communities in ecosystems (Bulgarelli et al 2015, Franzosa et al 2015, Lebeis et al 2015, Rodrigues et al 2013, van de Voorde et al 2011, van de Voorde et al 2012). High-resolution sequencing approaches have the potential to allow for the detection of the entire microbial community structure including the most dominant and the rarest species (Lynch and Neufeld 2015, Pester et al 2010). I applied these approaches in the study described here. Continuing advances in sequencing technology have allowed for studies on diverse microbiomes, ranging from natural environments (Mendes et al 2014) to the human body (Tremaroli and Backhed 2012). Although these approaches have been proven highly effective, there are still limitations based on the current DNA sequence-based methods. For example, upon application of these approaches, a clear

definition of a microbial species is still lacking (Nielsen et al 2015). Usually we rely on the sequence similarities of taxa-specific DNA subunits to distinguish microorganisms, using the term "Operational Taxonomic Unit, or OTU" rather than species. Most sequencing approaches provide, at best, species-level taxonomic resolution, but many important phenomena may be present at the strain level. Furthermore, uncultured microorganisms represent the majority of microbial diversity, and, as their functional potential is largely unknown, the databases used for annotation of the functional genes underrepresent dramatically the overall microbial functional potential. In addition, a fundamental limitation of metagenome sequencing is that the presence of a functional gene does not necessarily represent its activity, as host organisms may be dormant, inactive or only active in certain condition. Thus, additional integrated approaches, such as RNA (transcriptomics) and proteins (proteomics), are required to fully describe a microbial community and it's functioning.

The network analysis described in Chapter 4 showed a more tighter and complex network of rhizosphere communities than that of bulk soil communities, including more keynote species mainly belonging to different genera. These key members of the rhizosphere microbial communities may also be the key intermediaries in plant-microbe associations. Indeed, in Chapter 5, two groups, *i.e. Arthrobacter* and *Planctomycetaceae*, which were identified as strongly enriched families in the rhizosphere and important intermediates in the networks in the rhizosphere (Chapter 4), were also identified as potential candidates to explain best the differences in plant biomass production after incubation of the undiluted 10^{-1} suspension using unsupervised multivariate analysis. Further partial correlation revealed that *Arthrobacter* was the taxonomical group most related to plant growth. However, *Arthrobacter* had a lower betweenness centrality, *i.e.* the extent of network interactions, than *Planctomycetaceae* in the 10^{-1} rhizosphere community (317 for *Planctomycetaceae* and 163 for *Arthrobacter*, respectively). This suggests that *Arthrobacter* might have been more important for plant-microbe interactions, while *Planctomycetaceae* mediated more network associations. The role of the other key intermediate groups of the rhizosphere and soil networks was not further assessed and, at least, their impact on plant biomass production was negligible as compared to that of *Arthrobacter* and *Planctomycetaceae.* This

points to both the power and the limitations of network analyses for detecting species associations in plant-soil systems.

6.2 The impact of soil on the structuring of soil bacterial communities

Previous studies have found that soil is one of the most important factors structuring microbial communities (Berg and Smalla 2009, Garbeva et al 2004, Kuramae et al 2012). Results from my study indicated (Chapters 2 and 3), indeed, that different soils had a strong steering, selective, effect on shaping bacterial communities.

As described in previous studies, soil type has been ranked as the most important factor determining the structure of microbial communities, followed by time, specific farming operations, management systems and spatial variation (Bossio et al 1998). The factors in soils that may potentially affect microbial communities and thus may explain differences and shifts in community structure are pH (Lauber et al 2009), phosphate availability (Faoro et al 2010), and organic matter content (Verbruggen et al 2010). The soils I used in the cross-dilution experiment differed in these factors. The Utrecht soil was characterized by low pH, the Clue soil was characterized by high phosphate content, and high organic matter while the Meijendel was characterized by a relatively high pH. All these soils contained a characteristic microbial community and the factors mentioned are likely the driving variables shaping the bacterial communities in these soils (Chapter 3).

Previously, studies have focused on the importance of single (a)biotic factors and much less on the integrated soil characteristics when evaluating the effects of soil on microbial community structure and function (Murty et al 2002, Torsvik and Ovreas 2002), despite the increasing recognition of the importance of the overall environment on the structuring of microbial communities in soil and their biodiversity (Fierer and Jackson 2006, Hogberg et al 2007). I showed that the structure of the bacterial community was changed dramatically after incubation in soil as compared to the structure of the community in the suspension (Chapters 2 and 3). This strongly points to the overriding impact of soil, as an important, decisive, factor in the assemblage of bacterial communities in soil, which is likely due to the integrated physical and chemical

characteristics of the soils. Therefore, I suggest that a combination of abiotic factors, and not only pH or another single factor determines the structure of soil bacterial communities.

6.3 Bacterial community assemblage in soil and rhizosphere

Plants are known to significantly select for specific microorganisms in the rhizosphere (Haichar et al 2008, Mendes et al 2014). This is called the 'rhizosphere effect'. It is known that plant species have rather specific effects on the structure of the rhizosphere microbial communities, even at the genotype level, (Berg and Smalla 2009, Duineveld et al 2001, Haichar et al 2008). I also observed a considerable effect of the presence of plants on the bacterial communities in the rhizosphere at both taxonomic and functional levels (Chapter 4).

