

# **Phylogenetic and functional diversity of soil prokaryotic communities in temperate deciduous forests with different tree species**

Dissertation  
for the award of the degree  
"Doctor of Philosophy" Ph.D. Division of Mathematics and Natural  
Sciences  
of the Georg-August-Universität Göttingen  
within the doctoral program in Biology  
of the Georg-August University School of Science (GAUSS)

submitted by  
Amélie Dukunde  
from Nyamyumba (Rwanda)

Göttingen, 2018

### **Thesis committee**

Prof. Dr. Rolf Daniel, Department of Genomics and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August Universität Göttingen

PD. Dr. Michael Hoppert, Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August Universität Göttingen

### **Members of the examination board**

Reviewer: Prof. Dr. Rolf Daniel, Department of Genomics and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August Universität Göttingen

Second Reviewer: PD. Dr. Michael Hoppert, Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August Universität Göttingen

### **Further members of the examination board:**

PD. Dr. Fabian Commichau, Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August Universität Göttingen

Prof. Dr. Kai Heibel, Department of Molecular Microbiology and Genetics, Institute of Microbiology and Genetics, Georg-August Universität Göttingen

PD. Dr. Ina Meier, Plant Ecology and Ecosystem Research, Albrecht von Haller Institute of Plant Sciences

Prof. Dr. Stefanie Pöggeler, Department of Genetics of Eukaryotic Microorganisms, Institute of Microbiology and Genetics, Georg-August Universität Göttingen

Date of the oral examination: 17.05.2018

# Table of contents

---

|   |           |
|---|-----------|
| <b>Chapter 1: Literature review</b> .....   | <b>5</b>  |
| 1.1 Overview of forest biomes .....   | 5         |
| 1.2 Temperate deciduous forests .....   | 7         |
| 1.2.1 Structure and vegetation type .....   | 7         |
| 1.2.2 Forest microbial diversity .....  | 9         |
| 1.3 Research overview and design .....  | 13        |
| 1.3.1 The Hainich National Park .....   | 13        |
| 1.3.2 Objective of thesis .....   | 14        |
| <b>Chapter 2: Tree species shape soil bacterial community structure and function in temperate deciduous forests</b> ..... | <b>15</b> |
| 2.1 Abstract .....  | 16        |
| 2.2 Introduction .....  | 17        |
| 2.3 Methods and materials .....   | 19        |
| 2.3.1 Sample site description .....   | 19        |
| 2.3.2 Sampling and environmental nucleic acid isolation .....   | 20        |
| 2.3.3 Processing of 16S rRNA gene sequence data and statistical analyses .....  | 21        |
| 2.3.4 Accession numbers .....   | 22        |
| 2.4 Results and Discussion .....  | 23        |
| 2.4.1 Edaphic properties reflect stand-related variations .....   | 23        |
| 2.4.2 Stand-specific patterns in bacterial richness and diversity .....   | 24        |
| 2.4.3 Tree species effect of bacterial community structure .....  | 24        |
| 2.4.4 General patterns in bacterial community composition and structure .....   | 26        |
| 2.4.5 Taxa-habitat association patterns .....   | 30        |
| 2.4.6 Bacterial functional profiles across forest stands .....  | 32        |
| 2.5 Conclusion .....  | 36        |
| Acknowledgments .....   | 36        |
| 2.6 Supplementary information .....   | 37        |
| <b>Chapter 3: Bioprospecting for novel biocatalysts</b> .....   | <b>49</b> |
| 3.1 Harvesting the soil metagenome .....  | 49        |
| 3.2 Functional screening for lignocellulolytic enzymes .....  | 52        |
| 3.2.1 Esterases .....   | 52        |
| 3.2.2 Laccases .....  | 54        |

|   |            |
|---|------------|
| 3.2.3 Lignocellulases .....   | 54         |
| 3.2.4 Objectives:.....  | 55         |
| 3.3 Methods and materials .....   | 56         |
| 3.3.1 General molecular cloning techniques .....  | 56         |
| 3.3.2 Function-based screening.....   | 57         |
| 3.4 Results and discussion.....   | 59         |
| 3.4.1 Construction metagenomic plasmid libraries.....   | 59         |
| 3.4.2 Screening for lignocellulolytic enzymes .....   | 59         |
| <b>Chapter 4: A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum .....</b> | <b>65</b>  |
| 4.1 Publication .....   | 65         |
| 4.2 Supporting material.....  | 77         |
| <b>Chapter 5: Summary .....</b>   | <b>82</b>  |
| <b>References .....</b>   | <b>84</b>  |
| <b>Acknowledgements .....</b>   | <b>102</b> |
| <b>Appendices .....</b>   | <b>104</b> |
| List of non-standard abbreviations.....   | 104        |
| Academic curriculum vitae.....  | 105        |
| Thesis Declaration.....   | 106        |

# Chapter 1

---

## Literature review

### 1.1 Overview of forest biomes

The forest biome is the largest terrestrial ecosystem and covers approximately 4 billion hectares or 30% of Earth's land area (FAO, 2006). It has global significance as the largest habitat of the of Earth's terrestrial biodiversity and the greatest source of plant biomass (80%), which contributes as much as 75% to gross primary productivity (Baldrian, 2017; Pan et al., 2013). As such, forests act as major sinks of stored carbon and facilitates biogeochemical cycling of other elements and regulating the global climate (Lladó et al., 2017). Characterized by the dominance of trees and layered vegetation, forests are distributed across several biogeographic and climate zones (Figure 1.1). Forests also have numerous benefits for humans, by offering goods and services such as timber, biofuels, clean water, agricultural products (food and medicine) and shelter, cultural and recreational services (Macdicken et al., 2015; Pan et al., 2013; Thompson et al., 2011).

With such fundamental global implications, the ecology of forest ecosystems is of great interest. Management practices to preserve such an integral part in ecosystem functioning is the focus of several global environmental initiatives (MacDicken, 2015). Recent advances in tools to measure qualitative and quantitative changes in forests at high resolution have led to a better understanding of the dynamic processes at play, and aid in developing solutions to maintain and improve the world's forests.

Definitions of forests often vary depending on administrative, land-use and land cover contexts (see **Box 1.1 – Defining Forests**), however, three main forest types are recognized: boreal or taiga, temperate, and tropical forests (Amano et al., 2000; Anon. and FAO, 2016).

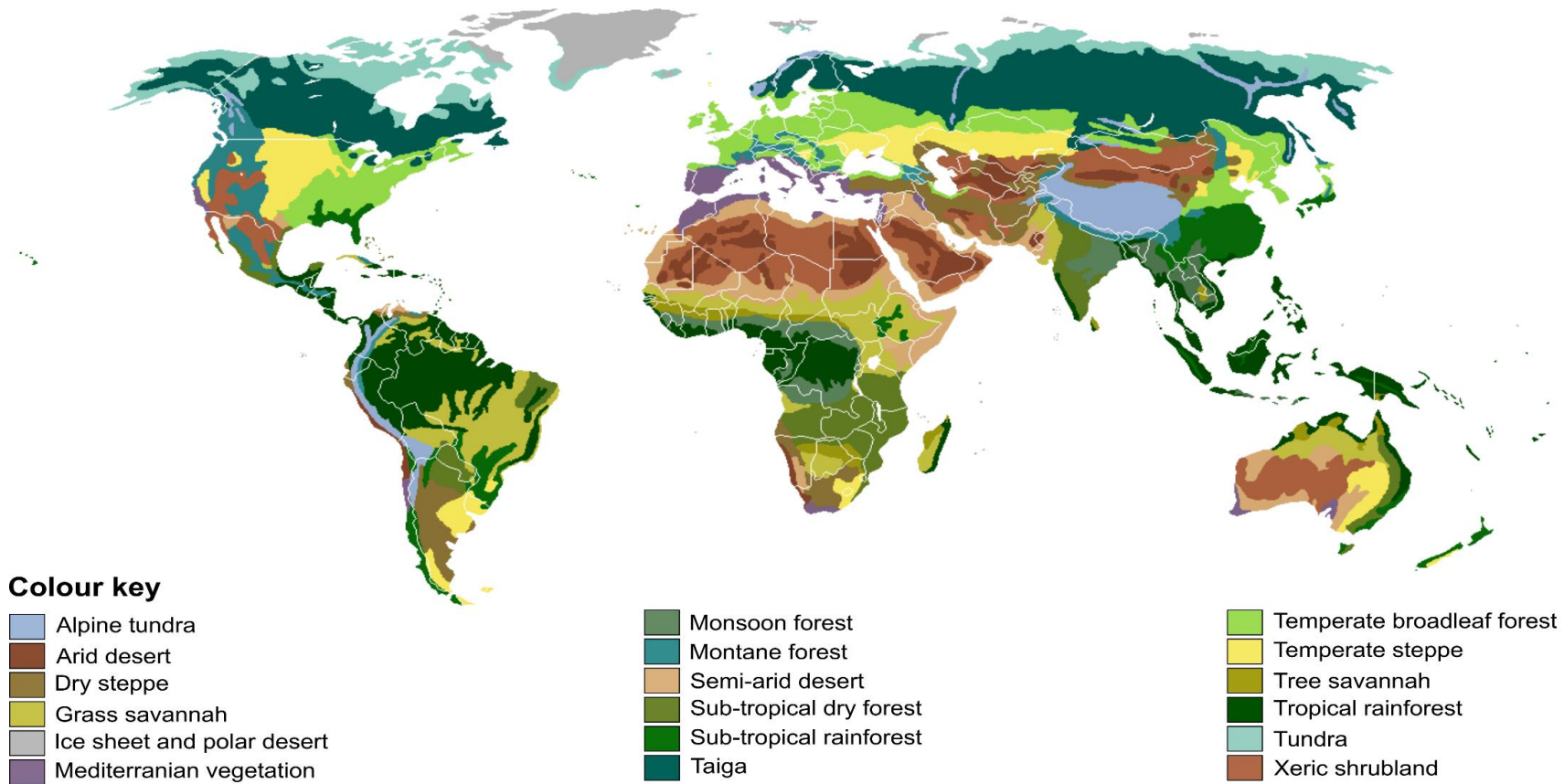


Figure 1.1 Distribution of terrestrial biomes, including forest biomes, across the globe. Image adapted from Ville Koistinen (user Vzb83) CC BY-SA 3.0 (<https://commons.wikimedia.org/w/index.php?curid=1700408>).

The highest terrestrial biodiversity and forest carbon stocks are found in tropical forests, located about the equatorial belt (Bonan, 2008). The high density of evergreen or semi-deciduous tree species (200 – 300 trees per hectare) and tree species diversity makes tropical forest ecosystems vary from region to region, compared to boreal and temperate forests, which have lower tree species diversity (Condit et al., 2002; Malhi et al., 1999). Boreal forests are located close to the poles and are characterized by a low diversity of coniferous tree species, such as pine (*Pinus*) and spruce (*Picea*), and a few deciduous species such as birch (*Betula*) and poplar (*Populus*) (Malhi et al., 1999).

Temperate forests occupy mid-latitude regions between polar regions and the Tropics of Cancer and Capricorn, respectively in the northern and southern hemisphere (Figure 1.1). A prominent feature of temperate forests is the sensitivity to seasonality, and broad-leaved tree species which undergo distinct seasonal changes: growing leaves in spring and undergoing leaf fall in the autumn, as an adaptation to cold winters (Reich and Frelich, 2002). As the focus of thesis, temperate forests are described further.

## 1.2 Temperate deciduous forests

Temperate forests represent 21.8 % of forested area (Crowther et al., 2015), and have been extensively impacted by human activity, which has drastically altered their diversity and structure (Augusto et al., 2002). Since human settlement in temperate regions, communities used forest resources as fuel, building materials and converted swaths of land for settlement and agriculture. The advent of the industrial revolution in the 18<sup>th</sup> century placed an even greater demand for forestry products on temperate forests in Europe and northern America. Consequently, temperate forests have undergone exploitation for much longer and more intensively than tropical and boreal forests (Reich and Frelich, 2002; Schulze et al., 2016).

### 1.2.1 Structure and vegetation type

Temperate deciduous forests grow in a characteristic four-layer arrangement (Figure 1.2). Mature tree species form the forest canopy, the tallest layer. This layer is densest in summer and limits the amount of sunlight reaching the understory (Barbier et al., 2008).

### **Box 1.1 – Defining Forests**

**Administrative:** A governing body can declare a forest as a legislative region, i.e. “Any lands falling within the jurisdiction of Department XYZ”, with no relation to vegetation characteristics of the area.

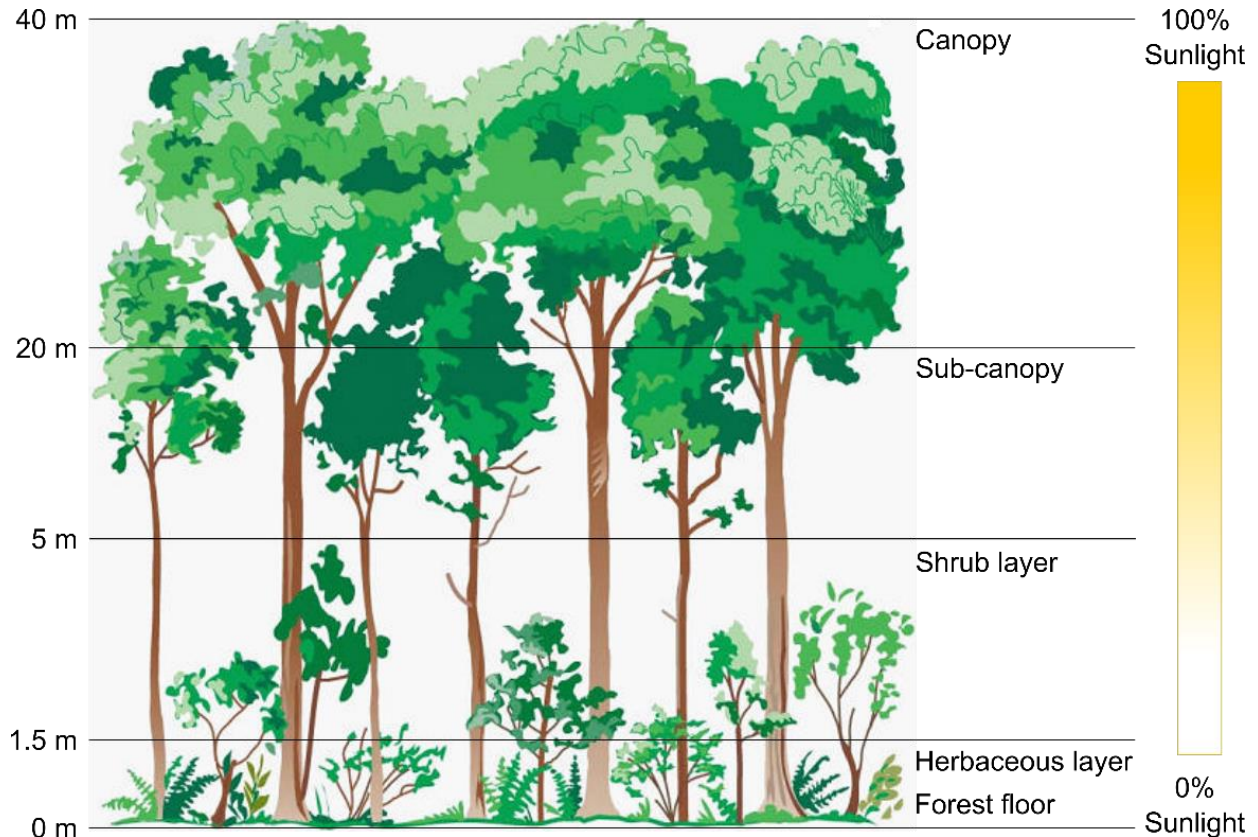
**Land cover:** Forests may also be defined by the crown cover, i.e. the ground area covered by the canopy shade and tree density. Although this considers vegetation, the vegetation type is not specified and may include urban parks and agricultural tree crops.

**Land use:** An area is designated a forest based on the management practice taking place, from an undisturbed (pristine) environment to a region with a mixture of natural vegetation and occasional management practices which select for beneficial forest products, such as logging.

Based on a global consensus with individual countries, the combination of land cover and land-use is used to define forest biomes. According to the Food and Agriculture organization of the United Nations (FAO), a forest is as an area with a minimum size of 0.5 hectares, 10% canopy cover and trees with a minimum height of 5 m.

From Watson et al., (2000) and FAO, (2016).

The sub-canopy lies below the main tree canopy, consisting of smaller trees, young saplings awaiting to break into the main canopy, shrubs and small bushes (Senécal et al., 2018). Most plants in this layer are shade tolerant and have adapted to the light-limiting canopy (Lu et al., 2016). The herbaceous layer grows very close to the forest floor and is made up of ephemeral flowers, mosses and other herbaceous species, which have adapted to the closed overstory by flowering much earlier in the spring, before the emergence of tree leaves (Fridley, 2012; Gilliam, 2007).



**Figure 1.2 Diagrammatic representation of the four vegetation layers of a temperate deciduous forest. Height above ground of forest layers is an approximation and is not drawn to scale. Image source (Cain et al., 2011).**

Temperate forests can have needleleaf, coniferous species such as fir (*Abies*) or cedar (*Cedrus*); broadleaf species, such as maple (*Acer*) or ash (*Fraxinus*); or a mixture of both leaf forms. In warmer temperate regions, in Australasia and South America, broadleaf evergreen species such as *Eucalyptus* and *Nothofagus* dominate (Ciesla, 2002; Gandolfo et al., 2011). Most temperate broadleaf forests are concentrated in the northern hemisphere and comprise deciduous tree species. Hereon, the focus will be on European temperate deciduous forests.

Due to intensive land-use practices, the diversity of European deciduous forests has been greatly reduced, and most are dominated by the *Fagus* genus (beech). Tree species diversity decreases with stand age; thus, it is more common to find a monoculture of beech stands, rather than mixed species (Decocq et al., 2004). Nevertheless, Central European forests host at least 55 forest tree species, extensively reviewed by Schulze et al., (2016).



## 1.2.2 Forest microbial diversity

Microorganisms inhabit almost all habitats on our planet including forests and play an important role by driving biogeochemical processes, through the decomposition of plant-/animal-derived organic material and metabolism of plant root exudates in forest ecosystems (Thoms and Gleixner, 2013). Their small size and diverse functional adaptations enables them to colonise several, often overlapping niches within the forest, including vegetation foliage, wood bark and deadwood, atmosphere, rhizosphere and rock surfaces, streams, and invertebrates (Baldrian, 2017). As such, microbial habitats create continuous connection of all forest layers, in a manner more complex than plant vegetation (Baldrian, 2017; Gilliam, 2007).

### 1.2.2.1 Soil bacterial diversity

The soil environment is one of the largest terrestrial reservoirs of microbial diversity (Torsvik and Øvreås, 2002). Physical characteristics such as mineral particle size, texture, pore size and content of organic and inorganic compounds show large spatial heterogeneity (Fierer et al., 2009). Therefore, as an abiotic matrix, it is complex in composition and enables the formation of a multi-layered and multi-functional habitat for soil bacteria. The soil microbiota consists of soil microfauna, such as nematodes and rotifers, fungi and prokaryotes, and plant roots (Baldrian, 2017; Chen et al., 2007; Devetter et al., 2017). Hereon, the focus will be on soil bacterial communities.

Understanding drivers of soil bacterial diversity and community structure improves our ability to predict structural and functional environmental shifts, ecosystem responses to such shifts, and formulate solutions to counter those changes that are detrimental (Nacke, 2011). Bacterial communities are prevalent in the rhizosphere and bulk soil. Their distribution is often dependent on edaphic properties and aboveground vegetation (Colin et al., 2017; Jeanbille et al., 2016; Mendes et al., 2013; Prescott and Grayston, 2013). Despite high bacterial taxonomic diversity, and the heterogeneity of soils, a ‘fingerprint’ of soil habitats has emerged; most of the bacterial diversity is represented by a few phyla (Fierer et al., 2009) mainly, high populations of *Acidobacteria*, *Proteobacteria*, *Actinobacteria* and lower populations of *Bacteroidetes*, and *Firmicutes* (Herzog et al., 2015; Li et al., 2014b; López-Mondéjar et al., 2015; Uroz et al., 2013). Variations in bacterial community structure are the result of disturbances in the microhabitat, such as fire, rather than global changes in the parent soil material (Fierer et al., 2009).

Soil pH is an important predictor of bacterial diversity and structure (Lauber et al., 2009; Romanowicz et al., 2016), as dominant phyla have shown significant changes along pH gradient (Peralta et al., 2013; Tripathi et al., 2013; Uroz et al., 2016). However, taxa-specific patterns in vertical stratification have also been documented, whereby bacterial diversity decreased with increasing soil depth due to a decrease in soil organic matter and nutrients (Eilers et al., 2012; Lladó et al., 2017; López-Mondéjar et al., 2015).

### 1.2.2.2 Impact of tree species on soil bacterial communities

As the dominant vegetation type, trees make the central force around which forest productivity rotates. The relationship between aboveground plant diversity and belowground microbial diversity is of significant ecological interest, due to the complex interactions between biotic and abiotic drivers of

ecosystem functioning. Aboveground tree diversity alters the soil environment through the chemistry of litterfall, nutrient uptake at the roots and efficiency in water and light use (Lang and Polle, 2011; Langenbruch et al., 2012). At the interface, microorganisms mediate the transformation of nutrients for forest carbon and nitrogen stocks, and sulfur and phosphorus for the global biogeochemical cycle (Cline et al., 2017).

Bacteria, which are present in all ecological niches in the forest, participate at all levels of biochemical transformation processes (Figure 1.3) (Lladó et al., 2017). Bacterial communities generally carry out reactions associated with a subset of a biogeochemical pathway, instead of the complete pathway (Castelle and Banfield, 2018). Elements from this pathway may be used as reaction intermediates in a different cycle. This results in the coupling of biogeochemical cycles, whereby the availability of elements from one cycle limits some pathways of another biogeochemical cycle (Castelle and Banfield, 2018; Finzi et al., 2011; Schlesinger et al., 2011).

Soil bacteria are involved in mineralization processes and conduct nutrient uptake from plants through the degradation of lignocellulosic material or metabolism of exudates at the root/soil interface. Nutrient exchange takes place more efficiently at the fine root/soil interface rather than in large, coarse and woody roots (Kubisch et al., 2015). Fine-root biomass is a species-dependent trait in trees and determines the surface area available for nutrient exchange. Additionally, as fine root colonization decreases with depth, associated bacteria tend to inhabit the uppermost horizons of topsoil (0 – 10 cm) (Thoms et al., 2010). Lastly, fine root tissue retains transient concentrations of C and N (Jacob et al., 2013; Kubisch et al., 2015), another trait that is tree species-specific. Taking all into consideration, these traits have potential to select for specific bacterial groups with the ability to exploit available resources.

Bacteria act as decomposers on litterfall and deadwood. Different functional traits in trees affect the quality of litter and deadwood debris, which in turn influence microbial communities differently (Lang and Polle, 2011; Seven and Polle, 2014), and promote preferential colonization by some taxa and not others. Bacteria are especially more sensitive to differences in nutrient quality, thereby resulting in a complex taxonomic structure.