Earlier studies indicated that plants influence the composition and activity of the rhizosphere microbiota by selecting specific microbial populations from the soil-borne microbial reservoirs (Berg 2009, van Overbeek and van Elsas 2008), and, thus, the microbial community in the rhizosphere is a subset of the bulk soil (Duineveld et al 2001). Results presented in this thesis clearly indicate that the composition of the rhizosphere communities was dramatically different from that of the soil communities in terms of the dominant species. That does not hold for the abundant phyla of *Proteobacteria*, which showed to be highly diverse both in soil (Chapters 2 and 3) and rhizosphere (Chapter 4), which is consistent with the common concepts on the lifestyle of *Proteobacteria* (Fierer et al 2007). In agreement with earlier observations, within the phylum of the *Preoteobacteria*, bacteria from the families of *Pseudomonadaceae* (DeAngelis et al 2009) or of *Burkholderiaceae* (Pastorelli et al 2011, Uroz et al 2010) are among the most abundant members of the rhizosphere communities. It is, therefore, remarkable that 'transporters' genes that could be assigned to *Pseudomonaceae* were overrepresented in the bulk soil and not in the rhizosphere. As the occurrence of 'transporters' genes was found to be a determinative factor explaining plant biomass production (Chapter 5), this questions the significance of this group of bacteria as plant growth promoting organisms. Also, the relative abundance of *Actinobacteria* was found to be significantly larger in the rhizosphere of *Senecio* plant than in

the bulk soil while the Shannon diversity index for this phylum was significantly lower in the rhizosphere, which could be explained by the large relative abundance in the rhizosphere samples of one family, *i.e. Micrococcaceae* to which *Arthrobacter* belongs (Chapter 4). Interestingly, in line with these observations we clearly showed in Chapter 5 that *Arthrobacter* was significantly positively correlated to plant biomass more than any other group of bacteria (Chapter 5).

Based on the possibilities provided by the advanced sequencing approaches available, the concept of 'rhizosphere effect' should not only be limited to species but should be extended to the selection of functional genes in the soil microbiome (Mendes et al 2014, Ofek-Lalzar et al 2014). One of main goals of metagenomics has always been to link functional genes to particular organisms (DeLong 2009). The results presented in chapter 4 clearly illustrate the process of rhizosphere selection both at the community composition and functioning levels. We showed that the enrichment processes in the rhizosphere selects for microorganisms with specific functional traits including 'transporters', 'Embden Meyerhof Parnas' (EMP) and 'hydrogen metabolism'. The genes related to 'transporters' have been described in earlier observations by Mark et al (2006) and Mendes et al (2014) who showed that transporter systems are frequently enriched in the rhizosphere. The 'transporters' genes were positively related to plant growth. Because they were enriched in the rhizosphere compared to the soil we can conclude that plants positively affect bacterial species with such genes and this may in part explain the positive correlation with plant growth. At the same time bacterial species with these 'transporters' genes may stimulate plant growth, which would contribute to the positive correlation. Whether either one of the two or both explanations are true cannot be concluded from our experiments with certainty. Another over-represented group of functions in the rhizosphere is linked to EMP cycling. The EMP pathway is the most common bacterial glycolytic pathway for cellular energy production (Flamholz et al 2013). Considering that plants provide a wider and more complex range of substrate in the rhizosphere than is available in the soil, and thus provide better conditions for bacterial growth and activity we could expect, indeed, that the genes related to energy production will be over-represented in the rhizosphere metagenome as compared to the soil metagenome. Similarly, 'hydrogen metabolisms' also involve genes related to energy-generating mechanisms of specific microbial species such as nitrogen-

fixing bacteria (Eisbrenner and Evans 1983). Therefore, the group of genes related to 'hydrogen metabolism' might also be over-represented in the rhizosphere than in the soil, as we discussed earlier.

6.4 Impact of the rhizosphere microbiome on plant growth

One of the main results of the metagenomics analysis was the identification of particular functional genes activated in the rhizosphere, which determines plant microbe interactions. As mentioned above the results described in Chapter 4 demonstrated that selection of functions took place in the rhizosphere resulting in over-representation of particular functional genes in the rhizosphere compared to bulk soil. Although earlier studies have identified particular functions beneficial to plant growth, including nitrogen fixation or disease suppression (Quecine et al 2012, Tittabutr et al 2013), generally, the microbial functional traits that contribute to plant fitness have been largely unknown. In Chapter 5 I identified both the species and functional genes that potentially had most influence on plant growth by unsupervised multivariate analysis. As mentioned earlier based on unsupervised multivariate analysis, *Arthrobacter* and *Planctomycetaceae* were selected as potential candidates to explain the differences in plant biomass production, with *Arthrobacter* having the strongest impact. Tahir et al (2015) showed, after analysis of the wheat rhizosphere using 16S rRNA gene sequencing, that *Arthrobacter* belonged to the plant health promoting rhizobacteria. Several species of *Arthrobacter* have been described as plant growth promoting rhizobacterium (Gusain et al 2015, Ullah and Bano 2015). *Planctomycetales* is also a rhizosphere species (Tahir et al 2015), but its functionality is until now largely unknown.

The results presented in Chapter 5 also illustrated the importance of particular functional gene category for regulating plant growth. Interestingly, the functional genes of 'transporters' in the rhizosphere, which I already observed as being positively selected in the rhizosphere in Chapter 4, also appeared to be positively correlated to plant growth (Chapter 5). This provides evidence that plants may select for particular functional genes that promote their own growth. Interestingly, the frequency of 'transporters' genes was higher in *Arthrobacter* than in most other components of the bacterial community. By using partial correlation analysis, I proved that *Arthrobacter*

was not significantly correlated to plant biomass when taking 'transporters' genes into account, which suggests, that the functional genes explained better the plant-bacteria interactions than the community composition. Specifically, the 'monosaccharide transporters' genes were significantly positively correlated to plant biomass when all three dilutions samples were taken together, and this group of genes increased significantly upon dilutions in the rhizosphere. So, it is not enough to know who is there, but more importantly is to know what are they doing (Xu et al 2014).

However, we should exercise caution with the assertion on the importance of certain functional traits because it is extremely unlikely that a single function determines the differences in plant biomass production. Indeed, we also observed that 'nucleic acid metabolism' genes were also positively correlated to plant biomass production. The nature of this particular relationship is still unclear to us, but this may be related to cellular growth processes, which indicates a higher bacterial abundance/activity in the rhizosphere than in the soil, as also shown in Chapter 5. Consequently, this group of genes may point to a positive relationship between bacterial and plant growth.

We also observed functional genes including 'cellular response to stress' and 'saccharide metabolisms' that were negatively correlated to plant biomass production. As was described above, these functional genes were under-represented in the rhizosphere as compared to their abundance in the soil, suggesting that plant selected against such genes in the rhizosphere. One of the explanations could be that plants create a less hostile environment for microbial community in the rhizosphere by the rhizodeposition processes. As a result, this may lead to a negative correlation between 'cellular response to stress' genes and plant growth. Similarly, if plants produce more saccharides that become available for the rhizosphere community, microbial genes related to their biosynthetic pathway might be suppressed in the rhizosphere. However, in that case, one would expect that the 'saccharide metabolisms' genes were over-represented in the bulk soil, which they were not.

6.5 Final conclusions and future perspectives

A few points to consider in future studies concern the theoretical concept on assemblage of microbial communities, the separation of rare versus abundant species, the plant soil feedback effects and the selection processes operating in the rhizosphere.

Although I did not include this in this thesis, I did assess the rules leading to the assemblage of microbial communities in soil and rhizosphere using the theories niche-based and neutral mechanisms. These theories are based upon macro-ecological concepts, but have been used frequently in microbial ecology to describe microbial community assembly processes. The niche-based assemblage concept predicts that the assemblage of a community is based on niche partitioning of the limited resources between competitive species or the differentiation of niche space within a community and have been used to explain microbial community assemblage processes in, for instance, lakes (Van der Gucht et al 2007), soil (Fierer and Jackson 2006, Lozupone and Knight 2007), rhizosphere (Mendes et al 2014) and human gut (Lu et al 2014). The neutral theory is based on the assumption that the differences between members of an ecological community of similar species or species from the same trophic level are "neutral," or irrelevant to their success. In the light of my results on the importance of functionality rather than of taxonomic composition for the functioning of microbial communities in terrestrial ecosystems, it is recommendable to extend these concepts focusing on species functionality in order to be able to better understand how microbial communities are shaped in soil and rhizosphere. It should be taken into account that these models for microbial communities are mostly applied to the entire microbial community. This can lead to a strong underestimation of the selection effects. Selection is most likely occurring within groups of organisms, such as pollinators, insect herbivores etc. that share important ecological features. However, scale is a significant problem in microbial ecology. The entire microbial community encompasses a very diverse set of such ecological features and by pooling all microbes into one group selection may appear neutral while in fact it is not. Finding the “pollinators” within microbial communities is one of the challenges for modern microbial ecology.

The existing assumption linked to the dilution approach is that the approach allows for the separation of rare from abundant species. However, I

showed that unique species were detected even in the most diluted suspension and that rare/less abundant species could become abundant in another environment. This may be effectuated by the dilution procedure in which a less competitive environment may be created for rare species so to flourish more than in a more competitive environment of a less diluted inoculum. Thus, it is impossible to investigate their importance in natural ecosystems by use of the dilution approach. As the long tail of less abundant/rare species is a typical characteristic of highly diverse natural microbial communities, there is still an urgent need to develop appropriate methodologies to separate less and more abundant microbes that allow for specific investigations of their behavior and activities.

Future studies also need to compare plant-soil feedback processes across ecosystems and across successional stages within these ecosystems. One of the most reported findings regarding plant-soil feedback effects is that these effects are negative regarding plant growth (Bever 2003, Lankau et al 2011). So, these issues need to be further consideration under controlled conditions and time scales in order to enable the determination of the potential factors explaining the feedback processes. In these studies functional traits rather than taxonomy should be the target of fundamental research.