Rates of decomposition in forest litter are tree species-dependent and likely to limit some bacteria (Cesarz et al., 2013b). Consequently, indicators of litter quality such as the carbon to nitrogen (C/N), carbon to phosphorus (C/P) and lignin to nitrogen (N) ratios are negatively correlated with decomposition rate, whereas N to calcium (Ca) ratio is positively correlated, are also plant-specific (Jacob et al., 2010). The availability of litter influences rates of nutrient uptake by soil bacteria, therefore, only growth-limiting decomposition reduces the diversity of soil bacteria in beech stands (Thoms and Gleixner, 2013), as beech trees produce low quality, recalcitrant litter which lowers soil pH (Pfeiffer et al., 2013).

Root exudates have garnered interest in their ability to impact bacterial community structure and composition. Studies on root exudates are not as numerous as those of forest plant litter, however, research indicates that exudates contain primarily labile compounds (Cesarz et al., 2013a) such as sugars and carbohydrates, amino acids, organic acids, secondary metabolites (Eisenhauer et al., 2017;

Zhalnina et al., 2018; Zwetsloot et al., 2018), and are therefore sources for C and N. Through root exudation, plants are able to recruit only certain beneficial soil bacterial communities in their environment (Zhalnina et al., 2018). Söderberg et al. (2004) demonstrated that the change in root C inputs, decrease in decomposable sugars and amino acids caused Gram-negative bacterial biomass to decrease with increasing soil depth and decreasing root mass, in contrast to Gram-Positive bacteria. Soil exchangeable C and N, and K vary with soil depth, and together with pH, act to restrict the growth of gram negative bacteria (Söderberg et al., 2004).

An overarching contributor to changes in tree functional traits is seasonality. As mentioned in section 1.1, seasonal changes characterize temperate deciduous forests. Changing temperatures induce changes in functional traits, such as leaf emergence and leaf-fall, the quality, quantity and timing of which are tree-specific (Rasche et al., 2011). Additionally, this directs the time at which nutrients, such as C and N, are available and thus indirectly affects ecosystem productivity (Kaiser et al., 2010). A feedback mechanism exists between plants, microbial communities and the environmental parameters of the habitat, such that a limitation in one component will bring about changes in all; however, the precise pathways of this mechanism are not well understood.

Forest management initiatives aim to increase forest productivity by the introduction of mixed species of forest stands. Mixed forest stands are advantageous over monocultures as they increase the substrate pool of available nutrients by offering complementary nutrient sources (Cardinale et al., 2012; Dawud et al., 2016). It remains to be seen whether tree species identity or tree species diversity are the drivers of productivity, and soil microbial composition. While this has been a successful dynamic in grasslands or microcosms (Eisenhauer et al., 2017), the same approach in temperate deciduous forests has variable outcomes (Khelifa et al., 2017).

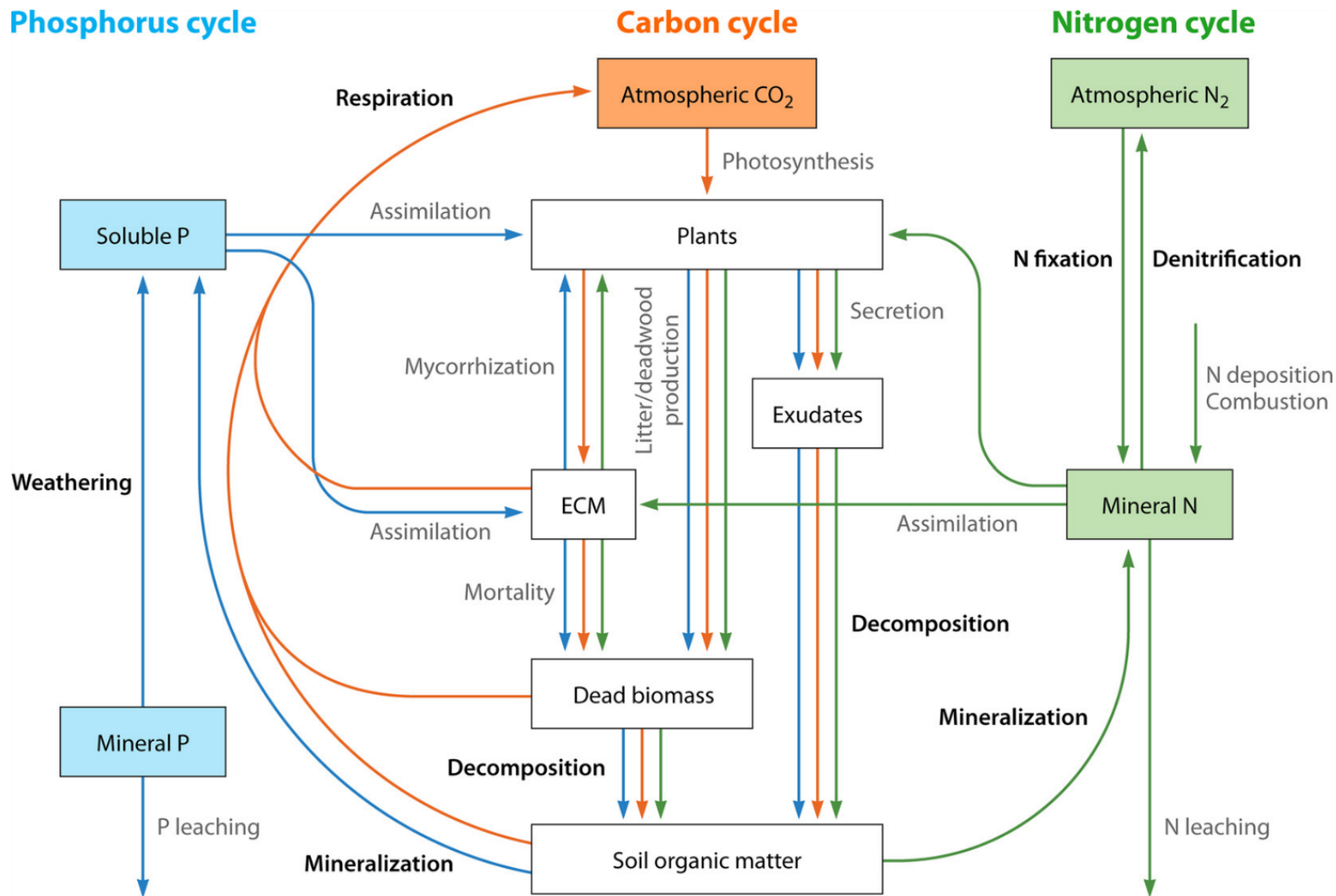


Figure 1.3 Diagrammatic representation of above ground and belowground interactions between forest vegetation and soil bacterial communities, as part of key biogeochemical cycles. Coloured arrows represent transfer of elements among different ecological niches. Image source: Lladó et al. (2017).

### 1.3 Research overview and design

The thesis is part of the Research Training Group 1086 “The Role of Biodiversity for Biogeochemical Cycles and Biotic Interactions in Temperate Deciduous Forests”, under the sub-project “A - Biodiversity analyses and biotic interactions”. Research aims to identify key relationships between tree species diversity and productivity, well as the implications for biotic interactions among organisms in a temperate deciduous forest.

#### 1.3.1 The Hainich National Park

All projects were carried out in the Hainich national park (Thuringia, Germany), the largest continuous wooded area in Germany. Previous silvicultural activities included coppicing and logging; however, all management activity ceased in 1990, and is therefore considered a semi-natural environment (Guckland, 2009). The area was declared a national park in 1997, and since 2011, a UNESCO World Heritage (Sommer, 2016). The forest grows on loess-derived Stagnic Luvisol on Triassic lime stone parent rock, with a pH range between  $4.5 \pm 0.5$  and  $5.8 \pm 0.6$  (Cesarz et al., 2013b). The tree layer is dominated by European beech (*Fagus sylvatica*) which covers more than 95% of the forest, with mature trees reaching an age of approximately 200 years (Seven, 2014). Interspersed within the beech community are ash (*Fraxinus excelsior*), maple (*Acer sp.*), lime or linden (*Tilia sp.*), oak (*Quercus sp.*), hornbeam (*Carpinus betulus*), Scotch elm (*Ulmus glabra*) and sweet cherry (*Prunus avium*) species (Seven, 2014).

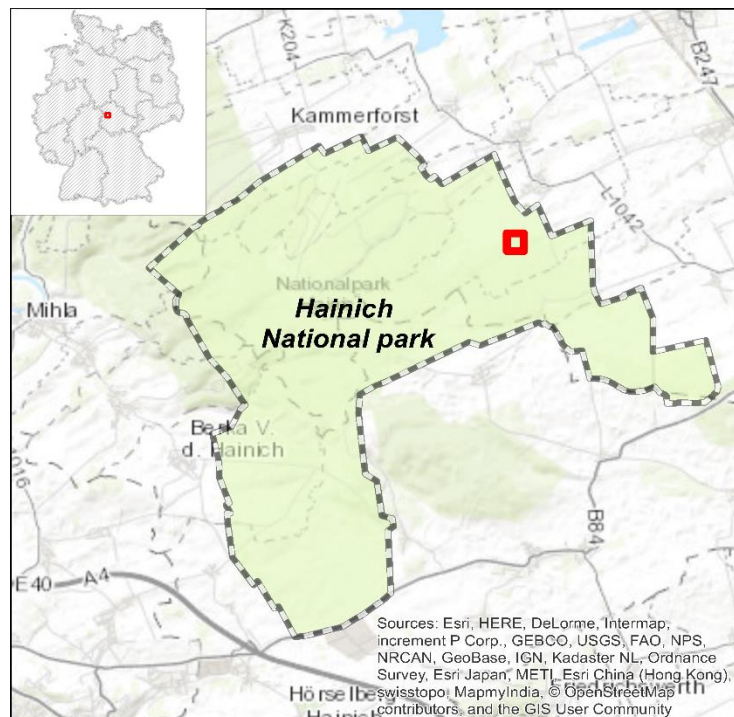


Figure 1. 4. Map of Hainich National Park in Thuringia, Germany (inset). Red point marks the location of the sampling area (51°05'37.0 N, 10°30'10.6 E).

The study was conducted on a 25-h area with four main tree species: beech, hornbeam, lime and oak. Oak species in this area comprised *Q. petraea* and, to lesser extent, *Q. robur*, while lime species belonged to *T. cordata* Mill (dominant) and *T. platyphyllos* (Schmidt et al., 2015). Trees belong to the *Rosid* clade; however, lime belongs to the *Marvales* order and is far removed from beech, hornbeam and oak, which belong to the *Fagales* order. Additionally, beech and oak belong to the *Fagaceae* family, unlike hornbeam, which belongs falls under *Betulaceae*.

### 1.3.2 Objective of thesis

Previous studies on the Hainich forest dynamics from the group concentrated on demonstrating the influence of tree species on endo/ectomycorrhizal interactions (Lang, 2008; Seven and Polle, 2014). This work aims to investigate tree species diversity and their impact on belowground soil bacterial community in a temperate deciduous forest, with the following specific aims:

1. To examine soil bacterial taxonomic structure and functional diversity as a function of tree species diversity, focusing on mono-specific and mixed species stands of beech, hornbeam, lime and oak trees (**Chapter 2**).
2. To recover potentially novel biocatalysts from the forest soil environment through a culture-independent approach: the construction of short-insert metagenomic plasmid libraries to screen for lignocellulosic enzymes (**Chapter 3**).
3. To characterize a biocatalyst from a metagenome-derived clone, as the first step to harvesting the soil metagenome for industrially relevant enzymes (**Chapter 4**).

# Chapter 2

---

## Tree species shape soil bacterial community structure and function in temperate deciduous forests

Amélie Dukunde<sup>1</sup>, Dominik Schneider<sup>1</sup>, Marcus Schmidt<sup>2</sup>, Rolf Daniel<sup>1\*</sup>

The following chapter has been prepared for submission.

<sup>1</sup> Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August University of Göttingen, Grisebachstraße 8, 37077, Göttingen, Germany

<sup>2</sup> Büsgen Institute – Soil Science of Tropical and Subtropical Ecosystems, Georg-August University of Göttingen, Büsgenweg 2, 37077 Göttingen, Germany

**Keywords:** forest soil bacterial community, soil bacteria, Hainich national park, temperate deciduous forest soils, tree species diversity, soil bacterial diversity

To whom correspondence should be addressed:

\* Rolf Daniel

Department of Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University Göttingen, Grisebachstraße 8, Germany.

Phone: 0049-551-393827 – Fax: 0049-551-3912181 – Email: [rdaniel@gwdg.de](mailto:rdaniel@gwdg.de)

Author contributions:

Conceived and designed the experiments: RD, **AD**

Performed the experiments: **AD**, MS

Analysed the results: **AD**, DS

Wrote the publication: **AD**, DS, RD

## 2.1 Abstract

Amplicon-based analysis of 16S rRNA genes and transcripts was used to assess the effect of tree species composition on soil bacterial community structure and function in a temperate deciduous forest. Samples were collected from mono and mixed stands of *Fagus sylvatica* (beech), *Carpinus betulus* (hornbeam), *Tilia* sp. (lime), and *Quercus* sp. (oak) in spring, summer and autumn. Total (DNA-based) and potentially active (RNA-based) soil communities exhibited a similar composition, with fewer taxa comprising the active community. Members of *Xanthobacteraceae* and *Bradyrhizobiaceae* dominated both total and active bacterial communities, followed by *Acidobacteriales* and *Solibacterales* (*Acidobacteria*), *Rhodospirillales* and *Xanthomonadales* (*Proteobacteria*). Bacterial communities at total and active community level significantly responded to tree species identity (mono stands) and to a lesser extent to tree species richness (mixed stands). Approximately 58 and 64 % of significant indicator OTUs were associated with only one mono stand at total and active community level, respectively, indicating a strong impact of tree species on soil bacterial community composition. Soil C/N ratio, pH and P content exhibited significant impact on soil bacterial communities and were attributed to direct and indirect effects of forest stands. Seasonality was the strongest driver of predicted metabolic functions related to C fixation and degradation, and N metabolism. Both processes were significantly abundant in spring, while C degradation gene abundances increased from summer to autumn, corresponding to increased litterfall and decomposition. These results provide insight into specific tree species-mediated mechanisms that structure belowground bacterial communities.



## 2.2 Introduction

Forests offer many ecosystem services of ecological and economic significance (Cardenas et al., 2015). They provide a diverse, multi-layered habitat for most terrestrial plants and animals, and resources for humans ranging from timber to recreational facilities (Wood et al., 2017). The high biodiversity of forests makes them important drivers of primary productivity (approximately 50%) and a major global sink for carbon (approximately 45% of global stocks) (Bonan, 2008). It is important to establish sustainable management practices, (Parviainen and Frank, 2003), to maintain high productivity (Cardenas et al., 2015; Schulze, 2017).

Soil acts as an essential biological matrix in which microbial communities execute key ecosystem functions, including biogeochemical cycling through decomposition and mineralization, processes mediated by prokaryotes and fungi (Lladó et al., 2017; Uroz et al., 2016). Factors shaping soil microbial community structure are crucial for predicting how bacterial-mediated processes drive ecosystem responses to environmental change (Nemergut et al., 2014).

Soil bacterial communities are shaped by several edaphic factors, including soil texture and chemistry, and biotic factors such as plant roots, aboveground litter and other decomposing organic matter (Thoms et al., 2010). Management regimes, which range from intensely managed agricultural land to undisturbed forested environments, have demonstrated the high impact of human-driven activities in shaping soil bacterial community structure and function (Mendes et al., 2014; Schneider et al., 2015b; Veresoglou et al., 2015).

Plant species significantly alter bacterial diversity (Schlatter et al., 2015). In forests, trees exert a stronger influence on the soil than other perennial vegetation due to their longevity. Over time, changing features such as forest canopy, root biomass and exudates, as well as oxygen and water consumption change soil temperature, chemistry, porosity and soil moisture (Augusto et al., 2002, 2015), which in turn shape belowground bacterial communities. The extent to which soil bacteria are shaped by trees depends on the tree species and the stand type and concentration (mono species or mixed species stands), in addition to macroflora and micro/macrofauna associated with the trees (Klimek et al., 2016; Uroz et al., 2016).

Only a limited number of studies are available for forest tree species effects on bacterial community structure and function, compared to grassland or agricultural ecosystems. Nacke et al. (2011) compared mono-specific stands of beech and spruce, demonstrating a strong effect of tree species and pH, with higher bacterial diversity in beech stands. Pfeiffer et al. (2013) conducted mesocosm experiments simulating the forest environment of the Hainich national park with single and mixed species of five different trees, showing that beech and ash had a strong effect on soil bacterial community richness and evenness. A pan-European study on land-use intensity and microbial co-occurrence indicated that forest soil communities form far richer networks than grassland or farmed soils (Creamer et al., 2016). Furthermore, studies focused on forest stand influences on both entire and active bacterial community carry out their studies in soils with wide spatial heterogeneity (Chodak et al., 2016; Klimek et al., 2016; Siles and Margesin, 2017). Therefore, the impact of trees is not fully reported. Despite frequent referencing of the abiotic environment, particularly pH, as a strong driver

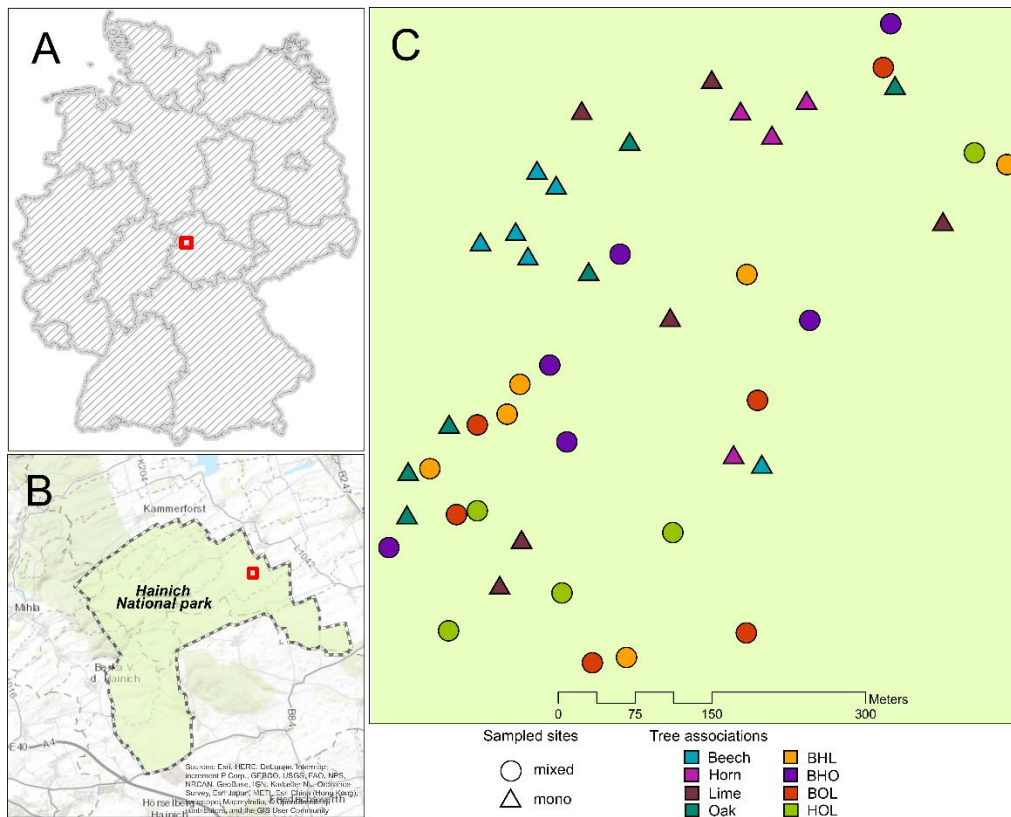
of microbial diversity (Kaiser et al., 2016; Lauber et al., 2008), nevertheless, tree species can exhibit a stronger impact on community structure than the soil environment (Bonito et al., 2014).

The aim of this study was to investigate the influence of tree species on soil bacterial communities in a broad-leaved temperate forest, the Hainich National Park (Figure 1A and B). The park is the largest unmanaged deciduous, broad-leaved forest ecosystem in central Germany, situated near the village of Weberstedt (350 m a.s.l; 51°05'37.0 N, 10°30'10.6 E) (Mölder et al., 2006; Schmidt et al., 2015). The limestone bedrock is covered by Stagnic Luvisol soil, which contains high silt (approximately 74%) and low sand contents (less than 5%) (Guckland et al., 2009). The region has a typically semi-humid climate with mean annual temperature of 7.7 °C, mean annual precipitation of approximately 590 mm/year. Former management activities included coppicing with standards and selective logging. Since the 1960s, the area underwent very minor management, which were discontinued after its declaration as a national park in 1997 (Kubisch et al., 2015). The environment is therefore described as a natural ecosystem. Core tree species in the mixed-forest ecosystem include the European beech (*Fagus sylvatica*), small-leaved and large-leaved lime (*Tilia cordata* Mill. and *T. platyphyllos* Scop., respectively), oak (*Quercus petraea* and *Q. robur*), and hornbeam (*Carpinus betulus*) (Mölder et al., 2006). We hypothesized that i) tree species identity drives bacterial community diversity and structure at entire and potentially active bacterial community level, ii) metabolic functions are also driven by tree species. We predicted that pure (mono species) and mixed stands are responsible for changes in soil physicochemical properties and that bacterial communities in different stands are shaped by season but to a lesser extent than by tree species. In addition, functional profiles of the soil bacterial communities with respect to tree species were predicted and compared. We hypothesize that bacterial metabolic function follows stand-specific composition.

## 2.3 Methods and materials

### 2.3.1 Sample site description

The study was conducted over a 25-ha area in the Hainich National Park located in Thuringia, Germany (Figure 2.1A and B). Recorded values for mean annual temperature and precipitation for the study year 2012 were 9.5 °C and 470 mm/year, respectively. Data were retrieved from the nearby weather station at Weberstedt. The sample sites were selected based on the following criteria: a) each stand should have a similar surrounding tree species composition displayed, b) trees must have been actively growing and have a well-defined canopy and c) homogeneity in soil parameters including colour, texture, drainage and slope of sampling area at initial sampling. Individual stands consisted of 4 to 8 trees, with an area between 68 – 313 m<sup>2</sup>. Mono-species stands (hereafter referred to as mono stands) comprised beech, hornbeam, lime and oak, while mixed-species stands (hereafter referred to as mixed stands) consisted of three out of the four tree species: BHL (beech, hornbeam, lime), BHO (beech, hornbeam, oak), BOL (beech, oak, lime) and HOL (hornbeam, oak, lime). Each stand was replicated 6 times, resulting in a total of 48 plots (Figure 2.1C).



**Figure 2. 1. Area of study in the Hainich national park. A) Map of Germany showing location of the Hainich national park in Thuringia, Germany (51°05'37.0 N, 10°30'10.6 E). B) Boundary of the Hainich national park. C) Plan view of the sampling area showing distribution of mono stands and mixed stands used for sample collection. Abbreviations stand for the following tree species associations in mixed stands: beech-hornbeam-lime (BHL), beech-oak-hornbeam (BHO), beech-oak-lime (BOL) and hornbeam-oak-lime (HOL).**

### 2.3.2 Sampling and environmental nucleic acid isolation

Soil samples were collected in the spring (April), summer (July) and autumn (September) of 2012. Two soil cores (10 cm in diameter, 5 cm depth) from the A horizon (topsoil) were extracted from randomly selected polar sites within a plot and pooled to generate a composite sample. To preserve environmental DNA and RNA, samples were stored at  $-20\text{ }^{\circ}\text{C}$  and  $-80\text{ }^{\circ}\text{C}$ , respectively. Determinations of soil physicochemical properties such as nitrogen (N), carbon (C), phosphorus (P) content, carbon/nitrogen (C/N) ratio and moisture content were performed and described by Schmidt et al. (2015) and are presented in Supplementary Information (Figure S2.1).