Finally, further studies should address how microbial communities are structured and selected at both the taxonomic and functional levels, in distinct soil types and in the presence of specific plant species. Metagenomics analysis has provided information about which microorganisms are present and what they are capable of doing. However, the detection of functional genes is not evidence of their activity. Further functional gene expression analysis including metatranscriptome analysis can provide information about what microorganisms are actually doing. Therefore, integrated experimental, including sequencing approaches, together with computational analysis, are needed to improve our understanding of microbial functionality in specific niche and plant- microbe interactions.

6.6. References

Berendsen RL, Pieterse CMJ, Bakker PAHM (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci* **17:** 478-486.

Berg G (2009). Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. *Appl Microbiol Biot* **84:** 11-18.

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

Bever JD (2003). Soil community feedback and the coexistence of competitors: conceptual frameworks and empirical tests. *New Phytol* **157:** 465-473.

Bossio DA, Scow KM, Gunapala N, Graham KJ (1998). Determinants of soil microbial communities: Effects of agricultural management, season, and soil type on phospholipid fatty acid profiles. *Microbial Ecol* **36:** 1-12.

Bulgarelli D, Garrido-Oter R, Munch PC, Weiman A, Droge J, Pan Y *et al* (2015). Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley. *Cell Host Microbe* **17:** 392-403.

Chaparro JM, Badri DV, Vivanco JM (2014). Rhizosphere microbiome assemblage is affected by plant development. *ISME J* **8:** 790-803.

Cottrell MT, Kirchman DL (2003). Contribution of major bacterial groups to bacterial biomass production (thymidine and leucine incorporation) in the Delaware estuary. Limnol. *Oceanogr*. 48, 168-178.

DeAngelis KM, Brodie EL, DeSantis TZ, Andersen GL, Lindow SE, Firestone MK (2009). Selective progressive response of soil microbial community to wild oat roots. *ISME J* **3:** 168-178.

DeLong EF (2009). The microbial ocean from genomes to biomes. *Nature* **459:** 200-206.

Duineveld BM, Kowalchuk GA, Keijzer A, van Elsas JD, van Veen JA (2001). Analysis of bacterial communities in the rhizosphere of chrysanthemum via denaturing gradient gel electrophoresis of PCR-amplified 16S rRNA as well as DNA fragments coding for 16S rRNA. *Appl Environ Microb* **67:** 172-178.

Eisbrenner G, Evans HJ (1983). Aspects of Hydrogen Metabolism in Nitrogen-Fixing Legumes and Other Plant-Microbe Associations. *Annu Rev Plant Phys* **34:** 105-136.

Faoro H, Alves AC, Souza EM, Rigo LU, Cruz LM, Al-Janabi SM *et al* (2010). Influence of Soil Characteristics on the Diversity of Bacteria in the Southern Brazilian Atlantic Forest. *Appl Environ Microb* **76:** 4744-4749.

Fierer N, Jackson RB (2006). The diversity and biogeography of soil bacterial communities. *P Natl Acad Sci USA* **103:** 626-631.

Fierer N, Bradford MA, Jackson RB (2007). Toward an ecological classification of soil bacteria. *Ecology* **88:** 1354-1364.

Flamholz A, Noor E, Bar-Even A, Liebermeister W, Milo R (2013). Glycolytic strategy as a tradeoff between energy yield and protein cost. *P Natl Acad Sci USA* **110:** 10039-10044.

Franklin RB, Mills AL (2006). Structural and functional responses of a sewage microbial community to dilution-induced reductions in diversity. *Microb Ecol* **52:** 280-288.

Franzosa EA, Hsu T, Sirota-Madi A, Shafquat A, Abu-Ali G, Morgan XC *et al* (2015). Sequencing and beyond: integrating molecular 'omics' for microbial community profiling. *Nature Reviews Microbiology* **13:** 360-372.

Fuhrman JA (2009). Microbial community structure and its functional implications. *Nature* **459:** 193-199.

Garbeva P, van Veen JA, van Elsas JD (2004). Microbial diversity in soil: Selection of microbial populations by plant and soil type and implications for disease suppressiveness. *Annu Rev Phytopathol* **42:** 243-270.

Garland JL, Lehman RM (1999). Dilution/extinction of community phenotypic characters to estimate relative structural diversity in mixed communities. *Fems Microbiol Ecol* **30:** 333-343.

Gaston KJ (2012). Ecology the Importance of Being Rare. *Nature* **487:** 46-47.

Griffiths BS, Ritz K, Wheatley R, Kuan HL, Boag B, Christensen S *et al* (2001). An examination of the biodiversity-ecosystem function relationship in arable soil microbial communities. *Soil Biol Biochem* **33:** 1713-1722.

Gusain YS, Kamal R, Mehta CM, Singh US, Sharma AK (2015). Phosphate solubilizing and indole-3-acetic acid producing bacteria from the soil of Garhwal Himalaya aimed to improve the growth of rice. *J Environ Biol* **36:** 301-307.