Extraction of environmental DNA and RNA was, performed with the MoBio Powersoil DNA isolation kit and RNA Powersoil Total RNA isolation kit, respectively, as recommended by the manufacturer (MoBio Laboratories, Carlsbad, USA). Minor modifications were performed in the extraction of total RNA as follows: after eluting bound RNA from the RNA Capture column, a 2.5-fold volume of ethanol and 0.1-fold volume of 3 M sodium acetate (pH 5.2) were added to the RNA, vortexed and stored at  $-80\text{ }^{\circ}\text{C}$  until required. To recover RNA solutions were centrifuged at  $10,000 \times g$  for 1 h ( $4\text{ }^{\circ}\text{C}$ ), subsequent ethanol/sodium acetate supernatants were discarded, and RNA pellets were dried before suspension in 100  $\mu\text{l}$  distilled water. Residual DNA was removed from extracted RNA by using TURBO DNA-Free kit (Ambion Applied Biosystems, Darmstadt, Germany), and confirmed by 16S rRNA gene PCR according to Wemheuer et al. (2012). Resulting RNA was concentrated with the Rneasy MinElute cleanup kit (Qiagen GmbH, Hilden, Germany). Quantification of DNA and RNA concentrations was performed with the NanoDrop ND-1000 UV-Vis spectrophotometer by following the instructions of the manufacturer (Peqlab Biotechnologie GmbH, Erlangen, Germany). Purified RNA (approximately 300 – 400 ng) was converted to cDNA using SuperScript III reverse transcriptase (Invitrogen, Karlsruhe, Germany) and the bacterial reverse primer V5rev\_B (5'- CTATGCGCCTTGCCAGCCCGCTCAG-MID-CCGTCAATTCMITTTGAGT-3') (Wang and Qian, 2009).

Environmental DNA and cDNA were used as template to amplify the V3-V5 regions of the 16S rRNA gene by PCR. The 50  $\mu\text{l}$  PCR reaction mixture contained 25 ng of environmental DNA or cDNA as template, 10  $\mu\text{l}$  of 5X Phusion GC buffer, 0.2  $\mu\text{M}$  of each of the four deoxynucleoside triphosphates (dNTPs), 0.4  $\mu\text{M}$  of each primer, 2.5  $\mu\text{l}$  DMSO, 0.15  $\mu\text{l}$   $\text{MgCl}_2$  and 1 U of Phusion high-fidelity DNA polymerase. The V3-V5 region was amplified with the following set of primers comprising the Roche 454 pyrosequencing adaptors (underlined), a key (TCAG), a unique 10-bp multiplex identifier (MID) and template-specific sequence per sample: the forward primer V3for\_B (5'- CGTATCGCCTCCCTCGCGCCATCAG-MID-TACGGRAGGCAGCAG-3') and reverse primer V5rev\_B (5'- CTATGCGCCTTGCCAGCCCGCTCAG-MID-CCGTCAATTCMITTTGAGT-3') from Liu et al. (2007) and Wang and Qian, (2009), respectively. All PCR reactions were performed in triplicate employing the following thermal cycling parameters for amplification: an initial denaturation at  $98\text{ }^{\circ}\text{C}$  for 5 min, followed by 25 cycles of denaturation at  $98\text{ }^{\circ}\text{C}$  for 45 s, annealing at  $65\text{ }^{\circ}\text{C}$  for 45 s and extension at  $72\text{ }^{\circ}\text{C}$  for 30 s, and a final extension at  $72\text{ }^{\circ}\text{C}$  for 5 min. Resulting amplicons were analysed by gel electrophoresis for size, pooled and purified with the Qiagen Qiaquick gel extraction kit as recommended by the manufacturer (Qiagen GmbH, Hilden,

Germany). Quantification of amplicons was determined by using the Quant-iT dsDNA BR assay kit and Qubit fluorometer as recommended by the manufacturer (Invitrogen GmbH, Karlsruhe, Germany). Amplicon sequencing was performed by the Göttingen Genomics Laboratory with the 454 GS-FLX pyrosequencer and titanium chemistry as recommended by the manufacturer (Roche, Mannheim, Germany).

### 2.3.3 Processing of 16S rRNA gene sequence data and statistical analyses

The resulting 16S rRNA gene and transcript sequences were processed and analysed with the QIIME (1.9.1) software package (Caporaso et al., 2010) by employing the scheme outlined by Schneider et al. (2015a). This involved removal of sequences shorter than 300 bp, containing unresolved nucleotides, exhibiting a low average quality score ( $< 25$ ) or harbouring long homopolymers ( $> 8$  bp). Forward and reverse primer sequences were removed with the *split\_libraries.py* script. Remaining reverse primer sequences, pyrosequencing noise and chimeric sequences, were removed with cutadapt (Martin, 2011), Acacia (Bragg et al., 2012) and UCHIME (Edgar et al., 2011), respectively. Operational taxonomic units (OTUs) were assigned at 97 % genetic similarity (species-level), by employing the UCLUST algorithm (Edgar, 2010) and the *pick\_open\_reference\_otus.py* script, against the Silva database. Taxonomic classification of subsequent OTUs was performed by using similarity searches with BLAST (Altschul et al., 1990) against the Silva SSU database release 128 (Quast et al., 2013). The *filter\_otu\_table.py* script was used to remove singletons, chloroplast sequences, extrinsic domain OTUs and unclassified OTUs.

For all statistical tests, a  $p$ -value of  $\leq 0.05$  was regarded as significant. Alpha and beta diversity indices and rarefaction curves were calculated with QIIME by using *alpha\_rarefaction.py* employing the same level of surveying effort (13,000 randomly selected bacterial sequences per sample). Additional analyses were performed in R (R Development Core Team, 2017). Non-metric multidimensional scaling (NMDS) was performed with the ‘vegan’ package (Oksanen et al., 2015), employing weighted UniFrac distance matrices to calculate beta diversity (Lozupone et al., 2011). Environmental parameters were fitted on the NMDS with the *envfit* function of the ‘vegan’ package in R (Gergs and Rothhaupt, 2015). A combination of analysis of similarities (ANOSIM) and permutational multivariate analysis of variance (PERMANOVA), implemented in QIIME, was used to improve robustness of multivariate analyses on the effect of tree stands on bacterial community using weighted UniFrac distance matrices (Hartmann et al., 2015). Association networks between tree species and OTUs were determined by mapping significant point biserial correlation values calculated by the ‘indicpecies’ package in R (De Cáceres, 2013). Subsequent network visualisations for taxa/tree stand associations were generated with Cytoscape v3.5 by using the ‘edge-weighted spring embedded layout’ algorithm, whereby network edges were weighted by association value (Cline et al., 2007; Shannon et al., 2003). Prediction of functional pathways and corresponding enzymes was performed with Tax4Fun (Abhauer et al., 2015). Subsequent NMDS ordinations and boxplots were calculated with the *vegdist* function in ‘vegan’ and ‘ggplot2’ package, respectively (Wickham, 2009).

#### 2.3.4 Accession numbers

The 16S rRNA gene and transcript sequences were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the accession number PRJNA342582.

## 2.4 Results and Discussion

### 2.4.1 Edaphic properties reflect stand-related variations

Soil properties retrieved from Schmidt et al. (2015) reveal distinct patterns with respect to mono and mixed stands. Edaphic parameters displayed significantly different stand-specific variations ( $p < 0.05$ ), except for moisture content (Figure S2.1, Supplementary material). Among mono stands, lime exhibited the highest mean pH ( $5.9 \pm 0.6$ ). Beech and oak shared a similarly low pH environment ( $4.6 \pm 0.3$  and  $4.5 \pm 0.5$ , respectively). Lime-associated stands displayed pH values of  $5.4 \pm 0.4$  (BOL) and  $5.2 \pm 0.5$  (HOL), whereas beech and hornbeam-associated mixed stands ranged from  $5 \pm 0.43$  (BHL) to  $4.8 \pm 0.8$  (BHO).

The highest carbon to nitrogen (C/N) ratio was recorded in beech ( $15.2 \pm 0.7$ ) and oak ( $16.1 \pm 1.3$ ) mono stands and the lowest in lime mono stands ( $12.4 \pm 0.6$ ). Corresponding beech and oak mixed stands (BHO and BOL) also had higher C/N ratios ( $15.5 \pm 3.5$  and  $15 \pm 4.1$  respectively) compared to lime-hornbeam mixed stands, BHL ( $13.2 \pm 0.7$ ) and HOL ( $13.1 \pm 0.2$ ). Plant available phosphorus (P) was consistently higher in soils of hornbeam and lime mono stands ( $585 \pm 240$  and  $536 \pm 109$  mg/kg, respectively), while beech and oak had significantly lower P content ( $340 \pm 41$  and  $335 \pm 58$  mg/kg, respectively) than hornbeam and lime mono stands. Furthermore, hornbeam and lime-associated stands (BHL and HOL) exhibited significantly higher P content than beech and oak-associated stands (BHO and BOL).

Carbon/nitrogen ratios between 8 and 16 indicate complete breakdown of organic material and, consequently, higher nutrient availability while phosphorus is an essential component for the growth and maturity of plants (Lang et al., 2016; Lauber et al., 2008). Therefore C/N and P content are used as indicators of soil productivity. Low C/N ratios, high available N and pH, as observed in lime mono stands, promote tree productivity (Figure S2.1). In contrast, beech and oak mono stands exhibited higher C/N ratios, low soil available nitrogen and low pH. One explanation is that lime litterfall has a higher decomposition rate compared to beech litter and contributed to the observed effect on soil chemistry (Jacob et al., 2009; Scheibe et al., 2015). The results present a clear impact of tree species identity on soil characteristics. This is confirmed in a study, also from the Hainich national park, that demonstrated similar soil acidification in beech stands in comparison to ash and lime species (Langenbruch et al., 2012).

Generally, traits observed in mono stands were also recorded in associated mixed stands. Beech and oak mixed stands (BHO and BOL) exhibited an effect similar to individual beech and oak mono stands, whereas hornbeam and lime mixed stands (BHL and HOL) show traits similar to hornbeam and lime mono stands. This indicates that tree identity and stand type create distinct microenvironments. Mixed stands produce intermediate effects in stand productivity compared to mono stands. The ability of individual species in mixed stands to promote stand growth (facilitation) or compete for nutrient resources (competition) is species dependent (Cavard et al., 2011; Jacob et al., 2013; Schmidt et al., 2015). For example, lime grows better in mono stands, but in mixed stands it improves overall stand productivity by contributing litter high in pH and low in C/N and lignin (De Jaegere et al., 2016; Schmidt et al., 2015). Nevertheless, the results consistently showed a more acidic

environment of beech- or oak-dominated stands, compared to lime- or hornbeam-dominated stands, as previously reported (Berger and Berger, 2014; Brunet et al., 1997; Falkengren-Grerup et al., 1998; Salehi et al., 2007).

#### 2.4.2 Stand-specific patterns in bacterial richness and diversity

Soil DNA and RNA were collected from a temperate forest soil environment. Subsequently, soil bacterial community structure was determined based on 16S rRNA gene and transcript analyses. After processing, quality-filtering and taxonomic clustering at 97% similarity, 40,385 for entire (DNA-based) and 52,277 OTUs for active (RNA-based) bacterial communities were recovered for analyses (Table S2.1). The higher number of bacterial OTUs in the active community is in contrast to some other studies of bacterial community diversity in forest (Baldrian et al., 2012; Romanowicz et al., 2016) and grassland soils (Herzog et al., 2015). However, Gill et al. (2017) reported higher RNA-based OTUs from urban soils, similar to our results. We attribute the different abundances in total and active communities to the detection of very low-abundant but active rare taxa at RNA level.

Diversity indices showed several consistent significant patterns in mono and mixed species tree stands with respect to species richness and evenness, as indicated by Shannon ( $H'$ ), Chao 1 and phylogenetic diversity (PD) ( $p < 0.1$ ) (Figure S2.2). Rarefaction curves were close to saturation and, in addition, Good's coverage estimator across all stands remained above 75 %, indicating that the sampling effort recovered most of the observed bacterial diversity.

Amongst mono stands, lime showed the highest bacterial community richness ( $H'$   $10.5 \pm 0.3$ ; PD  $169 \pm 33$ ), in both total and active communities. Beech mono stands were similar to hornbeam mono stands, and the corresponding mixed stand BHO ( $H'$   $10.1 \pm 0.2$ ; PD  $159 \pm 24$ ). Lime-associated mixed stands (BOL and HOL) were very similar to lime mono stands ( $H'$   $10.4 \pm 0.3$ ; PD  $170 \pm 32$ ). Oak mono stands showed significantly lower bacterial richness and diversity ( $H'$   $9.8 \pm 0.4$ ; PD  $140 \pm 27$ ) compared to other mono stands. We assume that the bacterial richness and diversity differences are due to tree species-specific effects on the bacterial community. Changes in bulk soil chemistry, i.e. pH differences, drive bacterial community structure and diversity (Jeanbille et al., 2016). Chemical changes exerted on soil by trees subsequently impact belowground bacterial communities (Thoms et al., 2010). In our study, this effect resulted in lime mono stands having the highest bacterial richness and diversity, while oak had the lowest. Favourable physiological conditions (high pH, low C/N ratio, high exchangeable P) in lime and hornbeam mono stands promote bacterial diversity compared to the acidic environment of beech and oak mono stands. Studies in a mixed deciduous forest showed similar bacterial community response to mono stands and 2-species mixed stands of lime, oak and alder (*Alnus* sp.), birch (*Betula* sp.), larch (*Larix* sp.) and spruce (*Picea* sp.) (Šnajdr et al., 2013; Urbanová et al., 2015).

#### 2.4.3 Tree species effect of bacterial community structure

Multivariate analyses confirmed that bacterial communities are distinct and can be delineated along tree species (tree species,  $p = 0.001$ ,  $r^2 = 0.506$ ) (Figure 2.2A and Table S2.2). Communities of each tree species group similarly in the total and active community (template,  $p = 0.001$ ,  $r^2 = 0.439$ ). A pattern in which communities in mono stands are separate from those in mixed stand was also detected (stand type,  $p = 0.043$ ,  $r^2 = 0.068$ ). An exception was for bacterial communities in hornbeam

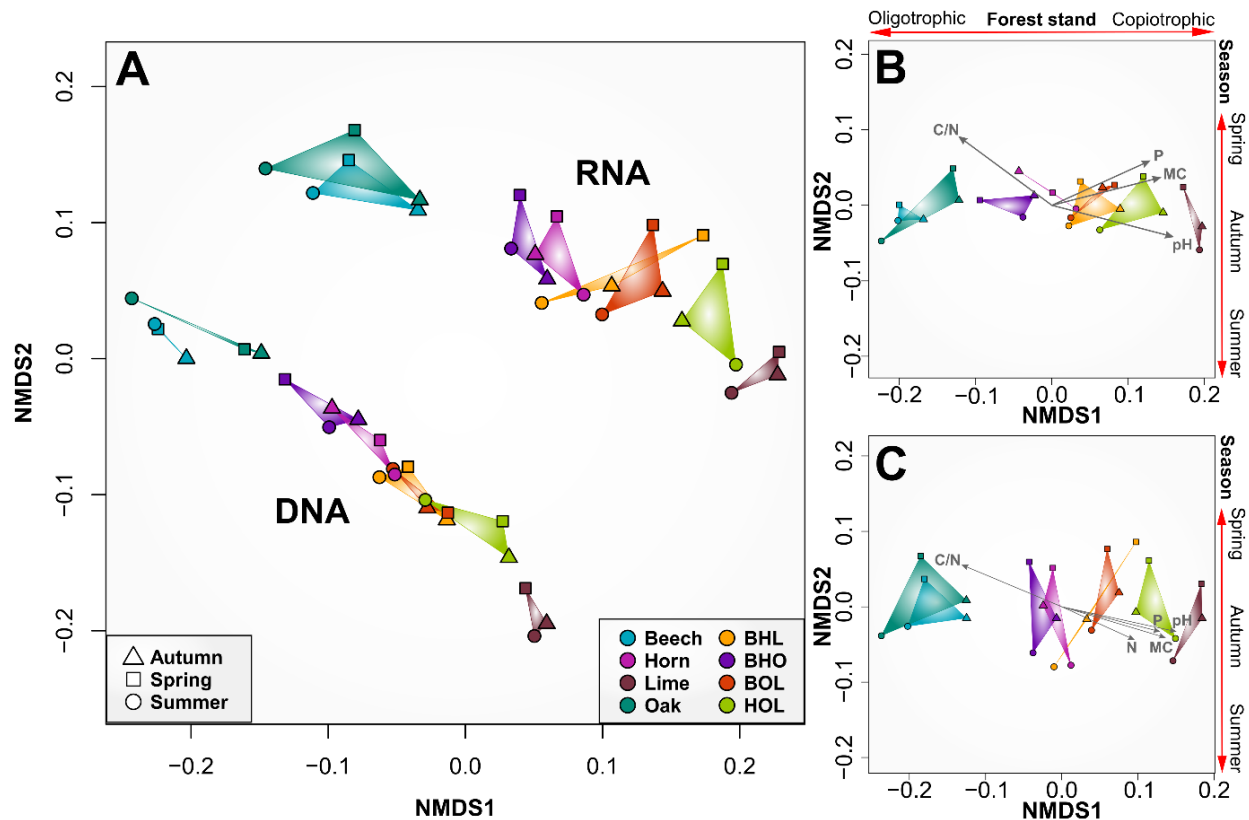


mono stands, which grouped with mixed stands. Results of the NMDS were supported by ANOSIM and PERMANOVA tests, which indicate that tree species ( $p = 0.001$ , ANOSIM and PERMANOVA) is a strong driver of bacterial community structure at total and active community level (Table S2.2). Stand type as a driver of community structure was detected in the potentially active community ( $p = 0.002$  and  $p = 0.044$ , ANOSIM and PERMANOVA, respectively). Seasonal effects in our data were less pronounced (season,  $p = 0.858$ ,  $r^2 = 0.0142$ ). The results supported the trends observed during alpha diversity analysis, in which bacterial richness and evenness follow stand-specific variation, and confirm our hypothesis that trees drive bacterial composition.

ANOSIM and PERMANOVA analyses indicate that season had no significant impact on composition of the total community ( $p = 0.464$  and  $0.529$ , respectively), but ANOSIM showed a significant seasonal impact on active communities ( $p = 0.038$ ). Although temporal seasonality is an important factor in temperate deciduous forests, which influences plant phenology, and subsequent root exudation and nutrient uptake processes (Oh et al., 2012), it does not have a strong effect on bacterial beta diversity. An experimental warming experiment of a temperate forest soil only showed a shift in bacterial community composition after 20 years, and only in the organic horizons (DeAngelis et al., 2015). An explanation is that small changes in bacterial community composition across seasons are driven instead by ‘conditionally rare taxa’ that periodically increase during favourable environmental changes (Shade et al., 2014). As environmental RNA almost exclusively originates from viable organisms (Moran et al., 2013; Pochon et al., 2017), the active community is more sensitive to environmental disturbance than the total community (Pochon et al., 2017).

Edaphic parameters also contributed significantly to shape the bacterial community ( $p = 0.001$ ) (Figure 2.3B and C, Table S2.2). Soil pH, P, N and soil moisture showed greater impact on bacterial communities in lime and hornbeam mono stands, and in the corresponding BHL, BOL and HOL mixed stands. However, bacterial communities in beech and oak mono stands, and the BHO mixed stand responded to an increase in C/N ratio. Nutrient availability in soil has been linked to soil bacterial structure (Bergkemper et al., 2016). Our results show that the nutrient-rich soils of lime and hornbeam mono stands possess bacterial structures, which are distinct from that in comparatively nutrient-reduced soils in beech and oak mono stands.

Rather than tree species richness (mixed stands), tree species identity (mono stands) contributes to bacterial diversity and structure. A study in a mixed deciduous forest demonstrated a similar correlation of microbial diversity with tree species identity, while species richness influenced microbial composition (Khlifa et al., 2017). Plant-specific traits such as fine root biomass and density contribute to nutrient availability in soil by producing i.e. root exudates (Thoms et al., 2010). Beech stands were shown to have higher fine root biomass than hornbeam or lime (Jacob et al., 2013), which could have a stronger impact on bacterial communities than other mono stands. However, beech showed no overyielding of root biomass in mixed stands (Jacob et al., 2013; Langenbruch et al., 2012). This explains how the impact of mixed stands on bacterial communities is often similar to the constituent mono stands, as little competition exists in fine roots. A significant change in soil bacterial community was observed with increasing horizontal distance from trees and demonstrated shifts in nutrient availability, such as decreasing concentrations of organic nitrogen compounds (Nacke et al., 2016).



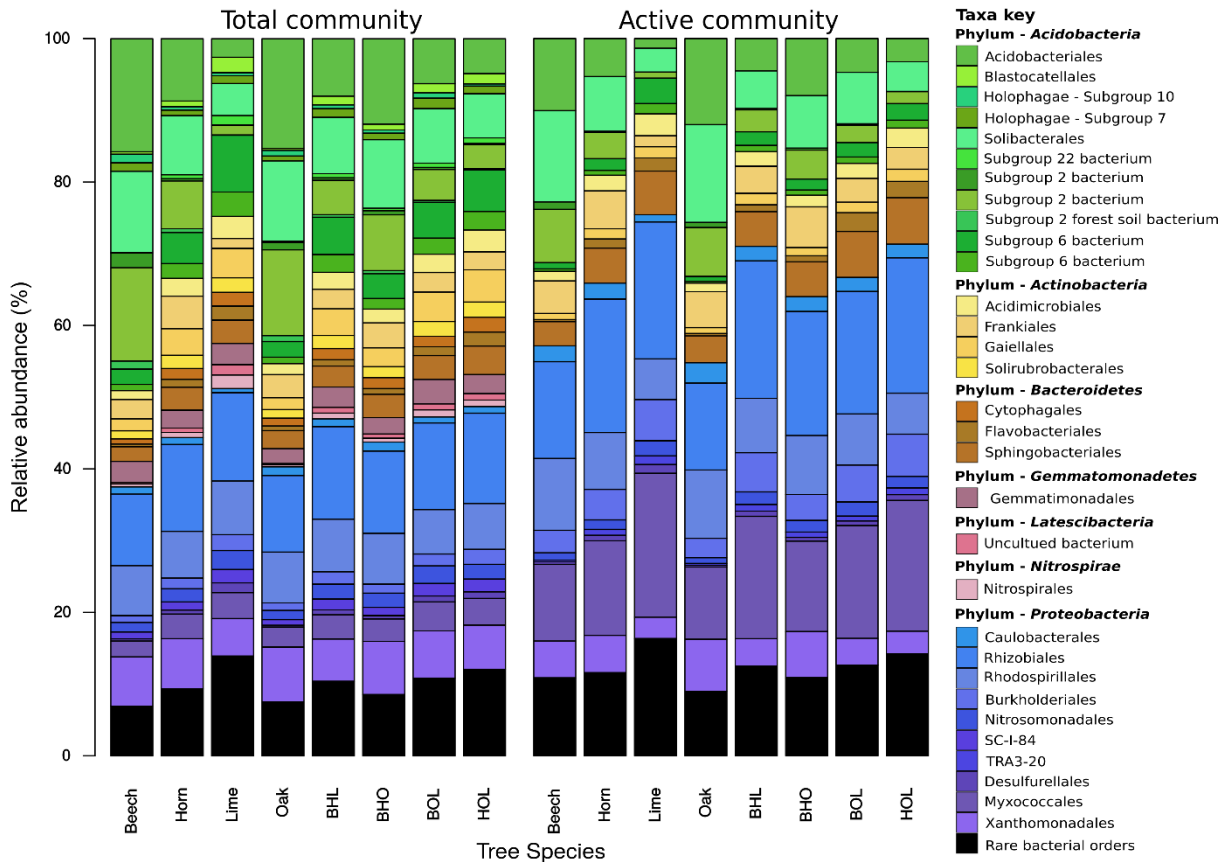
**Figure 2. 2. Non-metric multidimensional scaling (NMDS) ordination of soil bacterial communities in mono and mixed stands of a temperate deciduous forest. Ordination employed weighted UniFrac distance matrices of OTUs. Samples are grouped according to season ( $n = 3$ ) and represent data collected from 48 mono and mixed-species stands in the Hainich national park. A) NMDS plot for total and active community. B) and C) NMDS plots for total and active community, respectively, showing the goodness of fit for soil environmental parameters. Arrows outside the plot area have no statistical significance and are only a visual aid of observed trends in ordination. Carbon/nitrogen ratio (C/N), phosphorus content (P), nitrogen content (N), soil moisture (MC).**

#### 2.4.4 General patterns in bacterial community composition and structure

The entire dataset comprised 40 phyla, 155 classes, 385 orders, 704 families and 1,552 genera. At phylum level, the total community comprised 40 phyla and the active community 35. The higher diversity at phylum at DNA level compared to RNA level is consistent with results from other studies (Angel et al., 2013; Baldrian et al., 2012; Herzog et al., 2015).