Haichar FE, Marol C, Berge O, Rangel-Castro JI, Prosser JI, Balesdent J *et al* (2008). Plant host habitat and root exudates shape soil bacterial community structure. *ISME J* **2:** 1221-1230.

Hogberg MN, Hogberg P, Myrold DD (2007). Is microbial community composition in boreal forest soils determined by pH, C-to-N ratio, the trees, or all three? *Oecologia* **150:** 590-601.

Kulmatiski A, Beard KH, Stevens JR, Cobbold SM (2008). Plant-soil feedbacks: a meta-analytical review. *Ecol Lett* **11:** 980-992.

Kuramae EE, Yergeau E, Wong LC, Pijl AS, van Veen JA, Kowalchuk GA (2012). Soil characteristics more strongly influence soil bacterial communities than land-use type. *Fems Microbiol Ecol* **79:** 12-24.

Lankau RA, Wheeler E, Bennett AE, Strauss SY (2011). Plant-soil feedbacks contribute to an intransitive competitive network that promotes both genetic and species diversity. *J Ecol* **99:** 176-185.

Lauber CL, Hamady M, Knight R, Fierer N (2009). Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil Bacterial Community Structure at the Continental Scale. *Appl Environ Microb* **75:** 5111-5120.

Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald M *et al* (2015). Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* **349:** 860-864.

Lozupone CA, Knight R (2007). Global patterns in bacterial diversity. *P Natl Acad Sci USA* **104:** 11436-11440.

Lu HP, Lai YC, Huang SW, Chen HC, Hsieh CH, Yu HT (2014). Spatial heterogeneity of gut microbiota reveals multiple bacterial communities with distinct characteristics. *Scientific Reports* **4**.

Lynch MDJ, Neufeld JD (2015). Ecology and exploration of the rare biosphere. *Nat Rev Microbiol* **13:** 217-229.

Mandeel Q, Ayub N, Gul J (2005). Survey of Fusarium species in an and environment of Bahrain. VI. Biodiversity of the genus Fusarium in root-soil ecosystem of halophytic date palm (Phoenix dactylifera) community. *Cryptogamie Mycol* **26:** 365-404.

Mendes LW, Kuramae EE, Navarrete AA, van Veen JA, Tsai SM (2014). Taxonomical and functional microbial community selection in soybean rhizosphere. *Isme J* **8:** 1577-1587.

Murty D, Kirschbaum MUF, McMurtrie RE, McGilvray A (2002). Does conversion of forest to agricultural land change soil carbon and nitrogen? a review of the literature. *Global Change Biol* **8:** 105-123.

Nielsen UN, Wall DH, Six J (2015). Soil Biodiversity and the Environment. *Annu Rev Env Resour* **40:** 63-90.

Ofek-Lalzar M, Sela N, Goldman-Voronov M, Green SJ, Hadar Y, Minz D (2014). Niche and host-associated functional signatures of the root surface microbiome. *Nat Commun* **5**.

Pastorelli R, Landi S, Trabelsi D, Piccolo R, Mengoni A, Bazzicalupo M *et al* (2011). Effects of soil management on structure and activity of denitrifying bacterial communities. *Appl Soil Ecol* **49:** 46-58.

Pedros-Alio C (2006). Marine microbial diversity: can it be determined? *Trends Microbiol* **14:** 257-263.

Pedros-Alio C (2012). The Rare Bacterial Biosphere. *Annu Rev Mar Sci* **4:** 449-466.

Pester M, Bittner N, Deevong P, Wagner M, Loy A (2010). A 'rare biosphere' microorganism contributes to sulfate reduction in a peatland. *Isme J* **4:** 1591-1602.

Postma-Blaauw MB, de Vries FT, de Goede RGM, Bloem J, Faber JH, Brussaard L (2005). Within-trophic group interactions of bacterivorous nematode species and their effects on the bacterial community and nitrogen mineralization. *Oecologia* **142:** 428-439.

Prakamhang J, Tittabutr P, Boonkerd N, Teamtisong K, Uchiumi T, Abe M *et al* (2015). Proposed some interactions at molecular level of PGPR coinoculated with Bradyrhizobium diazoefficiens USDA110 and B-japonicum THA6 on soybean symbiosis and its potential of field application. *Appl Soil Ecol* **85:** 38-49.

Qin JJ, Li RQ, Raes J, Arumugam M, Burgdorf KS, Manichanh C *et al* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464:** 59-U70.

Quecine MC, Araujo WL, Rossetto PB, Ferreira A, Tsui S, Lacava PT *et al* (2012). Sugarcane Growth Promotion by the Endophytic Bacterium Pantoea agglomerans 33.1. *Appl Environ Microb* **78:** 7511-7518.

Rodrigues JLM, Pellizari VH, Mueller R, Baek K, Jesus ED, Paula FS *et al* (2013). Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. *P Natl Acad Sci USA* **110:** 988-993.

Shanmugam V, Kanoujia N, Singh M, Singh S, Prasad R (2011). Biocontrol of vascular wilt and corm rot of gladiolus caused by Fusarium oxysporum f. sp. gladioli using plant growth promoting rhizobacterial mixture. *Crop Prot* **30:** 807-813.

Tahir M, Mirza MS, Hameed S, Dimitrov MR, Smidt H (2015). Cultivation-Based and Molecular Assessment of Bacterial Diversity in the Rhizosheath of Wheat under Different Crop Rotations. *Plos One* **10**.

Thiele-Bruhn S, Bloem J, de Vries FT, Kalbitz K, Wagg C (2012). Linking soil biodiversity and agricultural soil management. *Curr Opin Env Sust* **4:** 523-528.

Tilman D, Knops J, Wedin D, Reich P, Ritchie M, Siemann E (1997). The influence of functional diversity and composition on ecosystem processes. *Science* **277:** 1300-1302.

Tittabutr P, Piromyou P, Longtonglang A, Noisa-Ngiam R, Boonkerd N, Teaumroong N (2013). Alleviation of the effect of environmental stresses using co-inoculation of mungbean by Bradyrhizobium and rhizobacteria containing stress-induced ACC deaminase enzyme. *Soil Sci Plant Nutr* **59:** 559-571.

Torsvik V, Ovreas L (2002). Microbial diversity and function in soil: from genes to ecosystems. *Current Opinion in Microbiology* **5:** 240-245.

Tremaroli V, Backhed F (2012). Functional interactions between the gut microbiota and host metabolism. *Nature* **489:** 242-249.

Ullah S, Bano A (2015). Isolation of plant-growth-promoting rhizobacteria from rhizospheric soil of halophytes and their impact on maize (Zea mays L.) under induced soil salinity. *Can J Microbiol* **61:** 307-313.

Uroz S, Buee M, Murat C, Frey-Klett P, Martin F (2010). Pyrosequencing reveals a contrasted bacterial diversity between oak rhizosphere and surrounding soil. *Env Microbiol Rep* **2:** 281-288.

van de Voorde TFJ, van der Putten WH, Bezemer TM (2011). Intra- and interspecific plant-soil interactions, soil legacies and priority effects during old-field succession. *J Ecol* **99:** 945-953.

van de Voorde TFJ, van der Putten WH, Bezemer TM (2012). The importance of plant-soil interactions, soil nutrients, and plant life history traits for the temporal dynamics of Jacobaea vulgaris in a chronosequence of old-fields. *Oikos* **121:** 1251-1262.

Van der Gucht K, Cottenie K, Muylaert K, Vloemans N, Cousin S, Declerck S *et al* (2007). The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. *P Natl Acad Sci USA* **104:** 20404-20409.

van Overbeek L, van Elsas JD (2008). Effects of plant genotype and growth stage on the structure of bacterial communities associated with potato (Solanum tuberosum L.). *Fems Microbiol Ecol* **64:** 283-296.

Verbruggen E, Roling WFM, Gamper HA, Kowalchuk GA, Verhoef HA, van der Heijden MGA (2010). Positive effects of organic farming on below-ground mutualists: large-scale

comparison of mycorrhizal fungal communities in agricultural soils. *New Phytol* **186:** 968-979.

Wall DH, Nielsen UN, Six J (2015). Soil biodiversity and human health. *Nature* **528:** 69-76.

Wall DH, Six J (2015). Give soils their due. *Science* **347:** 695-695.

Xu ZJ, Malmer D, Langille MGI, Way SF, Knight R (2014). Which is more important for classifying microbial communities: who's there or what they can do? *ISME J* **8:** 2357-2359.

Summary

Soil biodiversity is huge and determines largely the functioning of terrestrial ecosystems both at the 'macro' and the 'micro' level. Despite the general acceptance of the large impact of land use and other human activities on species loss in terrestrial ecosystems, their effects on microbial species reduction and the consequences are largely unknown. A major reason is the scarcity of experimental approaches to assess the relevance of soil microbial diversity for the functioning of soil ecosystems. The main goal of the study described in this thesis was to obtain better understanding of the diversity, structuring and functioning of bacterial communities in soil and and rhizosphere. With that purpose, we initially applied the rather old dilution approach to manipulate the diversity of microbial communities in soil by inoculation and subsequent incubation of more or less diluted soil suspensions in pre-sterilized soils.

In order to evaluate the potentials provided by the dilution approach to manipulate bacterial species diversity I combined this approach with next generation sequencing and advanced bioinformatics and statistical procedures. In Chapter 2, I revisited this approach and shed more light on the assemblage processes of bacterial communities in soil. The results presented in Chapters 2 and 3, showed the overriding impact of soil on the assemblage of bacterial communities from suspensions with different diversities. The dilution approach reduced dramatically the community diversity at the species or OTU level of the suspensions at the species or OTU level. Both composition and diversity of the regrown bacterial communities were altered dramatically compared to the original suspension. This indicates the soil's selective power during the assemblage of bacterial communities in terrestrial ecosystems. Remarkably, rare bacterial species could not be separated from abundant members as unique OTU's were detected in the most diluted suspensions and the soil communities that developed from these most diluted suspensions.