Across the entire DNA-based and RNA-based datasets the total and active bacterial community did not change considerably at phylum level. Dominant phyla (relative abundance  $>1\%$ ) comprise across the entire datasets *Acidobacteria*, *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, *Gemmatimonadetes*, *Chloroflexi*, *Latescimicrobia* and *Nitrospirae* (Figure 2.3). *Acidobacteria* dominated the total bacterial community (37.2% at DNA level and 23.2% at RNA level) and *Proteobacteria* the active bacterial community 36.8% at DNA level and 56% at RNA level). Soils generally have a few dominating phyla, mainly *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Proteobacteria* and *Verrucomicrobia* (Fierer et al., 2012). These phyla are characteristic constituents of bacterial communities derived from different cultivated and

natural forest soils (Ferrenberg et al., 2013; Schneider et al., 2015a; Shange et al., 2012; Shen et al., 2014; Shi et al., 2011). At total community level, the five most abundant orders were *Rhizobiales* (11.8 %), *Acidobacteriales* (9.2 %), *Solibacterales* (8.3 %), *Rhodospirillales* (6.9 %) and an uncultured order of acidobacterial subgroup 2 (6.7 %). In the active community, the dominant orders were *Rhizobiales* (17 %), *Myxococcales* (14.7 %), *Rhodospirillales* (7.7 %), *Solibacterales* (7.6 %) and *Acidobacteriales* (6.1 %).



**Figure 2. 3. Mean relative abundance of soil bacterial communities in mono and mixed stands. Results show order level composition at order level, where n = 3 seasons (spring, summer and autumn). Total and potentially active taxa were inferred from environmental DNA and RNA, respectively. Taxa with a relative abundance less than 1 % across all stands were grouped into ‘rare bacterial orders’.**

Bacterial community composition was stand-specific (Figure S2.3) and showed dependence on pH and nutrient availability (Figure S2.1). Most members of *Alpha*- and *Betaproteobacteria*, *Actinobacteria* and *Bacteroidetes* are known as copiotrophic and are adapted to nutrient-rich environments, as found in lime stands. In contrast, members of *Acidobacteria* are recognized as oligotrophs and adapted to nutrient-limiting conditions (Fierer et al., 2012; Koyama et al., 2014; Shange et al., 2012). Beech is frequently associated with acidic soils (Colin et al., 2017; Jeanbille et al., 2016), which favour acidobacterial groups (Jones et al., 2009; Kielak et al., 2016b; Lladó et al., 2016; Uroz et al., 2011).

*Proteobacteria* were evenly distributed across all sites, with slight differences between mono stands and mixed stands at total community level. *Rhizobiales* belonging to the *Alphaproteobacteria* were more abundant in lime and hornbeam mono stands at total community and active community level than in beech and oak mono stands (Figure 2.3 and Table S2.3). A similar trend was observed in the mixed stands BHL and HOL compared BHO and BOL. At genus level, we recorded that these differences were due to high abundances of unidentified members of the *Xanthobacteraceae*, *Bradyrhizobium*, *Rhizomicrobium* and *Variibacter*. *Rhizobiales* are broadly associated with nitrogen fixation, plant pathogenicity and organic matter decomposition (Carvalho et al., 2010). The *Rhodospirillales* order, which consisted of mostly uncultured groups at genus level, showed high abundance in soils of beech and lime mono stands. This is explained by the presence of the *Acetobacteraceae* family, (up to 4.7 % at DNA and RNA level) in beech mono stands and uncultured *Rhodospirillum* DA 111 in beech and lime mono stands (11.3 %). *Acetobacteraceae* contain some acidophilic genera of acetic acid bacteria (AAB) (Mamlouk and Gullo, 2013). Additionally, the *Reyranella* genus was enriched in lime mono stands, which is consistent with previous reports of forest soil communities (Felske et al., 1998; Kim et al., 2013). The presence of *Rhodospirillales* in both lime and beech mono stands, which affect the soil differently, illustrates the broad adaptation capacity of this phylogenetic group, ranging from aquatic to terrestrial acidic environments (Belova et al., 2009; Johnson et al., 2006; Kimoto et al., 2010).

*Betaproteobacteria* were represented by *Burkholderiales* (unidentified *Comamonadaceae* genus, *Variovorax* and *Rhizobacter*) and *Nitrosomonadales*, (unidentified *Nitrosomonadaceae* genus and *Nitrosospira*) (Figure S2.3 and Table S2.3). At total and active community level, members of these orders were more abundant (6.8 % and 9.5 %, respectively) in lime and hornbeam mono stands compared to beech and oak mono stands (3.8 % and 4.5 %, respectively). A similar distribution was also found in the mixed stands BHL and compared to mixed stands BHO and BOL. Both orders form part of the nitrogen-fixing bacterial community in forests soils, where they participate in symbiotic relationships with plants (Cherobaeva et al., 2011; Tkacz and Poole, 2015).

*Myxococcales* and *Desulfurellales* orders dominated within the *Deltaproteobacteria* (Figure 2.3). *Haliangium* was the most abundant genus within *Myxococcales* in the total community (1.3 %) and active soil community (6 %), followed by *Sorangium*. Both genera are commonly found in soils, but only members of *Sorangium* have been isolated from soil (Dawid, 2000; Fudou et al., 2002). Members of both genera exhibit a capacity for producing secondary metabolites with potential pharmaceutical use (Fudou et al., 2001; Li et al., 2014a). The *Desulfurellales* consisted mainly of *Desulfurellaceae* family members. *Desulfurellaceae* were abundant in lime and hornbeam mono stands compared to beech or oak mono stands at total and active community level (Table S2.3). *Desulfurellaceae* are obligate sulphur-metabolising thermophiles that contribute to the sulphur cycle (Flores et al., 2012; Wang et al., 2016). Their presence in forest soils is not widely reported but one study has reported *Desulfurellaceae* in farm soils (Wang et al., 2016).

The *Gammaproteobacteria* consisted primarily of *Xanthomonadales* (*Rhodanobacter*, *Acidibacter* and an unidentified genus), which includes agriculturally significant plant pathogens such as *Xanthomonas* (Naushad et al., 2015). Members of *Xanthomonadales*, including *Rhodanobacter*, have been reported to prefer environments with lower pH, and high C/N ratios, which promote efficient denitrification

(Green et al., 2012; Prakash et al., 2012; van den Heuvel et al., 2010). Our results are consistent with these findings as *Xanthomonadales* were enriched at total and active community in beech (6.9 and 5.1 %, respectively) and oak mono stands (7.7 and 7.3 %, respectively) compared to lime mono stands (5.2 and 3 %, respectively).

The composition of *Acidobacteria* varied considerably across forest stands, through several unidentified subgroups (Figure S2.3). *Acidobacteriales*, *Solibacterales* and subgroup 2 were more abundant in beech and oak mono stands compared to lime and hornbeam mono stands (Figure S2.3 and Table S2.3). In contrast, subgroup 6 showed greater abundance in lime and hornbeam mono stands (19 % at DNA level and 7.3 % at RNA level), compared to beech and hornbeam mono stands (6.3 % at DNA level and 2.2 % at RNA level). *Acidobacteria* broadly correlate negatively with pH and nutrient availability (Clivot et al., 2012; Kielak et al., 2016a; Miyashita, 2015; Sait et al., 2006). Reports also indicate that some groups, including subgroup 6, have copiotrophic lifestyles and are enriched in more neutral, nutrient-rich environments (Huber et al., 2016; Naether et al., 2012). This also is in accordance with our results. Differences in mixed stands were less striking. *Acidobacteriales*, *Solibacterales* and subgroup 2 in mixed stands exhibited similar trends observed in mono stands and were more abundant in beech and oak mixed stands BHO and BOL. Subgroup 6 showed higher abundance in lime and/hornbeam mixed stands BHL and HOL, in the total and active community. Consequently, the more acidic soil environment in beech and oak stands promote a high abundance of *Acidobacteria*.

The third-largest phylum *Actinobacteria* consists of physiologically diverse taxa, including plant pathogens and members capable of producing secondary metabolites (Barka et al., 2016; Ventura et al., 2007). *Actinobacteria*, was mainly represented by *Gaiellales*, *Frankiales*, *Acidimicrobiales* and *Solirubrobacterales* orders (Table S2.3). *Frankiales* abundance increased in the active community (4 %) compared to the total community (2.9 %). Many dominant genera within these orders were unclassified but we identified *Acidothermus* (2.5 % at DNA level and 4 % at RNA level) within *Frankiales*. *Acidothermus* showed lower abundance in lime mono stands (0.9 % at DNA level and 1.2 % at RNA level), compared to beech (2.5 % at DNA level and 4.5 % at RNA level), hornbeam (4 % at DNA level and 5.6 % at RNA level) and oak (3 % at DNA level and 5 % at RNA level). *Acidothermus* was recorded in high abundance (11.5 % at DNA level) from a natural forest by Kim et al. (2015). Additionally, the only cultured species (*Acidothermus cellulolyticus*) was isolated from an acidic hot spring which points to an acidophilic lifestyle for this genus, as shown in our study (Barabote et al., 2009; Mohagheghi et al., 1986).

Representatives of *Bacteroidetes* belonged predominantly to *Cytophagaceae* (*Cytophagales*), *Flavobacteraceae* (*Flavobacteriales*) and *Chitinophagaceae* (*Sphingobacteriales*) families. Additionally, all families showed higher abundance in lime and hornbeam mono stands, than beech and oak mono stands (Figure S2.3 and Table S2.3). We identified the genus *Flavobacterium* within *Flavobacteraceae*, which was more enriched at active community level. *Flavobacterium* was generally abundant in lime and hornbeam mono stands (2.8 % at DNA level and 2.9 % at RNA level) than in beech and oak mono stands (0.78 % at DNA level and 0.56 % at RNA level). *Flavobacterium* is a common soil inhabitant. Species of the genus have been reported in plant root associations and deadwood, and harbour broad physiological capabilities including lignolytic activity (Deshmukh and Sao, 2015; Dilly et al., 2000; Hoppe et al., 2015; Kolton

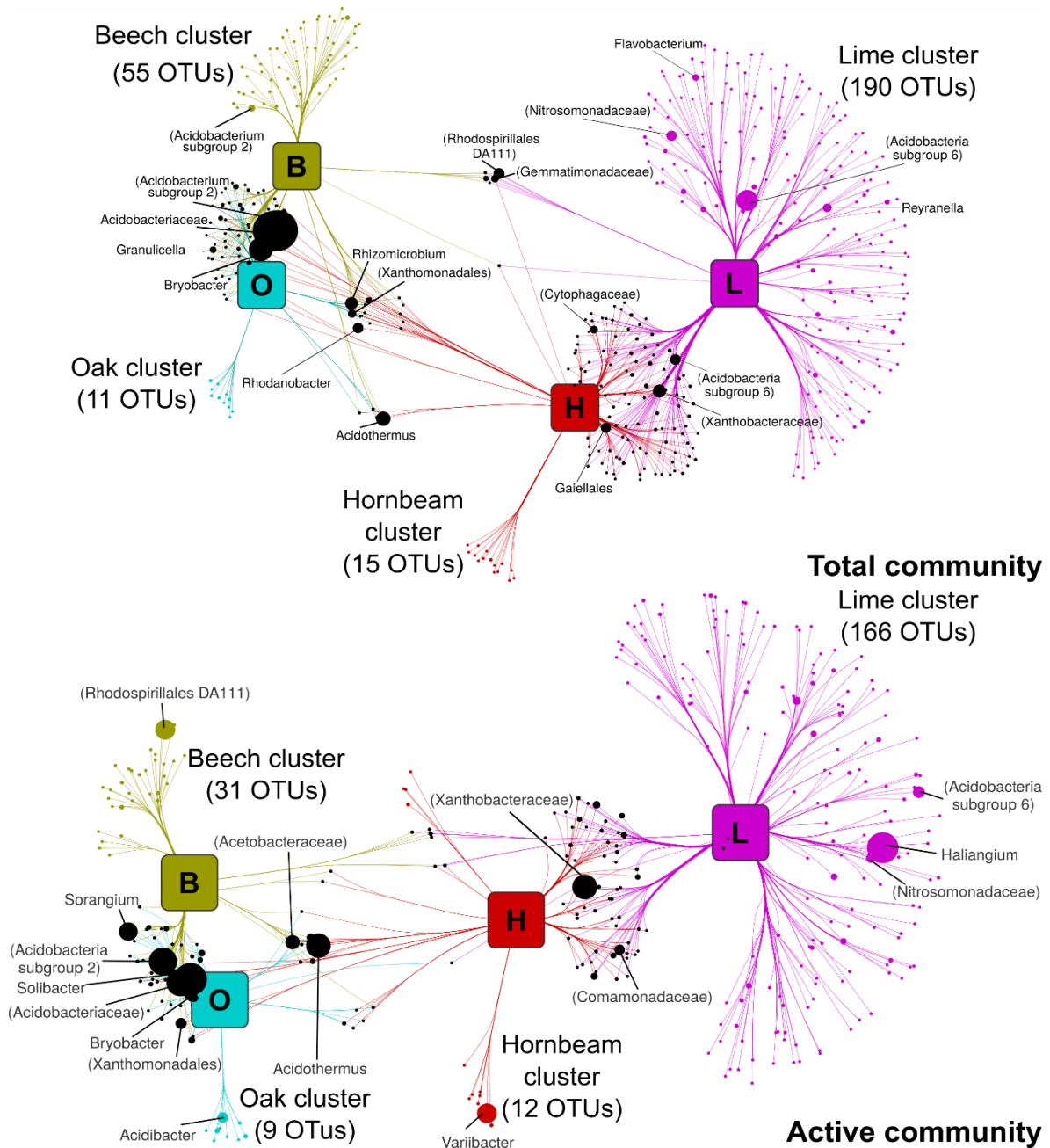
et al., 2013). In lime and hornbeam mono stands *Gemmatimonadaceae* (*Gemmatimonadetes*), *Nitrospirales* and *Latescibacteria* were more abundant than in other studied mono stands (Table S2.3). Farag et al., (2017) identified the presence of cellulosomes in *Latescibacteria*, which indicates a possible role in cellulose degradation.

#### 2.4.5 Taxa-habitat association patterns

Bipartite association networks provided insight into bacterial taxa that potentially drive the observed community structures across tree stands (Figure 2.4 and Figure S2.4). Networks were constructed from OTUs showing significant positive associations to specific stands or a combination of stands. The detected 466 and 348 indicator OTUs at total and active community level, respectively represent 1.2 and less than 1 % of all OTUs at total and active community level, respectively. The correlation-based network strongly mirrors the pattern of forest stands recorded during NMDS analysis (Figure 2.2). Shorter cross edges between beech and oak stands compared to hornbeam or lime stands indicate that communities in beech and oak stands are more closely associated with tree species than those associated with hornbeam or lime stands. No significant positive associations were observed between bacterial genera in oak and lime mono stands. Although genera in hornbeam and lime associate more closely, the strength of this association is not as strong as in beech and oak stands.

Unique clusters representing OTUs associated significantly with only one mono stand, accounted for 58 and 64 % of all network OTUs at total and active community level, respectively. The higher number of genera that define unique clusters in lime stands in the total (191) and active (169) community arises from the high bacterial diversity observed in lime stands, compared to hornbeam, beech and oak stands. Unique clusters in mixed stand were less common and instead, OTUs formed cross associations with two or more mixed stands (Figure S2.4). This provides further evidence that bacterial communities in mixed stands are composed of members associated with corresponding mono stands and illustrates the importance of tree species identity over tree species richness in shaping soil bacterial community.

Genera observed in the unique clusters belonged to previously described dominant phyla and orders. At entire community level, we identified unclassified OTUs from *Acidobacteria* subgroup 6, *Nitrosomonadaceae* and *Reyranella* in lime mono stands. Any overlap of significant bacterial cross-associations between beech and oak, and hornbeam and lime mono stands was provided by a few genera and comprise, among others, *Rhodanobacter*, *Rhizomicrobium*, *Acidothermus*, *Bryobacter*, *Granulicella*, *Gemmatimonadaceae* and *Cytophagaceae*. This pattern is similarly reflected in the active community but with fewer genera. Analysis of soil environments reveals shared taxonomic groups, but only a few genera participate in distinguishing one soil habitat from another (Hartmann et al., 2015; Rime et al., 2016). A study on microbial community conversion between organic and conventional farming showed that only 12 % of bacterial OTUs constituted the management-specific community, but 49 % of OTUs with significant association (Hartman et al., 2015). Similarly, our results follow this trend, as only a small fraction of OTUs define the difference between soil bacterial communities of different tree stands.



**Figure 2. 4. Association networks between soil bacterial communities (genus level) and mono specific stands. Source nodes (rounded squares) represent mono species tree stands and edges represent associations between stands and bacterial OTUs (circles, target nodes). Edges are coloured according to the source tree species and the length of edges is weighted according to association strength. Unique clusters, which associate with one tree species, consist of nodes coloured as the corresponding stand. Numbers of OTUs making up respective unique clusters are given in brackets. Black circles represent OTUs with significant cross association between two or more plots. Target node sizes represent mean relative abundance of OTUs across all mono plots. Data only represents OTUs that showed significant positive association with tree species ( $P \leq 0.05$ ). For ease of visualisation, edges were bundled together, with a stress value of 3.**

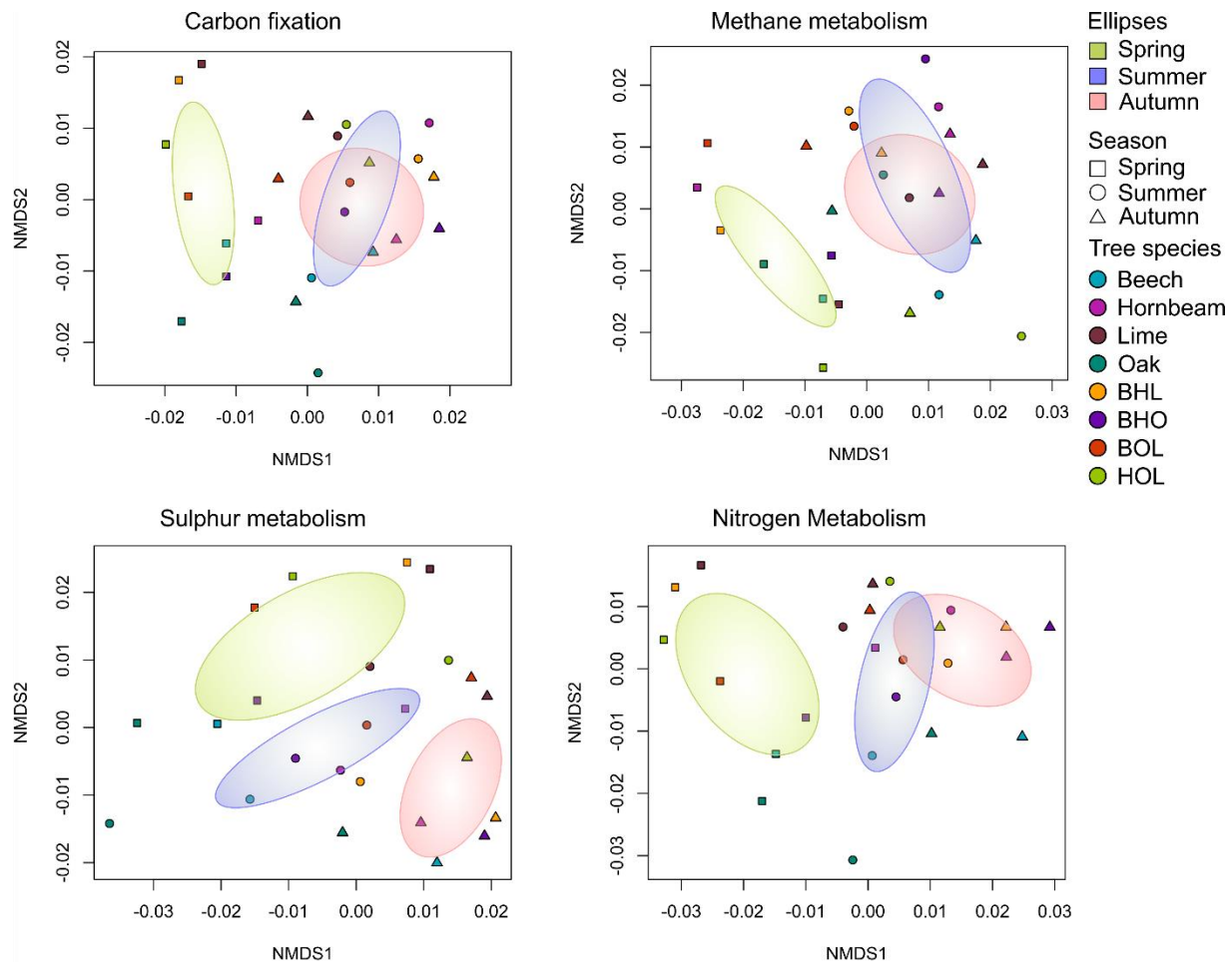
#### 2.4.6 Bacterial functional profiles across forest stands

Sequences recovered from environmental RNA were used to predict possible metabolic processes in forest soil. It is worth noting that rRNA abundance is only a qualitative index for activity but not direct measure of activity (Blazewicz et al., 2013). Therefore, the proposed functional profile does not necessarily reflect the direct activity of the studied organisms. Additionally, OTUs derived from unknown taxa limit functional predictions. However, it has been shown for bacterioplankton and soil bacterial communities, as well as for communities in other environments that 16S rRNA-derived functional prediction are in good agreement with those derived from direct sequencing of corresponding metagenomes and metatranscriptomes (ABhauer et al., 2015; Kaiser et al., 2016; Lüneberg et al., 2018; Wemheuer et al., 2017).

Predicted metabolic function did not follow the stand-specific trend of the taxonomic community but showed strong grouping according to season ( $p = 0.001$ ) (Figure 2.5 and Table S2.4). Only genes related to methane metabolism showed association with tree species ( $p = 0.033$ ) (Figure S2.5). This is in accordance with a report showing that methylotrophic bacteria respond to shifts in soil pH in a beech-dominated deciduous forest (Morawe et al., 2017). In general, genes associated with carbon and nitrogen metabolism in spring are distinct from summer and autumn, which group more closely. As a subset of carbon metabolism, gene function associated with methane metabolism also followed this trend. Genes for sulphur metabolism showed stronger separation between seasons. As light and temperature increase in spring (April), trees begin to form leaves and increase photosynthetic productivity, which peaks in summer (July) (Goodale et al., 2015). In autumn (October), light and temperature decline and trees generally respond through increased litterfall (Goodale et al., 2015; Žifčáková et al., 2016). However, most litter decomposition takes place during the summer (Sohng et al., 2014) and may continue into autumn. Thus, in belowground metabolic processes from summer extend into autumn give rise to similar bacterial functional profiles.

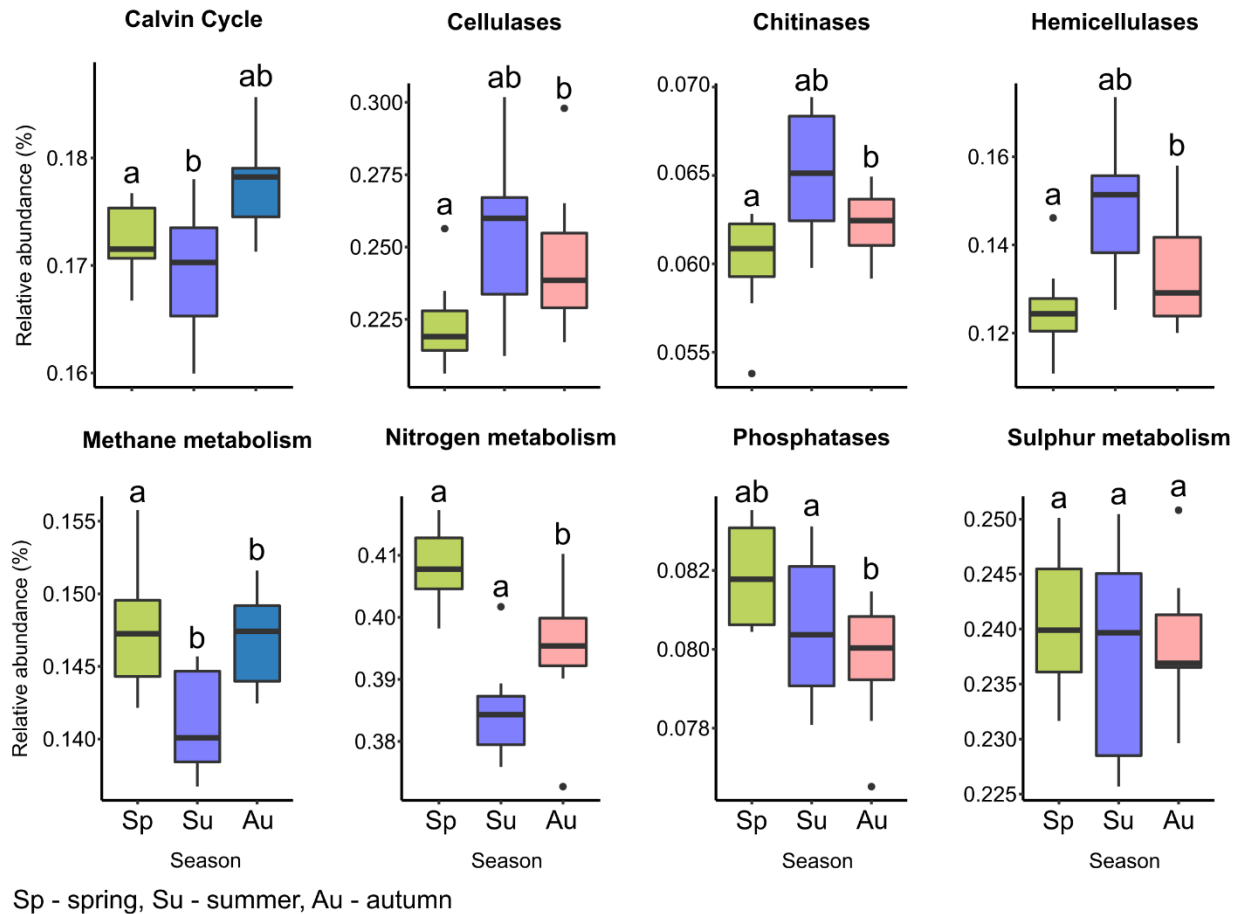
We identified several genes related to carbon (including methane) and nitrogen metabolism, key pathways in microbial biogeochemical cycling (Figure 2.6 and Table S2.5). Carbon turnover in soils is mediated by carbohydrate-active enzymes (CAZymes), which act on labile C compounds, recalcitrant cellulose or hemicellulose and fungal biomass (López-Mondéjar et al., 2016; Žifčáková et al., 2017). In our study, genes related to C fixation, including ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCo) showed significantly higher abundance in spring and autumn ( $p = 0.013$ ). However, the abundance of C degradation genes (cellulases, hemicellulases and chitinases) were more abundant in summer and autumn. Due to temperature dependence, CAZymes activity responds to seasonal changes and impacts C cycling (Žifčáková et al., 2017). The quality of soil C input from root exudates and litter varies seasonally (Siles and Margesin, 2017) and explains the changing abundance of C fixation and degradation genes. A similar increase in C degradation genes in autumn compared to spring was also reported by Siles and Margesin, (2017) in a mixed deciduous and coniferous forest ecosystem.





**Figure 2. 5. Non-metric multidimensional scaling ordination of functional genes in key biogeochemical processes. Ordination is based on Bray-Curtis distance matrices of functional genes (KEGG orthologs) predicted by Tax4Fun. Data points are summarized by season ( $n = 3$ ).**

The ability to utilise a wide range of carbon substrates for gaining metabolic energy is shared among several bacterial taxa, across phyla and enables several groups to inhabit the same environment with different nutrient niches (Lladó et al., 2017). Bacterial phyla comprise metabolically versatile genera that carry out the same general metabolic processes and are functional redundant (Burke et al., 2011). Therefore, the relative abundance of genes in an environment do not always correlate with metabolic activity of the corresponding pathways, as some bacteria are able to use more efficiently pathways to metabolise the same substrate (Rocca et al., 2015). As C allocation is mediated by bacteria through decomposition of organic matter, particularly through lignocellulose breakdown (Lladó et al., 2017), we identified cellulolytic genera in our study, which included *Burkholderia*, *Variovorax* and *Flavobacterium* (Štursová et al., 2012). Additional genera with potential cellulolytic potential identified were classified as rare (less than 1%) and included actinobacterial genera (*Arthrobacter*, *Cellulomonas*, *Kitasatospora*, *Oerskovia*, *Micromonospora*, *Streptomyces*), alphabroteobacterial genera (*Mesorhizobium*, *Methylobacterium*, *Sphingomonas*), the gammaproteobacterial genus *Dyella* and genera of the Firmicutes (*Bacillus*, *Paenibacillus*) (Haichar et al., 2007; Wirth and Ulrich, 2002).



**Figure 2. 6. Overview of bacterial genes involved in key energy pathways in soil bacteria. Genes were predicted from KEGG orthologs with Tax4fun. Values represent mean relative abundance in spring, summer and autumn, from 48 mon and mixed stands. Genes used are listed in Supplementary information, Table S2.4. Means with similar letters (determined by ANOVA with Tukey HSD post-hoc test, where  $P \leq 0.05$ ) share no significant differences among seasons.**

We also identified genes encoding methane monooxygenase (MMO), a key enzyme in the oxidation of methane to methanol, harboured by methylotrophic and methanotrophic bacteria (Hakemian and Rosenzweig, 2007; Hoppe et al., 2015). Methylotrophs metabolise single carbon substrates, as a by-product of lignin degradation (Hoppe et al., 2015). Potential methylotrophs were identified from *Rhizobiales* (*Methylocella*, and *Methyloferula*), gammaproteobacterial *Methylococcales* (*Methylococcaceae*) and *Verrucomicrobia* (*Methylacidiphilum*). Seasonal inputs of organic matter from decomposition explain the strong effect of season on abundance of methanotrophs. For example, Hoppe et al., (Hoppe et al., 2015), demonstrated an increase of methylotrophic *Rhizobiales* at different stages of deadwood decomposition in a beech and spruce deciduous forest.

The presence of genes for assimilatory (*sir*, *cys* genes) and dissimilatory (*dsr*) sulphate reduction revealed the presence of sulphur metabolising organisms. However, the abundance of these genes was not significant across season ( $p = 0.818$ ). This is explained by reports that sulphate reducing bacteria tend to be rare within the environment despite carrying out the major part of sulphur metabolism

(Hausmann et al., 2016; Yousuf et al., 2014). Candidates detected included *Desulfurellales*, *Desulfurbacterales* and *Desulfuromonadales* within *Deltaproteobacteria*, *Thiobalophilus* and *Thioalkalispira* from *Gammaproteobacteria*, *Comamonadaceae* (*Betaproteobacteria*), *Rhodopseudomonas* (*Alphaproteobacteria*) and *Desulfosporosinus* (*Firmicutes*). Several studies have identified these taxa as major components of sulphur cycling in soil and sediment ecosystems (Baker et al., 2015; Balk et al., 2015; Hausmann et al., 2016; Ling et al., 2015b).

We recovered genes such as *nir*, *nif*, *hao* and *amo* involved in (de)nitrification, nitrate reduction and anammox, the main processes in nitrogen cycling (Giles et al., 2012; Norton et al., 2002). Groups associated with N cycling from our study included *Nitrospira*, all *Nitroso-* genera and *Rhizobiales* (*Bradyrhizobium* and *Rhizobium*), *Rhodospirillaceae* (*Azospirillum*) and *Actinobacteria* (*Arthrobacter*). There was a significant abundance of N metabolic genes in spring and autumn ( $p \leq 0.01$ ), and similar to C metabolic genes, corresponding to periods of increased nutrient availability. Additionally, soil C allocation affects soil pH and C/N ratio, which subsequently impact nitrification and denitrification, and C cycling (Cardenas et al., 2018; Townsend et al., 2011). This explains the shared trend in seasonal abundance observed for C and N metabolic genes in our study.

## 2.5 Conclusion

Tree species had a strong impact on soil bacterial diversity and composition at entire and active soil bacterial community level, which supported our first hypothesis that tree species identity drives bacterial community structure in the entire and potentially active bacterial community. This is result of both direct and indirect influences such as litterfall and root exudates, which change soil pH, C/N ratio, N and P availability. Beech and oak mono stands displayed low pH and high C/N ratio and subsequently, showed a higher abundance of oligotrophic and lower abundance of copiotrophic bacterial taxa compared to lime and hornbeam mono stands. Mono stands showed a higher number of indicator OTUs corresponding to organisms closely associated with each stand, compared to mixed stands. Most indicator OTUs belonged to *Rhizobiales*, indicating the widespread physiological adaptation of its members to different environments. Our second hypothesis that bacterial community in structure across forest stands is shaped by season to a lesser extent than by tree species was not supported. Season showed a non-significant effect in forest stands at total and entire community level. We did not observe stand-specific effects on predicted bacterial functions, except for genes related to methane metabolism. In contrast, the predicted functional metabolic profiles were significantly influenced by season. This is attributed to functional redundancy across taxonomic groups, which results in lower functional diversity. However, we identified more genes associated with C fixation and degradation (including methane metabolism), compared to nitrogen metabolism. This supports evidence that forest ecosystems play a central role in carbon storage and contribute to global carbon cycling. The abundance of C and metabolic genes showed similar increase in spring and autumn, which is linked to shared metabolic pathways across different bacterial taxa. An increase in genes for both processes corresponded to periods of increased soil nutrient availability, as a response to increased root productivity (spring) and litterfall (autumn). The ability to monitor taxonomic and functional relationships in individual microbial taxa provides greater insight into specific impacts of trees in shaping the soil microbial communities. This improves our understanding of how potential conversion of forest stands effects changes in belowground microbial community.

## Acknowledgments

The authors wish to thank all personnel who contributed in sample collection and transportation. This work is supported by the Deutsche Forschungsgemeinschaft (DFG) as part of the Research training group GRK 1086 (<http://www.unigoettingen.de/de/82664.html>).

## 2.6 Supplementary information

### Tree species shape soil bacterial community structure and function in temperate deciduous forests

Amélie Dukunde<sup>1</sup>, Dominik Schneider<sup>1</sup>, Marcus Schmidt<sup>2</sup>, Rolf Daniel<sup>1\*</sup>

#### Content

Figure S2. 1. Boxplots showing soil environmental parameters in mono and mixed stands,

Figure S2. 2. Alpha diversity metrics across different tree stands.

Figure S2. 3. Bacterial orders showing significant composition across different tree stands.

Figure S2. 4. Bipartite association networks between soil bacterial communities (genus level) and tree stands.

Figure S2. 5 Non-metric multidimensional scaling (NMDS) ordination of methane metabolic genes.

Table S2. 1. Summary of sequence data generated with high throughput pyrotag sequencing.

Table S2. 2. Statistical tests of tree stand effects on bacterial community.

Table S2. 3. Summary of abundant soil bacterial genera across mono and mixed stands.

Table S2. 4. Statistical tests of tree stand effects on bacterial community function.

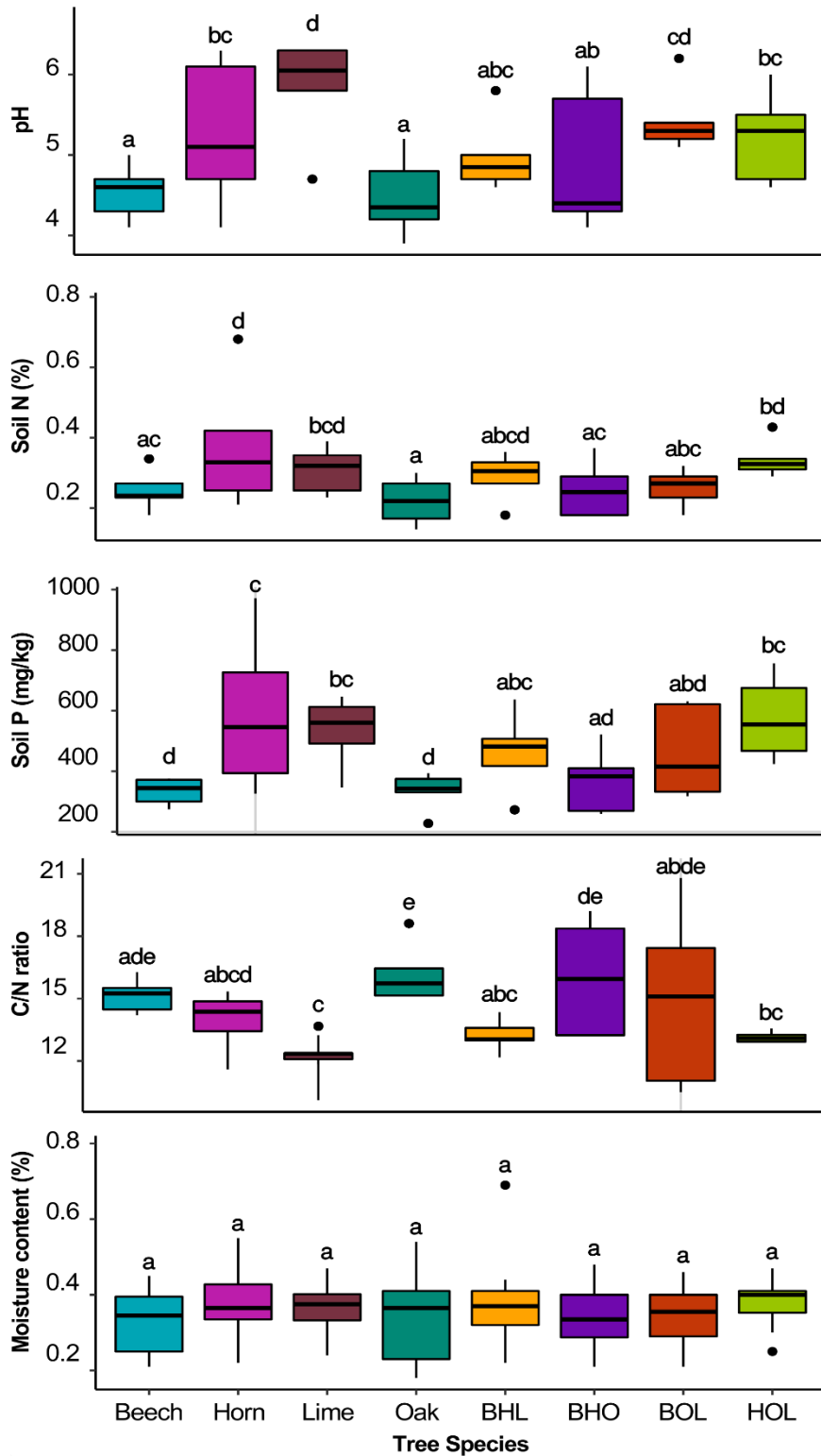
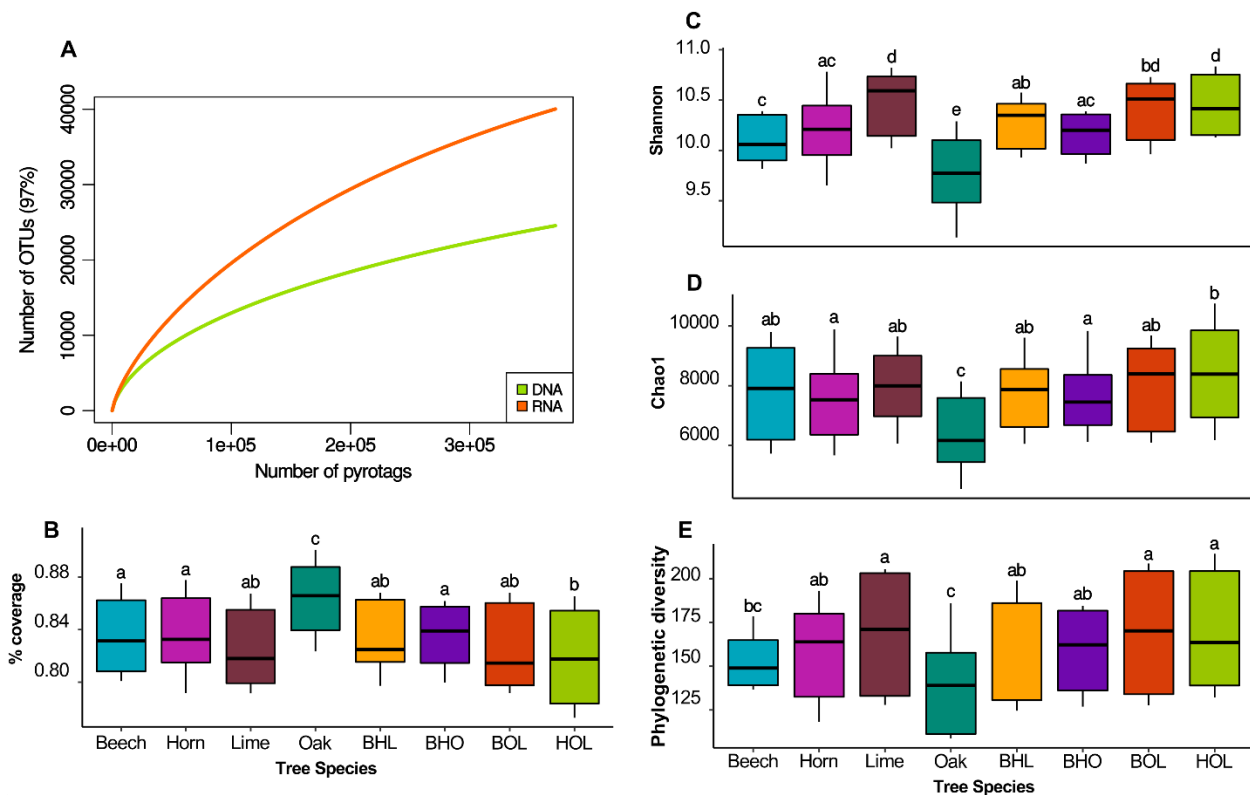
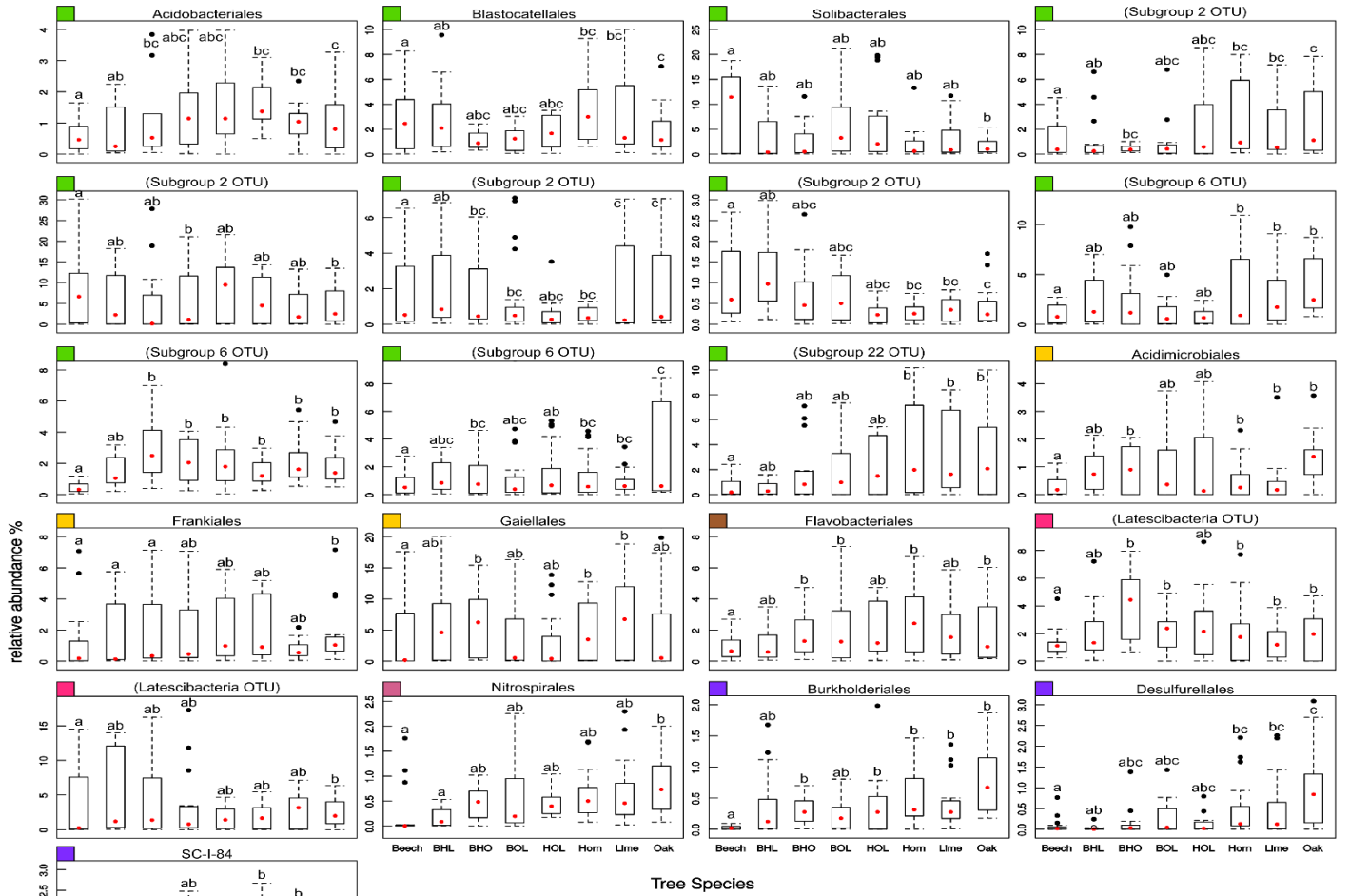


Figure S2. 1. Boxplots showing soil environmental parameters in mono and mixed stands. Mean values, where  $n = 6$  replicate stands, are given. Horizontal bars show median value. Mean values with similar letters (determined by ANOVA with Tukey HSD post-hoc test, where  $p \leq 0.05$ ) share no significant differences. Soil data was previously collected and described by Schmidt et al. (2015).