I also assessed the bacterial community assemblage in the rhizosphere as well as the functionality of the communities formed after inoculation of suspensions with communities with different diversities (Chapter 4). The bacterial diversity as well the composition of the community and functionality were significantly different between soil and rhizosphere. Network analysis revealed stronger interactions among bacterial OTU's in the rhizosphere than in

the soil. The enrichment processes in the rhizosphere selected microbes with particular functional genes mainly related to transporters, Embden-Meyerhof-Parnas, EMP, pathway and hydrogen metabolism. This selection was not random across bacteria with these functional traits, but specific for particular groups of bacteria containing these functions. Hence, I concluded that the selection processes in the rhizosphere are mainly driven by the functionality of the bacteria.

The potential effects of the bacterial communities in the rhizosphere on plant growth were further investigated in detail in Chapter 5. Plant biomass was on average the lowest and showed the highest variation in soils inoculated with the undiluted suspension harboring the most diverse community. A combination of unsupervised multivariate statistics and partial correlations showed that *Arthrobacter* was the taxonomical group that was most strongly related to plant biomass. A similar analysis showed that 'transporters' genes were the functional genes that were most strongly correlated to plant biomass. The frequency of 'transporters' genes was higher in *Arthrobacter* than in other components of the bacterial community. Partial correlation indicated further that the frequency of 'transporters' genes was more important than the frequency of *Arthrobacter* species to explain the effects on plant biomass production. Based on the results presented in Chapters 4 and 5, it can be concluded that plants select for specific bacterial species in the rhizosphere based on their functional properties, which benefit plant growth.

Samenvatting

De biodiversiteit in de bodem is enorm en bepaalt in hoge mate het functioneren van terrestrische ecosystemen, zowel op 'macro' als op 'micro' niveau. Ondanks het feit dat de grote invloed van land gebruik en menselijke activiteiten op het verlies aan soorten algemeen geaccepteerd is, zijn de effecten van het verlies aan microbiële soorten grotendeels onbekend. Een belangrijke reden hiervoor is het gebrek aan geschikte experimentele benaderingen om de relevantie van de microbiële biodiversiteit in de bodem voor het functioneren van bodem ecosystemen vast te stellen. De primaire doelstelling van het onderzoek dat in dit proefschrift beschreven wordt, was het beter begrijpen van de vorming, diversiteit en functioneren van bacteriële gemeenschappen in de bodem en de rhizosfeer. Voor dat doel, hebben we in eerste instantie de tamelijk oude verdunningsmethode toegepast voor het manipuleren van de diversiteit van microbiële gemeenschappen in de bodem door de inoculatie van meer of minder verdunde bodem suspensies in vooraf gesteriliseerde bodems.

Om de mogelijkheden van de verdunningsmethode voor het manipuleren van de diversiteit van bacteriën te evalueren, heb ik deze methode gecombineerd met de nieuwste sequentie analyse bepalingsmethodieken en geavanceerde bioinformatica en statistische technieken. In hoofdstuk 2, heb ik deze benadering getoetst en heb ik beter zicht gekregen op de vorming van bacteriële gemeenschappen in de bodem. De resultaten die in dit hoofdstuk en het volgende hoofdstuk 3 beschreven worden, laten het overheersende effect van de bodem op de vorming van bacteriële gemeenschappen vanuit geïnoculeerde suspensies met verschillende diversiteiten zien. Door toepassing van de verdunningsmethode kon de diversiteit van bodem suspensies op soort en OTU niveau aanzienlijk gereduceerd worden. Zowel de samenstelling als de diversiteit van de uiteindelijk gevormde gemeenschappen in de bodem verschilden enorm ten opzichte van de oorspronkelijke geïnoculeerde suspensies. Dit illustreert eens te meer de selecterende kracht van de bodem gedurende de vorming van bacteriële gemeenschappen in terrestrische ecosystemen. Het was opmerkelijk dat zeldzame bacteriële soorten niet konden worden gescheiden van de veelvuldig aanwezige soorten aangezien unieke OTU's ook werden gedetecteerd in de meest verdunde suspensies en in de bodem gemeenschappen die waren gevormd uit deze meest verdunde suspensies

Ik heb ook de vorming van bacteriële gemeenschappen in de rhizosfeeer als mede de functionaliteit van de gemeenschappen die waren gevormd na inoculatie van suspensies met gemeenschappen met verschillende diversiteit (hoofdstuk 4). De diversiteit en samenstelling van de gemeenschappen en hun functionaliteit verschilden significant tussen (wortelvrije) bodem en rhizosfeer. Netwerk analyses lieten een sterkere interacties zien tussen bacteriën in de rhizosfeer dan in de bodem. De selectie in de rhizosfeer betrof vooral bacteriën met specifieke genen die voornamelijk gerelateerd waren aan transport processen, de Embden-Meyerhof-Parnas, EMP, metabole route en waterstof metabolisme. Deze selectie was niet willekeurig verspreid over alle bacterie soorten. Daaruit heb ik geconcludeerd dat de selectie processen in de rhizosfeer voornamelijk gedreven worden door de functionele eigenschappen van de bacteriën.