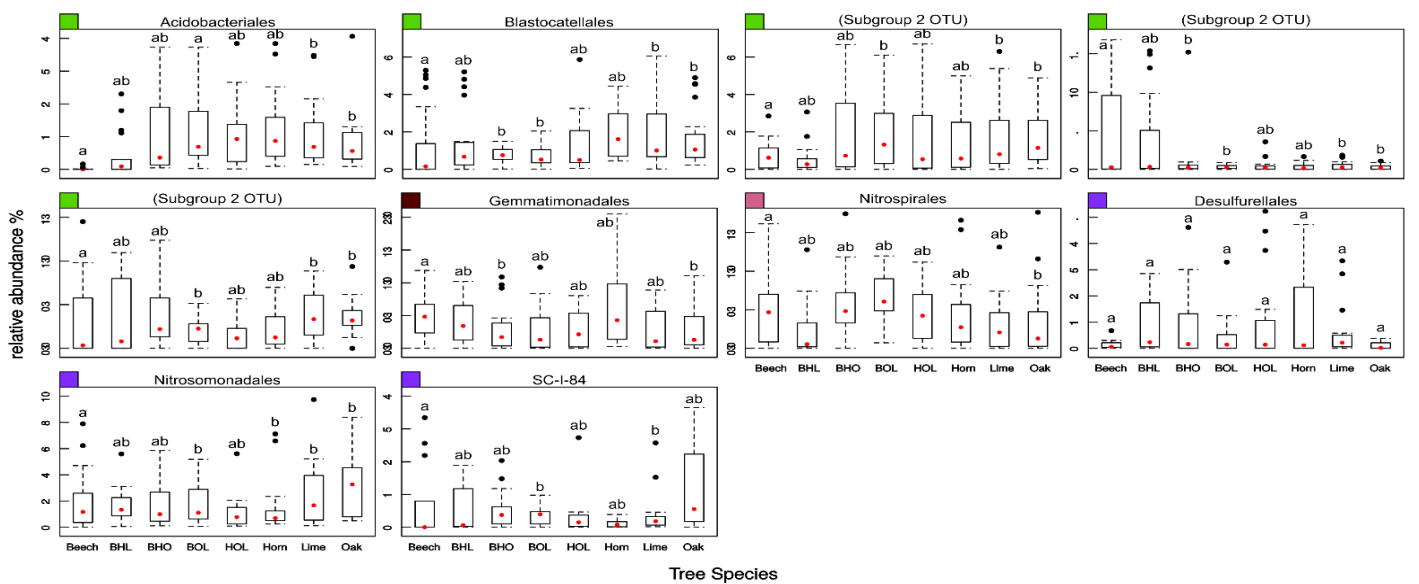


**Figure S2. 2. Alpha diversity metrics across different tree stands. A) Rarefaction curves of subsampled OTUs for environmental DNA and RNA, at 97% similarity. B) Good's coverage, C) Shannon richness index, Chao1 estimator and phylogenetic diversity (PD). Mean values are given, where  $n = 6$  replicate stands for three seasons. Horizontal bars represent the median value. Mean values with similar letters indicate no significant differences between stands (determined by ANOVA with Tukey HSD post-hoc test, where  $P \leq 0.05$ ).**

### Total community



### Active community

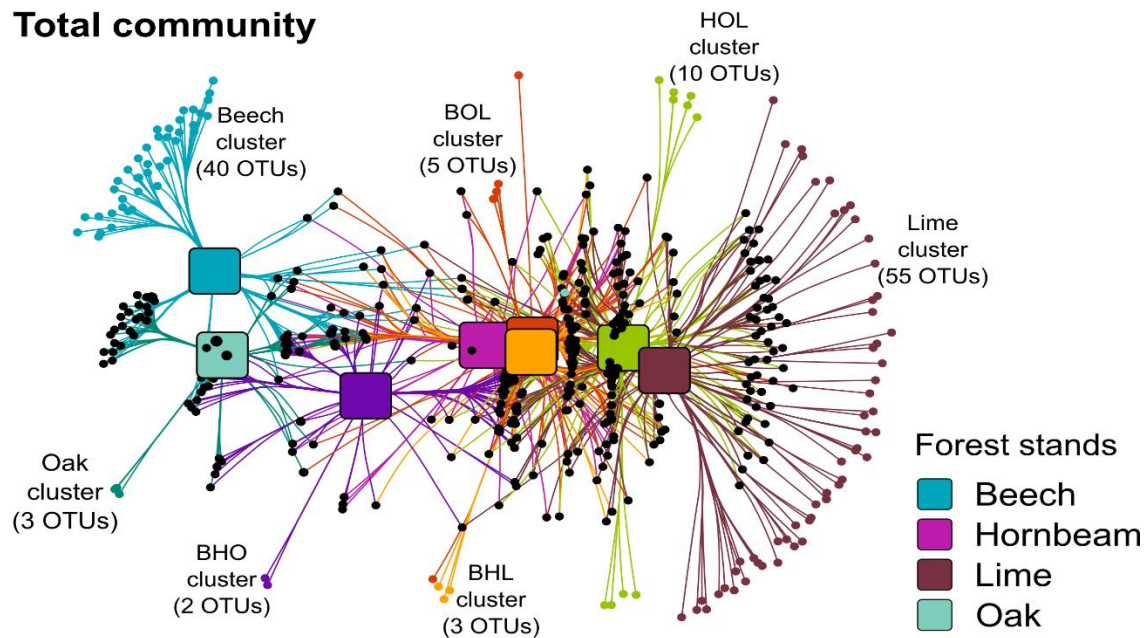


■ Acidobacteria 
 ■ Actinobacteria 
 ■ Bacteroidetes 
 ■ Gemmatimonadetes 
 ■ Latescibacteria 
 ■ Nitrospirae 
 ■ Proteobacteria



**Figure S2. 3. Bacterial orders showing significant composition across different tree stands. Only bacteria orders belonging to dominant phyla (greater than 1 %) are shown. Mean values, where  $n = 6$  replicate stands for three seasons, are given. Statistical calculations employed Kruskal-Wallis test with Dunn's post hoc test. A  $p \leq 0.05$  was considered statistically significant. Mean values with the same letter significant difference in composition. Red dots represent the median value. Black dots represent outlying data.**

## Total community



## Active community

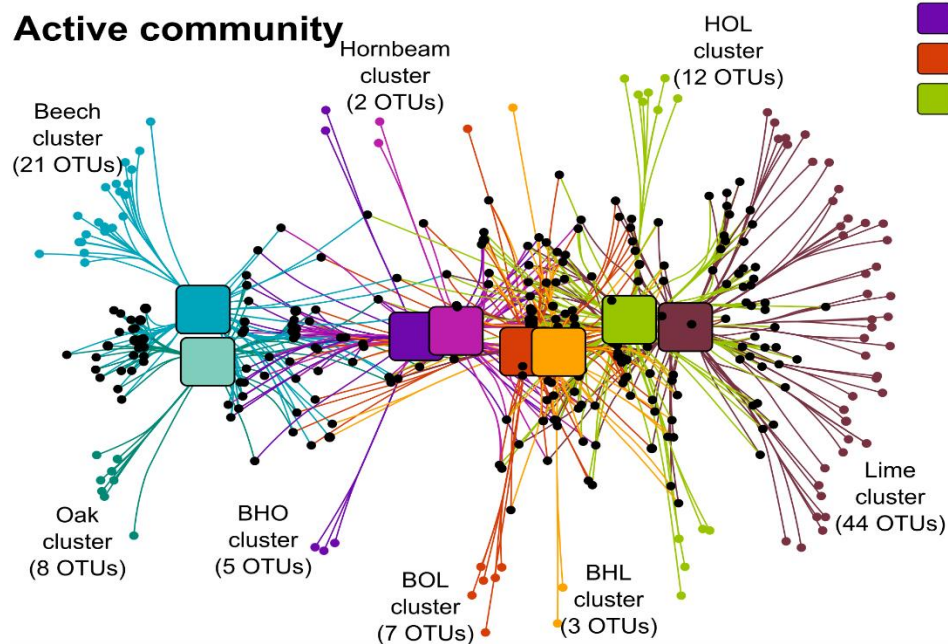


Figure S2. 4. Bipartite association networks between soil bacterial communities (genus level) and tree stands. Source nodes (rounded squares) represent tree stands and edges represent associations between stands and bacterial OTUs (circles, target nodes). Edges are coloured according to the source tree stand and the length of edges is weighted according to association strength. Unique clusters, which associate with one tree species, consist of nodes coloured as the corresponding stand. Numbers of OTUs making up respective unique clusters are given in parenthesis. Black circles represent OTUs with significant cross association between two or more stands. Target node sizes represent mean relative abundance of OTUs across all mono plots. Data only represents OTUs that showed significant positive association with tree stands ( $P \leq 0.05$ ). For ease of visualisation, edges were bundled with a stress value of 3.

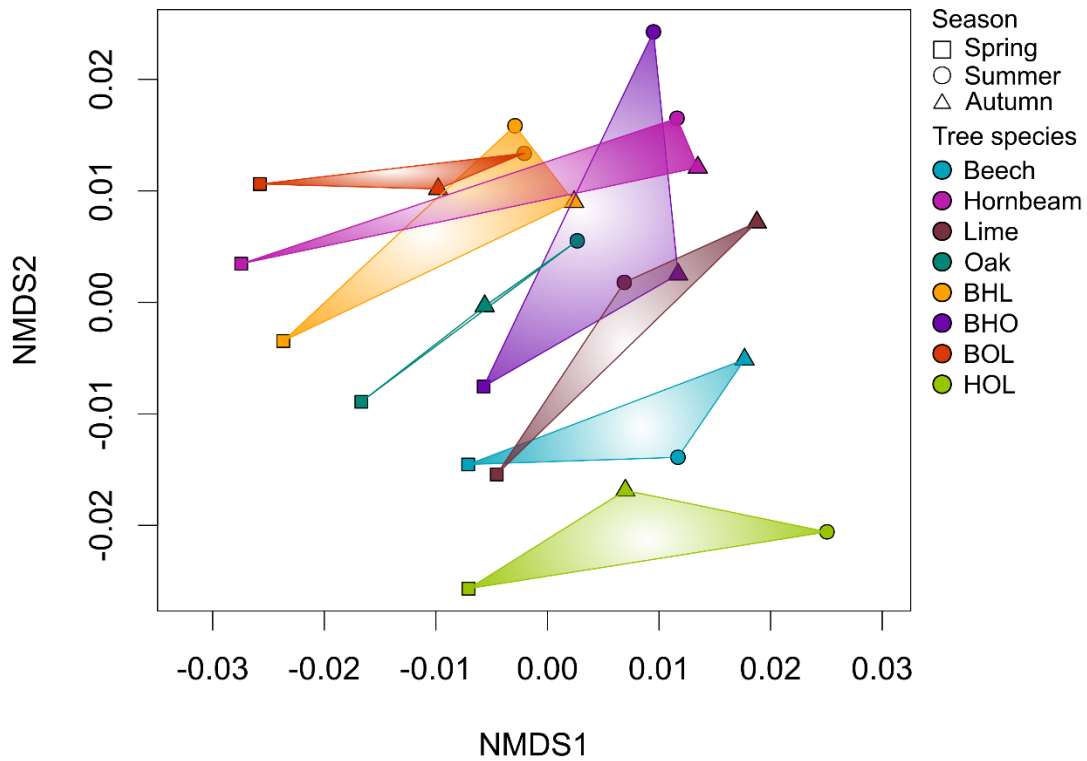


Figure S2. 5. Non-metric multidimensional scaling (NMDS) ordination of methane metabolic genes. Significant grouping of metabolic genes according to tree species ( $p = 0.033$ ) is given. NMDS was constructed with a Bray Curtis distance matrix using KEGG orthologs predicted with Tax4Fun. Genes used to construct the plot are listed in Supplementary information, Table S2.5.

**Table S2. 1. Summary of sequence data generated with high throughput pyrotag sequencing.**

| Reads after key steps of sequencing pipeline |                               | Template      |                  |
|--|-------------------------------|---------------|------------------|
|  |                               | DNA           | RNA <sup>a</sup> |
| 1.   | Raw reads                     | 3 335 821     | 1 308 705        |
| 2.   | Quality filtering             | 3 052 200     | 1 291 122        |
| 3.   | Chimera and singleton removal | 2 402 787     | 857 868          |
| 4.   | Taxonomy filtering            | 2 401 078     | 857 464          |
| % Recovery                                   |                               | 72.4          | 65.7             |
| Total OTUs <sup>b</sup>                      |                               | <b>40 385</b> | <b>52 277</b>    |

<sup>a</sup> Reads for RNA were generated via cDNA

<sup>b</sup> OTUs generated after clustering of abovementioned reads. Numbers represent total count across all plots in spring, summer and autumn from subsampled datasets of environmental DNA and RNA.

<sup>c</sup> Recovery refers to the percentage of sequence reads after processing over the total reads

**Table S2. 2. Statistical tests of tree stand effects on bacterial community. Results of ANOSIM and PERMANOVA were calculated using weighted uniFrac matrices representing the total (DNA-based) and potentially active (RNA-based) bacterial community. A p value  $\leq 0.05$  was considered significant.**

| Test         | ANOSIM  |        |         |       | PERMANOVA |          |         |          |
|--------------|---------|--------|---------|-------|-----------|----------|---------|----------|
|              | DNA     |        | RNA     |       | DNA       |          | RNA     |          |
|              | p-value | R      | p-value | R     | p-value   | Psuedo F | p-value | Psuedo F |
| Tree species | 0.001   | 0.72   | 0.001   | 0.63  | 0.001     | 9.4      | 0.001   | 4.9      |
| Stand type   | 0.008   | 0.22   | 0.002   | 0.23  | 0.1       | 2.4      | 0.044   | 2.9      |
| Season       | 0.464   | -0.009 | 0.038   | 0.132 | 0.529     | 0.73     | 0.132   | 1.7      |

**NMDS goodness of fit**

|              | p-value | R <sup>2</sup> |     |
|--------------|---------|----------------|-----|
| Template     | 0.001   | 0.439          | *** |
| Tree species | 0.001   | 0.506          | *** |
| Stand type   | 0.043   | 0.068          | *   |
| Season       | 0.858   | 0.014          |     |

**Environmental parameters**

|           | p-value | R <sup>2</sup> |     |
|-----------|---------|----------------|-----|
| pH        | 0.001   | 0.809          | *** |
| N         | 0.001   | 0.353          | *** |
| C/N ratio | 0.001   | 0.683          | *** |
| P         | 0.001   | 0.589          | *** |
| MC        | 0.001   | 0.655          | *** |

Abbreviations: N - nitrogen; C/N - carbon/nitrogen; P - phosphorus; MC - moisture content

Significant codes: 0 - \*\*\*, 0.001 - \*\*, 0.01 - \*, 0.05 ., 0.1, 1

**Table S2. 3. Summary of abundant soil bacterial genera across mono and mixed stands. Values represent mean relative abundance for 6 replicate stands in spring, summer and autumn. Values are given for total (DNA-based) community. Only genera from abundant orders (greater than 1 %) are given. Where the order or genus is unidentified, the taxonomic name at the highest known resolution is given in parenthesis.**

| Total community                | Order                      | Genus                          | Relative abundance (%) in forest stands |       |       |        |       |              |       |       |         |
|--------------------------------|----------------------------|--------------------------------|---|-------|-------|--------|-------|--------------|-------|-------|---------|
|                                |                            |                                | Mono stands                             |       |       |        |       | Mixed stands |       |       | Average |
|                                |                            |                                | Beech                                   | Horn  | Lime  | Oak    | BHL   | BHO          | BOL   | HOL   |         |
| <b><i>Acidobacteria</i></b>    |                            |                                |   |       |       |        |       |              |       |       |         |
|                                | <i>Acidobacteriales</i>    | ( <i>Acidobacteriaceae</i> )   | 13.743                                  | 6.886 | 1.171 | 12.836 | 5.240 | 9.314        | 5.109 | 3.315 | 7.202   |
|                                | <i>Solibacterales</i>      | <i>Bryobacter</i>              | 6.200                                   | 4.344 | 1.521 | 6.902  | 3.569 | 5.284        | 4.110 | 2.777 | 4.338   |
|                                |                            | <i>Candidatus Solibacter</i>   | 5.047                                   | 3.845 | 2.606 | 4.297  | 4.090 | 4.202        | 3.612 | 3.182 | 3.860   |
|                                | <i>Subgroup 2</i>          | ( <i>Subgroup 2</i> )          | 13.202                                  | 6.570 | 1.292 | 11.988 | 4.275 | 7.672        | 4.483 | 3.389 | 6.609   |
|                                | <i>Subgroup 6</i>          | ( <i>Subgroup 6</i> )          | 2.109                                   | 4.351 | 8.014 | 2.103  | 5.546 | 3.494        | 4.919 | 5.848 | 4.548   |
|                                |                            | ( <i>Subgroup 6</i> )          | 0.856                                   | 2.059 | 3.413 | 0.847  | 2.682 | 1.500        | 2.223 | 2.617 | 2.024   |
| <b><i>Actinobacteria</i></b>   |                            |                                |   |       |       |        |       |              |       |       |         |
|                                | <i>Acidimicrobiales</i>    | ( <i>Acidimicrobiales</i> )    | 0.768                                   | 0.989 | 0.953 | 0.826  | 0.928 | 0.933        | 1.022 | 1.066 | 0.936   |
|                                | <i>Frankiales</i>          | <i>Acidothermus</i>            | 2.528                                   | 4.032 | 0.906 | 3.023  | 2.212 | 3.185        | 2.251 | 1.951 | 2.511   |
|                                | <i>Gaiellales</i>          | ( <i>Gaiellales</i> )          | 1.147                                   | 2.235 | 2.142 | 1.083  | 2.297 | 1.779        | 2.466 | 2.624 | 1.972   |
|                                | <i>Solirubrobacterales</i> | ( <i>Solirubrobacterales</i> ) | 0.169                                   | 0.475 | 0.708 | 0.165  | 0.477 | 0.392        | 0.594 | 0.699 | 0.460   |
| <b><i>Bacteroidetes</i></b>    |                            |                                |   |       |       |        |       |              |       |       |         |
|                                | <i>Cytophagales</i>        | ( <i>Cytophagaceae</i> )       | 0.659                                   | 1.320 | 1.583 | 0.936  | 1.457 | 1.325        | 1.284 | 1.797 | 1.295   |
|                                | <i>Flavobacteriales</i>    | <i>Flavobacterium</i>          | 0.331                                   | 1.019 | 1.788 | 0.446  | 0.807 | 0.728        | 1.117 | 1.686 | 0.990   |
|                                | <i>Sphingobactriales</i>   | ( <i>Chitinophagaceae</i> )    | 0.965                                   | 1.006 | 0.891 | 1.037  | 0.971 | 1.061        | 1.093 | 1.039 | 1.008   |
| <b><i>Gemmatimonadetes</i></b> |                            |                                |   |       |       |        |       |              |       |       |         |
|                                | <i>Gemmatimonadales</i>    | <i>Gemmatimonas</i>            | 1.458                                   | 1.189 | 1.581 | 1.091  | 1.605 | 1.102        | 1.886 | 1.427 | 1.417   |
|                                |                            | ( <i>Gemmatimonadaceae</i> )   | 1.248                                   | 1.098 | 1.343 | 0.706  | 1.147 | 1.017        | 1.493 | 1.167 | 1.152   |
| <b><i>Latescibacteria</i></b>  |                            |                                |   |       |       |        |       |              |       |       |         |
|                                | ( <i>Latescibacteria</i> ) | ( <i>Latescibacteria</i> )     | 0.200                                   | 0.618 | 1.479 | 0.162  | 0.899 | 0.578        | 0.807 | 0.889 | 0.704   |
|                                |                            | ( <i>Latescibacteria</i> )     | 0.066                                   | 0.215 | 0.418 | 0.086  | 0.450 | 0.132        | 0.341 | 0.459 | 0.271   |
|                                |                            | ( <i>Latescibacteria</i> )     | 0.018                                   | 0.088 | 0.265 | 0.025  | 0.145 | 0.063        | 0.120 | 0.112 | 0.105   |
| <b><i>Nitrospirae</i></b>      |                            |                                |   |       |       |        |       |              |       |       |         |
|                                | <i>Nitrospirales</i>       | ( <i>Nitrospirales</i> )       | 0.042                                   | 0.438 | 1.380 | 0.062  | 0.617 | 0.349        | 0.557 | 0.616 | 0.508   |
|                                |                            | <i>Nitrospira</i>              | 0.350                                   | 0.255 | 0.507 | 0.200  | 0.301 | 0.186        | 0.371 | 0.303 | 0.309   |
| <b><i>Proteobacteria</i></b>   |                            |                                |   |       |       |        |       |              |       |       |         |
|                                | <i>Burkholderiales</i>     | ( <i>Comamonadaceae</i> )      | 0.305                                   | 0.570 | 0.917 | 0.368  | 0.739 | 0.469        | 0.681 | 0.932 | 0.623   |
|                                |                            | <i>Variovorax</i>              | 0.103                                   | 0.197 | 0.273 | 0.126  | 0.231 | 0.144        | 0.257 | 0.255 | 0.198   |
|                                |                            | <i>Paraburkholderia</i>        | 0.151                                   | 0.079 | 0.112 | 0.129  | 0.052 | 0.140        | 0.046 | 0.042 | 0.094   |
|                                |                            | <i>Rhizobacter</i>             | 0.033                                   | 0.068 | 0.141 | 0.030  | 0.110 | 0.058        | 0.088 | 0.140 | 0.083   |
|                                | <i>Desulfurellales</i>     | ( <i>Desulfurellaceae</i> )    | 0.294                                   | 0.538 | 1.325 | 0.248  | 0.735 | 0.474        | 0.770 | 0.859 | 0.655   |
|                                |                            | ( <i>Desulfurellaceae</i> )    | 0.001                                   | 0.022 | 0.058 | 0.000  | 0.023 | 0.013        | 0.023 | 0.035 | 0.022   |
|                                | <i>Myxococcales</i>        | <i>Haliangium</i>              | 0.854                                   | 1.379 | 1.381 | 1.174  | 1.409 | 1.303        | 1.718 | 1.507 | 1.341   |
|                                |                            | <i>Sorangium</i>               | 0.597                                   | 0.495 | 0.354 | 0.577  | 0.450 | 0.488        | 0.545 | 0.382 | 0.486   |
|                                | <i>Nitrosomonadales</i>    | ( <i>Nitrosomonadaceae</i> )   | 1.297                                   | 1.784 | 2.644 | 1.207  | 2.186 | 1.951        | 2.404 | 2.050 | 1.940   |
|                                |                            | <i>Nitrosospira</i>            | 0.003                                   | 0.014 | 0.005 | 0.004  | 0.014 | 0.005        | 0.007 | 0.010 | 0.008   |
|                                | <i>Rhizobiales</i>         | ( <i>Xanthobacteraceae</i> )   | 1.198                                   | 2.825 | 3.878 | 1.366  | 3.402 | 2.045        | 3.122 | 3.605 | 2.680   |

|                         |                             |       |       |       |       |       |       |       |       |       |
|-------------------------|-----------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|                         | <i>Bradyrhizobium</i>       | 1.794 | 2.572 | 2.448 | 2.197 | 2.741 | 2.606 | 2.516 | 2.834 | 2.464 |
|                         | <i>Rhizomicrobium</i>       | 2.768 | 2.247 | 0.745 | 3.150 | 2.050 | 2.816 | 1.706 | 1.471 | 2.119 |
|                         | <i>Variibacter</i>          | 1.672 | 1.798 | 1.505 | 1.478 | 1.815 | 1.522 | 1.751 | 1.763 | 1.663 |
| <i>Rhodospirillales</i> | ( <i>DA111</i> )            | 2.012 | 1.777 | 2.248 | 1.396 | 2.084 | 1.500 | 2.000 | 2.133 | 1.894 |
|                         | ( <i>Acetobacteraceae</i> ) | 1.969 | 1.671 | 0.885 | 2.549 | 1.665 | 2.325 | 1.202 | 1.021 | 1.661 |
|                         | <i>Reyranella</i>           | 0.618 | 1.457 | 2.469 | 0.835 | 1.877 | 1.122 | 1.367 | 1.695 | 1.430 |
| <i>Xanthomonadales</i>  | <i>Rhodanobacter</i>        | 1.865 | 1.874 | 1.087 | 2.171 | 1.832 | 1.538 | 2.680 | 1.851 | 1.862 |
|                         | <i>Acidibacter</i>          | 1.577 | 1.369 | 1.130 | 1.720 | 1.073 | 1.461 | 1.026 | 0.939 | 1.287 |
|                         | ( <i>Xanthomonadales</i> )  | 1.579 | 1.316 | 0.128 | 1.999 | 0.780 | 1.999 | 0.685 | 0.541 | 1.128 |