De potentiele effecten van bacteriële gemeenschappen in de rhizosfeer op de groei van planten is in meer detail onderzocht en beschreven in Hoofdstuk 5. De biomassa productie was gemiddeld het laagst en het meest variabel in bodems die geïnoculeerd waren met de onverdunde suspensies met de hoogste diversiteit. Door toepassing van een combinatie van niet-geleide (unsupervised) multivariate statistiek en partiële correlatie kon worden aangetoond dat *Arthrobacter* de taxonomische groep was die het meest gerelateerd was aan de planten biomassa productie. Op een zelfde manier kon worden vastgesteld dat genen die betrokken zijn bij transport processen het best correleerden met plant biomassa. De frequentie van ‘transporter’ genen was hoger in *Arthrobacter* dan in enig andere component van de bacteriële gemeenschap. Partiele correlatie wees verder uit dat de frequentie van ‘transporters’ genen belangrijker was dan de frequentie van *Arthrobacter* soorten om de effecten op plant biomassa productie te verklaren. Op basis van de resultaten die beschreven zijn in de hoofdstukken 4 en 5, kan worden geconcludeerd dat planten specifieke bacteriële soorten selecteren op basis van functionele eigenschappen die een positief effect hebben op planten groei.

Acknowledgements

The four years PhD journey is definitely one of the best periods in my life! I really enjoyed keeping on learning and the countless inspiring discussions.

Many thanks to the Plant Ecology Group. I enjoyed the time that we spent together. Special thanks to Karin who provided valuable help in collecting hundreds of kilos of soil and taught me so many experimental skills. Thanks to Tom who learned me a lot of statistics. Thanks to my roommates, Rocio and Saskia. It was a great pleasure to share my room with both of you and for the many discussions. Thanks to Dr. Wei who shared PhD journey with me in the past four years. Congratulations to my fellow graduates Xiaojie, Kaixuan, Yingjie, Suzanne, Yangan and Gang Chen; I wish you all well as you enter the next stage in your lives.

My thanks also go to the colleagues of the Department of Microbial Ecology of NIOO-KNAW. Everyone was always ready to help! Thanks for all the nice discussions during my presentations. Special thanks to the Kuramae research group. Learning Portuguese is one of my plans in the future. Special thanks also to our bioinfomatician, Mattias. Without your help in analyzing hundreds of samples, I could not have finished the work. Agaat, Saskia and Iris are acknowledged for valuable help in my experiments. Thanks also to Gera, Chunxu, Yani, Mauricio, Lucas, Thiago, Manoeli, Marcio, Nori, Olaf, Irene, Viviane and Ruth, your kind efforts and interesting discussions always help me solving problems. I really appreciated that.

My family was always there for me and provided strength to get me through the tough times. Especially, thanks to my dear parents, no matter what my choice is, all your unconditional love and encouragement made me a strong person. 爸爸妈妈永远是我们的强大后盾和精神支柱!

Yihao, we got both Bachelor and Master degree together, and now, the PhD degree. I am ready to be always with you to see each beautiful step of life. Keep on dreaming.

Yan Yan

Leiden, April, 2016

Curriculum Vitae

Yan Yan was born on January 10th 1986, in Shijiazhuang, Hebei province, China. In 2008 she obtained her Bachelor degree of Biological Engineering at the Agricultural University of Hebei, China. In the same year, she started her Master study at the College of Food Safety, Agricultural University of Hebei, China, supervised by Prof. dr. Hongtao Tian. From 2009-2011, she performed her Master internship in Prof. dr. Kunlun Huang's lab at the China Agricultural University in Beijing. She finished her Master thesis and obtained her master degree in 2011. The same year she started the PhD project described in this thesis at the Department of Microbial Ecology of the Netherlands Institute of Ecology (NIOO-KNAW) and at the Institute of Biology Leiden, Leiden University, under the supervision of Prof. dr. Hans van Veen, Prof. dr. Peter Klinkhamer and Dr. Eiko Kuramae.

Publications

Yan Y., De Hollander M., Kuramae E. E., Van Veen, J.A., Klinkhamer P.G.L. The effects of salicylic acid determine plant soil feedback. (*in preparation*)

Yan Y., De Hollander M., Klinkhamer P.G.L., and Van Veen, J.A., Kuramae E. E. Soil characteristics determine the assemblage of bacterial communities in terrestrial ecosystems. (*Manuscript*)

Yan Y., Kuramae E. E., De Hollander M., and Van Veen, J.A., Klinkhamer P.G.L. Functional genes rather than taxonomical composition of the rhizosphere bacterial community determine plant biomass in *Jacobaea vulgaris*. (*Manuscript*).

Yan Y., Kuramae E. E., De Hollander M., Klinkhamer P.G.L., and Van Veen, J.A. (2016). Functional traits determine rhizosphere selection of bacterial communities. (***ISME Journal*** *in revision*)

Yan Y., Kuramae E. E., Klinkhamer P.G.L., and Van Veen, J.A. (2015). Revisiting the dilution procedure used to manipulate microbial biodiversity in terrestrial systems. ***Applied and Environmental Microbiology***, 81 (13), 4246-4252.

Yan Y., Xu W, Luo Y, Tian H. Study on the methods of extraction of *Pleurotus ostreatus* genomic DNA. ***Science and Technology of Food Industry***, 2011 (09).