Table S2.3 *cont.* Summary of abundant soil bacterial genera across mono and mixed stands. Values represent mean relative abundance for 6 replicate stands in spring, summer and autumn. Values are given for the potentially active (RNA-based) community. Only genera from abundant orders (greater than 1 %) are given. Where the order or genus is unidentified, the taxonomic name at the highest known resolution is given in parenthesis.

| Active Community           |                                | Forest stands |       |       |       |              |       |       |       |         |
|----------------------------|--------------------------------|---------------|-------|-------|-------|--------------|-------|-------|-------|---------|
| Order                      | Genus                          | Mono stands   |       |       |       | Mixed stands |       |       |       |         |
|                            |                                | Beech         | Horn  | Lime  | Oak   | BHL          | BHO   | BOL   | HOL   | Average |
| <b>Acidobacteria</b>       |                                |               |       |       |       |              |       |       |       |         |
| <i>Acidobacteriales</i>    | ( <i>Acidobacteriaceae</i> )   | 8.292         | 4.069 | 0.698 | 9.534 | 3.438        | 5.447 | 3.211 | 1.883 | 4.572   |
| <i>Solibacterales</i>      | <i>Candidatus Solibacter</i>   | 6.811         | 3.446 | 2.090 | 6.468 | 2.691        | 3.555 | 3.734 | 2.480 | 3.909   |
|                            | <i>Bryobacter</i>              | 5.890         | 4.208 | 1.000 | 7.002 | 2.720        | 3.441 | 3.130 | 1.527 | 3.615   |
| <i>Subgroup 2</i>          | ( <i>Subgroup 2</i> )          | 7.570         | 3.765 | 0.841 | 6.741 | 3.163        | 3.819 | 2.369 | 1.724 | 3.749   |
| <i>Subgroup 6</i>          | ( <i>Subgroup 6</i> )          | 0.882         | 1.562 | 3.458 | 0.727 | 1.919        | 1.622 | 2.005 | 2.260 | 1.804   |
|                            | ( <i>Subgroup 6</i> )          | 0.337         | 0.631 | 1.425 | 0.267 | 0.857        | 0.745 | 0.897 | 1.065 | 0.778   |
| <b>Actinobacteria</b>      |                                |               |       |       |       |              |       |       |       |         |
| <i>Acidimicrobiales</i>    | ( <i>Acidimicrobiales</i> )    | 0.790         | 0.754 | 0.858 | 0.824 | 0.736        | 0.842 | 0.752 | 0.698 | 0.782   |
| <i>Frankiales</i>          | <i>Acidothermus</i>            | 4.510         | 5.560 | 1.189 | 5.036 | 3.735        | 5.687 | 3.107 | 2.897 | 3.965   |
| <i>Gaiellales</i>          | ( <i>Gaiellales</i> )          | 0.645         | 0.907 | 0.897 | 0.591 | 1.011        | 0.828 | 0.871 | 1.054 | 0.851   |
| <i>Solirubrobacterales</i> | ( <i>Solirubrobacterales</i> ) | 0.054         | 0.145 | 0.280 | 0.034 | 0.174        | 0.098 | 0.122 | 0.199 | 0.138   |
| <b>Bacteroidetes</b>       |                                |               |       |       |       |              |       |       |       |         |
| <i>Cytophagales</i>        | ( <i>Cytophagaceae</i> )       | 0.175         | 0.234 | 0.546 | 0.158 | 0.399        | 0.485 | 0.361 | 0.531 | 0.361   |
| <i>Flavobacteriales</i>    | <i>Flavobacterium</i>          | 0.240         | 1.114 | 1.787 | 0.321 | 0.880        | 0.825 | 2.716 | 2.227 | 1.264   |
| <i>Sphingobactriales</i>   | ( <i>Chitinophagaceae</i> )    | 1.891         | 2.481 | 2.754 | 2.168 | 2.351        | 2.151 | 3.214 | 2.878 | 2.486   |
| <b>Gemmatimonadetes</b>    |                                |               |       |       |       |              |       |       |       |         |
| <i>Gemmatimonadales</i>    | <i>Gemmatimonas</i>            | 0.477         | 0.310 | 0.455 | 0.342 | 0.554        | 0.278 | 0.459 | 0.432 | 0.413   |
|                            | ( <i>Gemmatimonadaceae</i> )   | 0.310         | 0.223 | 0.403 | 0.125 | 0.248        | 0.222 | 0.252 | 0.277 | 0.258   |
| <b>Latescibacteria</b>     |                                |               |       |       |       |              |       |       |       |         |
| ( <i>Latescibacteria</i> ) | ( <i>Latescibacteria</i> )     | 0.060         | 0.104 | 0.356 | 0.020 | 0.159        | 0.145 | 0.186 | 0.224 | 0.157   |
|                            | ( <i>Latescibacteria</i> )     | 0.044         | 0.077 | 0.184 | 0.031 | 0.166        | 0.077 | 0.178 | 0.237 | 0.124   |
|                            | ( <i>Latescibacteria</i> )     | 0.006         | 0.001 | 0.025 | 0.006 | 0.002        | 0.022 | 0.011 | 0.016 | 0.011   |
| <b>Nitrospirae</b>         |                                |               |       |       |       |              |       |       |       |         |
| <i>Nitrospirales</i>       | ( <i>Nitrospirales</i> )       | 0.188         | 0.092 | 0.168 | 0.076 | 0.122        | 0.118 | 0.130 | 0.096 | 0.124   |
|                            | <i>Nitrospira</i>              | 0.010         | 0.063 | 0.258 | 0.004 | 0.057        | 0.085 | 0.073 | 0.131 | 0.085   |
| <b>Proteobacteria</b>      |                                |               |       |       |       |              |       |       |       |         |
| <i>Burkholderiales</i>     | ( <i>Comamonadaceae</i> )      | 0.812         | 1.680 | 2.296 | 0.867 | 2.439        | 1.404 | 2.063 | 2.382 | 1.743   |
|                            | <i>Variovorax</i>              | 0.382         | 0.573 | 0.824 | 0.397 | 0.699        | 0.409 | 0.754 | 0.945 | 0.623   |

|                         |                              |       |       |       |       |       |       |       |       |       |
|-------------------------|------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|                         | <i>Rhizobacter</i>           | 0.099 | 0.254 | 0.567 | 0.134 | 0.450 | 0.257 | 0.430 | 0.565 | 0.344 |
| <i>Desulfurellales</i>  | ( <i>Desulfurellaceae</i> )  | 0.369 | 0.670 | 1.063 | 0.261 | 0.661 | 0.528 | 0.601 | 0.725 | 0.610 |
|                         | ( <i>Desulfurellaceae</i> )  | 0.002 | 0.044 | 0.160 | 0.001 | 0.058 | 0.041 | 0.048 | 0.084 | 0.055 |
| <i>Myxococcales</i>     | <i>Haliangium</i>            | 3.704 | 5.391 | 8.289 | 3.916 | 7.481 | 5.181 | 6.648 | 7.349 | 5.995 |
|                         | <i>Sorangium</i>             | 3.656 | 2.587 | 2.427 | 3.060 | 2.703 | 2.268 | 2.441 | 2.099 | 2.655 |
| <i>Nitrosomonadales</i> | ( <i>Nitrosomonadaceae</i> ) | 0.945 | 1.226 | 1.986 | 0.808 | 1.695 | 1.657 | 1.879 | 1.492 | 1.461 |
|                         | <i>Nitrosospira</i>          | 0.017 | 0.033 | 0.036 | 0.010 | 0.061 | 0.038 | 0.056 | 0.067 | 0.040 |
| <i>Rhizobiales</i>      | ( <i>Xanthobacteraceae</i> ) | 2.960 | 5.329 | 5.729 | 2.544 | 5.982 | 4.399 | 5.457 | 5.892 | 4.787 |
|                         | <i>Variibacter</i>           | 3.275 | 4.024 | 2.811 | 2.536 | 3.423 | 3.777 | 3.201 | 3.353 | 3.300 |
|                         | <i>Bradyrhizobium</i>        | 2.460 | 3.514 | 3.067 | 2.269 | 3.997 | 3.287 | 2.746 | 3.981 | 3.165 |
|                         | <i>Rhizomicrobium</i>        | 1.213 | 0.895 | 0.261 | 1.367 | 0.668 | 0.821 | 0.978 | 0.417 | 0.827 |
| <i>Rhodospirillales</i> | ( <i>DA111</i> )             | 4.054 | 3.105 | 2.941 | 2.777 | 3.742 | 2.820 | 2.860 | 2.884 | 3.148 |
|                         | ( <i>Acetobacteraceae</i> )  | 2.586 | 2.159 | 0.744 | 3.070 | 1.671 | 2.567 | 1.394 | 0.706 | 1.862 |
|                         | <i>Reyranella</i>            | 0.440 | 0.666 | 0.721 | 0.498 | 0.630 | 0.468 | 0.740 | 0.566 | 0.591 |
| <i>Xanthomonadales</i>  | <i>Rhodanobacter</i>         | 1.904 | 1.320 | 0.063 | 2.810 | 0.937 | 2.318 | 0.752 | 0.241 | 1.293 |
|                         | <i>Acidibacter</i>           | 1.418 | 1.431 | 0.929 | 1.931 | 1.036 | 1.506 | 0.984 | 0.910 | 1.268 |
|                         | ( <i>Xanthomonadales</i> )   | 1.091 | 1.148 | 0.469 | 1.227 | 0.683 | 1.099 | 0.712 | 0.557 | 0.873 |

**Table S2. 4. Statistical tests of tree stand effects on bacterial community function. NMDS was calculated using Bray Curtis distance matrices representing the genes predicted from the potentially active (RNA-based) bacterial community with Tax4Fun. Tax4Fun predicted assigned function to 45% of input OTUs from the potentially active community. A p value  $\leq 0.05$  was considered statistically significant.**

| NMDS goodness of fit       | p-value | R <sup>2</sup> | Significant code |
|----------------------------|---------|----------------|------------------|
| <b>Carbon metabolism</b>   |         |                |                  |
| Tree species               | 0.079   | 0.45           | .                |
| Stand type                 | 0.421   | 0.041          |                  |
| Season                     | 0.001   | 0.42           | ***              |
| <b>Methane metabolism</b>  |         |                |                  |
| Tree species               | 0.033   | 0.48           | *                |
| Stand type                 | 0.763   | 0.011          |                  |
| Season                     | 0.001   | 0.39           | ***              |
| <b>Sulphur metabolism</b>  |         |                |                  |
| Tree species               | 0.078   | 0.44           | .                |
| Stand type                 | 0.239   | 0.065          |                  |
| Season                     | 0.001   | 0.42           | ***              |
| <b>Nitrogen metabolism</b> |         |                |                  |
| Tree species               | 0.203   | 0.38           |                  |
| Stand type                 | 0.398   | 0.04           |                  |
| Season                     | 0.001   | 0.49           | ***              |

Significant codes: 0 - \*\*\*, 0.001 - \*\*, 0.01 - \*, 0.05 ., 0.1, 1

**Table S2. 5. List of selected genes associated with energy metabolism. Genes were predicted from the potentially active (RNA-based) bacterial community with Tax4Fun. Tax4Fun assigned function to 45% of input OTUs from the potentially active community.**

| KEGG ortholog              | Name   | KEGG ortholog         | Name   |
|----------------------------|--|-----------------------|--|
| <b>Calvin Cycle</b>        |  | <b>Hemicellulases</b> |  |
| K00174 - K00177            | 2-oxoglutarate ferredoxin oxidoreductase                           | K01181                | endo-1,4-beta-xylanase                       |
| K01601 - K01602            | ribulose-bisphosphate carboxylase                                  | K01198                | xylan 1,4-beta-xylosidase                    |
| K00239 - K00242            | succinate dehydrogenase  | K01218                | mannan endo-1,4-beta-mannosidase             |
| <b>Methane metabolism</b>  |  | K01224                | arabinogalactan endo-1,4-beta-galactosidase  |
| K14028 - K14029            | methanol dehydrogenase (cytochrome c)                              | K01684                | galactonate dehydratase                      |
| K00190 - K00198            | carbon monoxide dehydrogenase / acetyl-CoA synthase                | K01811                | alpha-D-xyloside xylohydrolase               |
| K03518 - K03520            | carbon-monoxide dehydrogenase small subunit                        | K15531                | oligosaccharide reducing-end xylanase        |
| K16154 - K16162            | methane monoxygenase   | K15921                | arabinoxylan arabinofuranohydrolase          |
| <b>Sulphur metabolism</b>  |  | K15924                | glucuronoarabinoxylan endo-1,4-beta-xylanase |
| K00380 - K00381            | sulfite reductase (NADPH) flavoprotein                             | <b>Cellulases</b>     |  |
| K00390                     | phosphoadenosine phosphosulfate reductase                          | K01179                | endoglucanase                                |
| K00392                     | sulfite reductase (ferredoxin)                                     | K01180                | endo-1,3(4)-beta-glucanase                   |
| K00394 - K00395            | adenylylsulfate reductase  | K01182                | oligo-1,6-glucosidase                        |
| K00860                     | adenylylsulfate kinase   | K01195                | beta-glucuronidase                           |
| K00955                     | bifunctional enzyme CysN/CysC                                      | K01225                | cellulose 1,4-beta-cellobiosidase            |
| K00956 - K00958            | sulfate adenylyltransferase  | K05349                | beta-glucosidase                             |
| K11180 - K11181            | sulfite reductase, dissimilatory-type                              | K05350                | beta-glucosidase                             |
| <b>Nitrogen metabolism</b> |  | K16213                | mannobiose 2-epimerase                       |
| K00360 - K00363            | nitrate reductase (NADH)   | <b>Chitinases</b>     |  |
| K00366 - K00367            | ferredoxin-nitrite reductase                                       | K01183                | chitinase                                    |
| K00368 - K00374            | nitrite reductase (NO-forming)                                     | K01452                | chitin deacetylase                           |
| K00376                     | nitrous-oxide reductase  | K03791                | putative chitinase                           |
| K00531                     | nitrogenase  | K03933                | chitin-binding protein                       |
| K01428 - K01430            | urease   | K13381                | bifunctional chitinase/lysozyme              |
| K02164, K02305, K02448     | nitric oxide reductase NorE, NorD                                  | <b>Phosphatases</b>   |  |
| K02586 - K02588            | nitrogenase molybdenum-iron protein (NifE, NifH)                   | K01077                | alkaline phosphatase                         |
| K02591 - K02596, K02597    | nitrogenase molybdenum-iron protein (NifN, NifT, NifV, NifX, NifZ) | K01078                | acid phosphatase                             |
| K03385                     | cytochrome c-552   | K01093                | 4-phytase / acid phosphatase                 |
| K04561                     | nitric oxide reductase subunit B                                   | K01113                | alkaline phosphatase D                       |
| K04747 - K04748            | nitric oxide reductase protein (NorF, NorQ)                        | K03788                | acid phosphatase (class B)                   |
| K10535                     | hydroxylamine oxidase  | K09474                | acid phosphatase (class A)                   |
| K10944 - K10946            | ammonia monoxygenase (AmoABC)                                      |                       |  |
| K15864                     | nitrite reductase (NO-forming) / hydroxylamine reductase           |                       |  |
| K15876                     | cytochrome c-type protein  |                       |  |



# Chapter 3

---

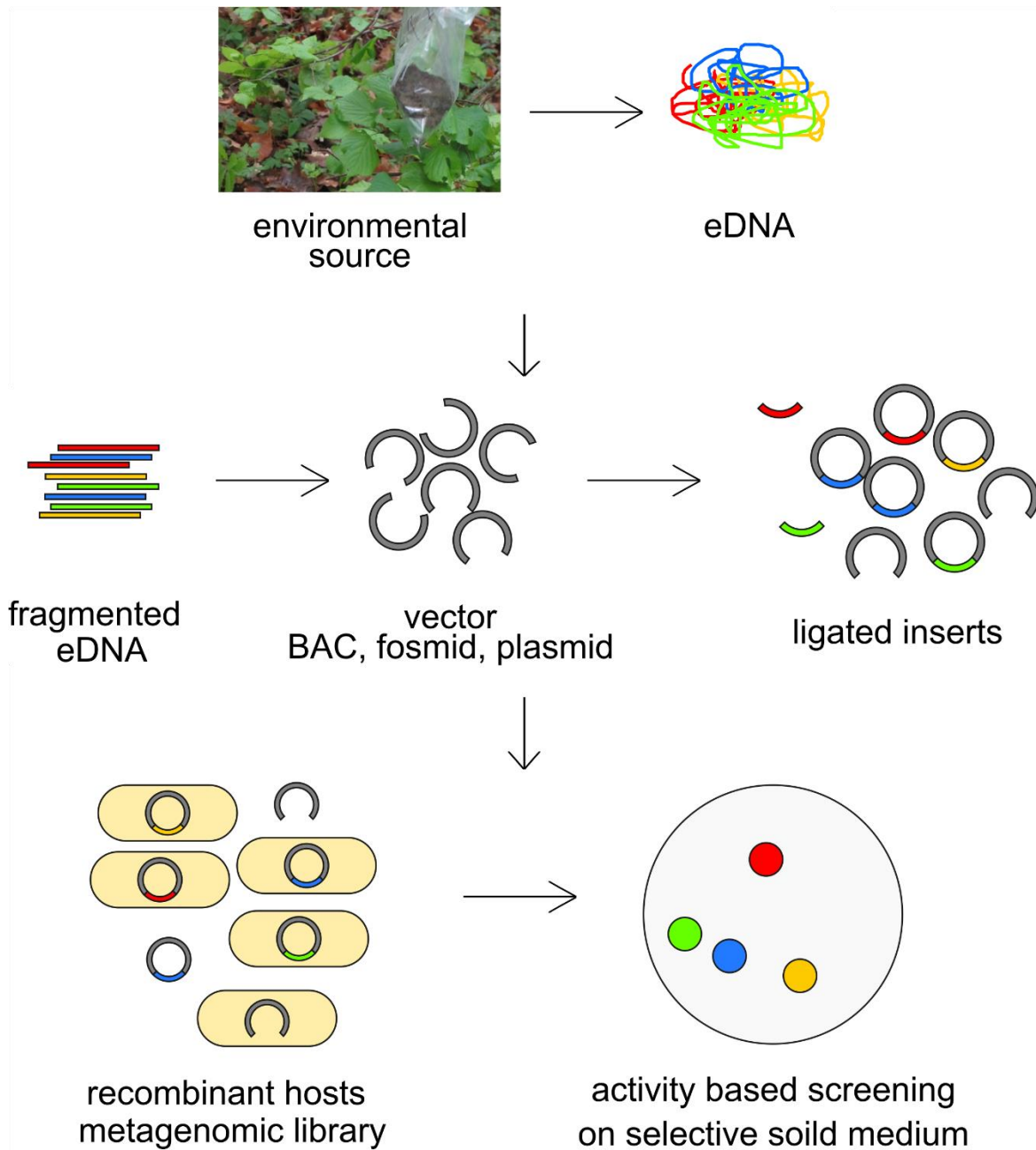
## Bioprospecting for novel biocatalysts

### 3.1 Harvesting the soil metagenome

Forests are a valuable hub of resources which benefit human society. From timber and food to recreational and cultural goods, forests have always been closely associated with human development. Soil microorganisms are important producers of natural products, from small antibiotic molecules to industrial biocatalysts used in environmental bioremediation (Demain, 2000). The diversity of uncultured bacteria far exceeds that of the cultured fraction (Banik and Brady, 2010). Current evaluations of prokaryotic diversity indicated that approximately 18.9 % of bacterial sequences originate from cultured organisms (Schloss et al., 2016). Consequently, the diversity of potentially useful bioactive compounds is also greater in this uncultured fraction (Banik and Brady, 2010). However, soil bacterial diversity remains largely untapped due to limitations in culturing methods (Torsvik and Øvreås, 2002). A study by Kaeberlein et al. (2002) proposed that this may be due to signaling mechanism between bacteria to identify an unfamiliar environment, even in the presence of nutrients. A recent innovation known as the Isolation chip or 'Ichip' seeks to improve parallel cultivation of bacteria from environmental samples (Nichols et al., 2010). The technique was subsequently important in the discovery of novel bacterium (*Eleftheria terrae*) expressing a new antibiotic, teixobactin, against methicillin-resistant *Staphylococcus aureus* (MRSA) and *Mycobacterium tuberculosis* (Ling et al., 2015a). Nevertheless, obtaining bioactive compounds from the environment remains a challenge.

Metagenomics, which describes the entire genetic information from an environmental community (Handelsman et al., 1998) has proved a powerful counter to this bottleneck. Current metagenomic tools have increased the potential for exploiting the forest microbiome by enabling culture-independent manipulation of bacteria, which are more accessible in this process (Lorenz and Eck, 2005; Rabausch et al., 2013). Products discovered through this approach include flavonoid-modifying glycosyltransferases (Rabausch et al., 2013), moderately thermostable and halotolerant cellulases (Ilmberger et al., 2012),  $\alpha$ -amylases, proteases, pectinases (Bashir et al., 2014), and antibiotic resistance biomolecules (Riesenfeld et al., 2004). Functional metagenomic screening also has wide applications beyond soil habitats, including marine habitats (Barone et al., 2014), extreme environments (Mirete et al., 2016), compost (Lopez-Lopez et al., 2014) and animals (Fang et al., 2012; Rashamuse et al., 2013). Thus, the potential for recovering novel biocatalysts has made functional metagenomic screening an indispensable tool in recent research methodologies.

Schloss and Handelsman (2003) outline two methods for the recovery of functional genes from metagenomes (Figure 3. 1): sequence-based and activity-based screening. In the sequence-based approach, primers are designed based on conserved regions of known genes and serve as anchors which flank the insert. After the construction of a metagenomic library, the primers hybridize with conserved regions and the gaps are filled through primer walking (Uchiyama and Miyazaki, 2009). Successful isolation of chitinases and dioxygenases were isolated with this approach (Hjort et al., 2010; Zaprasis et al., 2010). A limitation of this approach is the reduced potential for finding novel genes, if regions of homology in known sequences are highly conserved (Lorenz et al., 2002). For example, bacterial esterases (section 3.2.1) share several highly conserved motifs, while bacterial laccases (section 3.2.2) share very little sequence similarity and have few conserved domains. Activity-based functional screening is a simpler process whereby selection of the gene of interest relies on measuring direct activity and is sequence-independent (Schloss and Handelsman, 2003). A selective or indicator substrate, which in many cases induces a colour change in the colonies or surrounding medium, is introduced on solid agar plates. Subsequently, host cells containing the metagenomic library inserts are plated on the selective plates. Contact between the substrate and the expressed enzyme will induce a phenotypic change in the host cell, such as a zone of clearance around the colony (Coughlan et al., 2015). This is a highly useful approach as it allows for the discovery of new genes, which share no sequence similarity and is a common method for mining for novel enzymes (Lorenz et al., 2002; Mirete et al., 2016). In section 3.3 of this chapter, short-insert plasmid libraries were constructed and used in activity-based screening of lignocellolytic genes.



**Figure 3. 1. Outline of construction of metagenomic libraries from environmental DNA. Metagenomic libraries are pools of environmental DNA stored as fragments on extrachromosomal DNA (eDNA) such as plasmids, bacterial artificial chromosomes (BACs), cosmids or fosmids. The size of the library is dependent on the carrying capacity of the vector and the number of vectors carrying an insert +quality of eDNA. These vectors can then be replicated in a host strain to maintain or manipulate the library. Although several hosts and vectors, which require different practical approaches, are available for constructing metagenomic libraries, the key steps remain unchanged, and are outlined. Abbreviations: eDNA, environmental DNA; BAC, bacterial artificial chromosome. Image adapted from (Lorenz et al., 2002; Mirete et al., 2016).**

## 3.2 Functional screening for lignocellulolytic enzymes

Lignocellulose mass is derived from plant cell walls and is one of the most abundant raw biopolymers in nature (Bajpai, 2016). It comprises cellulose (40 %), hemicellulose (20 – 30%) and lignin (20 – 30%) in a complex matrix with other plant cell material which renders it recalcitrant to degradation (Chang et al., 2011; Chen, 2014). As an important material in renewable biofuel production, biocatalysts involved in its depolymerization are highly valued in industry (Obeng et al., 2017). Lignocellulose degradation occurs through the hydrolytic action of cellulases (endo/exoglucanases), hemicellulases (xylanases), laccases, pectinases, chitinases and esterases, among many others (Park et al., 2016; Tousehik et al., 2017). A short treatment of some of these enzymes and their mechanisms of action follows below.

### 3.2.1 Esterases

Esterases (EC 3.1.x) and lipases (EC 3.1.1.x) are a group of lipolytic enzymes involved in the breakdown of lipids into alcohol and free fatty acids (Figure 3. 2A). A distinction is generally made between carboxylesterases which hydrolyse water-soluble, short-chain glycerol esters (shorter than 6 carbons), and true lipases which hydrolyse longer-chain, water-insoluble triglycerides (Arpigny and Jaeger, 1999). Bacterial esterases are ubiquitous, intracellular or extracellular molecules, and have been isolated from environments as diverse as hot springs to deep-sea marine sediments (Lopez-Lopez et al., 2014).

Lipolytic enzymes are highly desired in biotechnology due to their wide substrate spectrum, high selectivity, high temperature and pH stability, and tolerance to salt and organic solvents. Additionally, some esterases exhibit stereoselective activity. Consequently, lipolytic biocatalysts are used in far-ranging applications such as enantiomeric drug production, additives in detergents, bioremediation agents in the environment or flavor development in the food industry (Bornscheuer, 2002; Coughlan et al., 2015; Garcia et al., 2015; Sharma and Kanwar, 2014). As the topic of biochemical characterization in Chapter 4, further insights into metabolic properties of esterases are provided in Dukunde et al. (Dukunde et al., 2017).

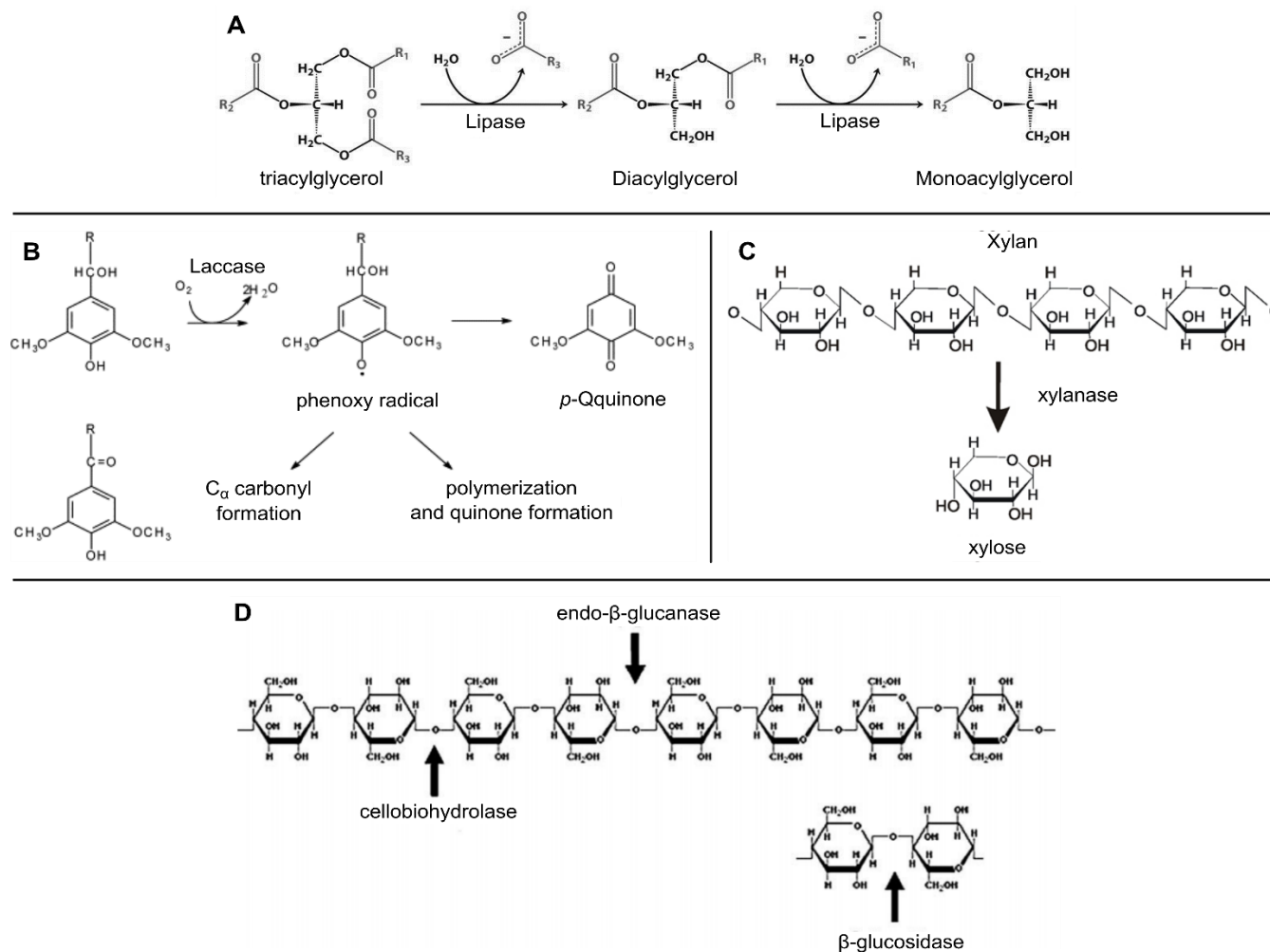


Figure 3. 2. Reaction mechanisms of select enzymes involved in lignocellulose degradation. A) hydrolysis of triacyl glycerol by lipases with concomitant release of water. B) Oxidation of a substrate molecule with oxygen as an electron donor. C) Xylanases hydrolyse glycosidic bonds between xylose monomers. D) Three types of cellulases act on different parts of a cellulose polymer. Images adapted from Berg et al. (2012), Minussi et al. (Minussi et al., 2002) and Kumar et al. (2008).

### 3.2.2 Laccases

Laccases (EC 1.10.3.2) are a subset of multicopper oxidases (MCOs), which act on phenolic and aromatic amines (Mtui, 2012). Although white-rot fungi can completely degrade lignin, bacterial laccases have emerged as more desirable catalysts due to the ease of manipulation (Pardo et al., 2012; Ricklefs et al., 2014).

Catalysis by laccases takes place via four copper-containing domains that form the active site; each domain catalyzes one electron oxidation of a substrate molecule with subsequent reduction of molecular oxygen to one water molecule (Figure 3. 2B). Functional screening media commonly employs guaiacol or 2,2'-azino-bis (3-ethylbenzothiazole-6-sulfonate) (ABTS), which form brown or blue oxidative rings around bacterial clones with potential activity, respectively (Fang et al., 2012; Šnajdr et al., 2013).

Bacterial laccases participate in an array of oxidative conversions for many substrates, including aromatic phenols, methoxyphenols, polyamines and aryl amines and some inorganic compounds (Ricklefs et al., 2014). In bacteria, such conversions aid in protection against UV light, harsh peroxidases and pigmentation, seen in the brown endospore coat of *Bacillus subtilis* (Dwivedi et al., 2011). As industrial biocatalysts, laccases do not produce toxic byproducts as their peroxidase counterparts, and bacterial laccases are reported to have better thermotolerance than fungal laccases (Santhanam et al., 2011). Laccases are used in the bleaching of textiles, Kraft pulp in paper making where they target lignin, and have been used in the clarification of wines and fruit juice (Arias et al., 2003; Tousehik et al., 2017; Virk et al., 2012). Additionally, the ability to degrade aromatic compounds makes them choice biocatalysts in decolorization of wastewater effluent (Virk et al., 2012). Functional metagenomic screening for laccases uses the sequence-based approach (Ausec et al., 2011; Fang et al., 2012; Jacquiod et al., 2014). However, in this thesis, we aim to use activity-based screening to increase the potential for mining novel laccases.

### 3.2.3 Lignocellulases

Cellulose degradation is catalyzed by four classes of cellulases, endoglucanases (EC 3.2.1.4), exoglucanases (EC 3.2.1.74), cellobiohydrolases (EC 3.2.1.91) and  $\beta$ -glucosidases (EC 3.2.1.21) (Saini et al., 2015). Hemicellulases are divided into endo-1,4- $\beta$ -xylanase (EC 3.2.1.8) and 1,4- $\beta$ -xylosidase (EC 3.2.1.37) and catalyze the hydrolysis of xylan, a major constituent of hemicellulose (Figure 3. 2C and D) (Alves-Prado et al., 2010). Their ability to degrade complex cellulosic polysaccharides make them important biocatalysts in renewable biofuel processes. Additionally, the use cellulases in animal feed can improve food digestion in livestock (Chang et al., 2011). Cellulose-free xylanases are used in biobleaching of kraft pulp in paper making, although the process is not well understood (Chakdar et al., 2016). However, xylanases and laccases are reported to have improved activity when used together in pulp biobleaching, which points to the presence of a feedback mechanisms between the two enzymes during lignocellulose degradation that is absent with cellulases (Virk et al., 2012). Cellulases also have had an increased role in the pretreatment of cotton fabrics (Araújo et al., 2008).

### 3.2.4 Objectives:

The aim of this chapter is to use forest soil-derived environmental DNA to construct short insert metagenomic plasmid libraries. Subsequently, metagenomic libraries will be used to conduct activity-based functional screening for potential lignocellulolytic biocatalysts.

### 3.3 Methods and materials

#### 3.3.1 General molecular cloning techniques

The following is an outline of general methods used, unless specified otherwise. Molecular cloning techniques, unless indicated otherwise, are based on or modified from Sambrook et al. (1989).

##### 3.3.1.1 Evaluating DNA and cell cultures

Agarose gel electrophoreses were performed by using 1 % agarose in 1X TAE buffer (50 mM tris, 20 mM acetic acid, 1mM EDTA) at 90 V and 400 mA. Gel images were documented with the ImageQuant 100 (GE Healthcare, CT, USA), at 320 nm. Quantitative and qualitative evaluation of DNA was performed with the NanoDrop Spectrometer ND1000 (Thermo Scientific, MA, USA). Growth of cell cultures was monitored by measuring optical density at 600 nm (OD<sub>600</sub>).

##### 3.3.1.2 Competent cells and bacterial transformation

Preparation of electrocompetent *E. coli* DH5 $\alpha$  (genotype; *F*-  $\phi$ 80*lacZ* $\Delta$ *M15* $\Delta$  (*lacZYA-argF*) *U169 endA1 recA1 hsdR17 (rk-, mk+) supE44 thi -1 gyrA96 relA1 phoA*), selected as hosts for subsequent metagenomic plasmid libraries, was performed as follows: a frozen cell pellet (– 80 °C) was used to inoculate 5 mL of LB medium (lysogeny broth, 10 g NaCl, 10g Tryptone, 5 g yeast extract per liter distilled water (dH<sub>2</sub>O) ) and grown overnight, at 37 °C on a shaking incubator at 150 rpm. The 5 mL starter culture was used to inoculate 250 mL of LB and grown at 30 °C, with shaking in an Aquatron water bath (Infors AG, Switzerland) to an OD<sub>600</sub> between 0.5 – 0.75. The culture was cooled on ice to 4 °C and centrifuged to pellet the cells, at 5,000 rpm and 4 °C, for 10 min. The supernatant was discarded, and the pellet was resuspended in 1 volume (vol) dH<sub>2</sub>O. Cells were centrifuged as previously described and the supernatant was discarded. This was followed by two more washing and centrifugation cycles. The cell pellet was resuspended in 10 % (v/v) glycerol, centrifuged for 10 min at 6,200 rpm and 4 °C, and the supernatant discarded. The resultant competent pellet was resuspended in 500  $\mu$ L 10 % (v/v) glycerol and stored as 40  $\mu$ L aliquots. Aliquots were flash-frozen with liquid nitrogen and stored at – 80 °C until required.

Transformation of *E. coli* strains was performed by mixing 40  $\mu$ L of competent cells with 75 ng of recombinant vector DNA and incubated on ice for 10 min. The *E. coli*/eDNA mix was transferred to an electroporation cuvette (Peqlab Biotechnologie GmbH, Erlangen, Germany) and electroporated with the Gene Pulser II electroporator (Bio-Rad Laboratories GmbH, Munich, Germany) set at 200  $\Omega$ , 25  $\mu$ F, 2.4 kV, to give a typical time constant of 4.95 ms. Cells were immediately rescued with 1 mL LB and incubated with shaking at 37 °C for 1 hour. Subsequently, 300  $\mu$ L of the cell culture was plated on LB agar (15g agar/liter LB) supplemented with kanamycin (Kan, 50  $\mu$ g/ $\mu$ L), along with additional screening substrate for the specific gene product under investigation. Plates were incubated at 37 °C to grow, and later transferred to 30 °C up to 12 days for activity-based screening.

##### 3.3.1.3 Plasmid DNA extraction and restriction analysis

Preparation of plasmid DNA from potential positive colonies (displaying halos on agar plates) was performed by selecting potential clones and culturing them overnight in 5 mL LB at 37 °C on a shaking incubator. Subsequently, plasmid DNA was extracted by using the NucleoSpin II kit (Macherey Nagel



GmbH Co. KG, Düren, Germany), as directed by the manufacturer. Restriction analysis of plasmid DNA was performed with the restriction endonuclease EcoRI (MBI Fermentas, St Leon Rot, Germany) to release the cloned fragment. Reactions were performed in 10  $\mu$ L volumes (1  $\mu$ L of 10-fold EcoRI buffer, 0.5  $\mu$ L EcoRI, 70 – 170 ng plasmid DNA, made to volume with dH<sub>2</sub>O) and incubated for 2 h at 37 °C and analyzed by agarose gel electrophoresis as described in 3.3.1.

### 3.3.2 Function-based screening

#### 3.3.2.1 Construction of metagenomic plasmid libraries

Short-insert plasmid libraries were constructed using microbial environmental DNA from soils in the litter and horizon layers in the Hainich national park (Figure 2.1), according to the procedure described by Nacke et al., (Nacke et al., 2011). For each library, environmental DNA was isolated from 10 g of soil using the MoBio Power Max Soil DNA extraction kit (MoBio Laboratories, Carlsbad, CA, USA), as directed by the manufacturer, with the exception that the bead-beating step was doubled to increase DNA fragmentation. Agarose gel electrophoresis was used to resolve fragments ranging between 3 – 8 kb, which were subsequently purified with the PeqGold Gel Extraction kit (PEQLAB Biotechnologie GmbH). Blunt-end polishing of purified fragments was performed with T4 DNA polymerase (MBI Fermentas), as directed by the manufacturer. Subsequently, DNA fragments were purified by using SureClean solution (Bioline GmbH, Luckenwalde, Germany) with the following modifications to improve DNA precipitation: incubation and centrifugation steps were increased to 30 min and 1 h, respectively. Fragments were resuspended in 35  $\mu$ L dH<sub>2</sub>O. To facilitate TA cloning, deoxyadenine (dATP) was added to 3' termini by adding the resuspended DNA (35  $\mu$ L) to 7  $\mu$ L of 10-fold Taq DNA polymerase buffer (MBI Fermentas), 6  $\mu$ L MgCl<sub>2</sub> (of 25 mM), 1  $\mu$ L dATP (100mM) solution, 1 $\mu$ L Taq DNA polymerase (5 U), and 20  $\mu$ L dH<sub>2</sub>O. The reaction was incubated at 72 °C for 30 min, purified with SureClean solution as previously described and resuspended in 15  $\mu$ L dH<sub>2</sub>O. The purified fragments were dephosphorylated by using Antarctic phosphatase (New England Biolabs, Ipswich, MA, USA), according to manufacturer directions. Resulting DNA fragments were ligated into pCR-TOPO-XL via complementary TA cloning using the TOPO XL PCR cloning kit (Invitrogen GmbH, Karlsruhe, Germany), according to manufacturer directions. Recombinant plasmids were used to transform competent *E. coli* DH5 $\alpha$ , as described in 3.3.1. Transformants were plated onto selective plates for activity-based screening, described below.

#### 3.3.2.2 Screening of lignocellulolytic enzymes

All activity-based screening was performed on LB<sub>Kan</sub> agar plates (50  $\mu$ g/ $\mu$ L), using *E. coli* DH5 $\alpha$  as a host strain. Agar plates were supplemented with guaiacol (0.01 % v/v), tributyrin (1 % v/v) to screen for laccase and lipolytic activity, respectively. To screen for potential xylanase and cellulose activity, approximately 0.165 g (dissolved in 2 mL ethanol) of AZCL-Xylan and AZCL-HE-Cellulose (Megazyme International, Ireland) per liter LB were used, respectively. Inoculated plates were incubated as described in 3.3.1 and activity was monitored through the formation of zones of (dis)coloration around potentially positive clones. That is, a brown ring of oxidation was expected around potential laccase-positive colonies, clear zones around lipolytic colonies, and blue rings around cellulose and xylanase-positive colonies.

### 3.3.2.3 Sequence analysis

Recombinant plasmids were prepared from potentially positive clones as described in section 3.3.1 and sequenced by the Göttingen Genomics Laboratory (Göttingen, Germany). Open reading frames (ORFs) were predicted from sequenced inserts with ORFinder (<https://www.ncbi.nlm.nih.gov/orffinder/>), FramePlot 4.0beta (<http://nocardia.nih.gov/jp/fp4/>) and SignalIP 4.0 (<http://www.cbs.dtu.dk/services/SignalP/>), by locating the presence of potential start and stop codons, signal peptides, and ribosomal binding sites (RBS) (Altschul et al., 1990; Petersen et al., 2011). Sequence similarity with existing proteins were performed by comparing amino acid sequences using the BLASTp suite (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) of the Basic Local Alignment Tool (BLAST) (Altschul et al., 1990). Multiple sequence alignment of amino acid sequences homologous to putative functional ORFs were retrieved from the National Center for Biotechnology Information (NCBI) database and aligned using MUSCLE (Edgar, 2004). Annotation of aligned sequences was performed with EsPript 3.0 (Robert and Gouet, 2014). Additional parameters were calculated with the ProtParam tool at Expasy (Gasteiger, E. et al., 2005).

### 3.4 Results and discussion

#### 3.4.1 Construction metagenomic plasmid libraries

Environmental microbial DNA from forest soil was used for functional screening of lignocellulolytic activity by constructing short-insert metagenomic libraries. Six plasmid libraries were constructed and used to screen for potential lipolytic, laccase, xylanase and cellulose activity. A full description of plasmid library properties is provided in Table 3. 1.

**Table 3. 1 Properties of short-insert plasmid libraries. All samples derived from a beech mono stand (plot 10) from litter (L) and horizon A (H) layers. Samples were collected in spring (Sp), summer (Su) and autumn (Au) from the Hainich National park. All inserts were cloned in the pCR-XL-TOPO vector (Invitrogen). Average insert size is based on the average insert sizes of 20 positive clones from each library.**

| Library | Number of clones | Average insert size (kb) | Insert frequency (%) | Estimated library size (kb) |
|---------|------------------|--------------------------|----------------------|-----------------------------|
| SpH10   | 10,379           | 5.5                      | 86                   | 51.3                        |
| SpL10   | 22,048           | 1.6                      | 64                   | 22.2                        |
| SuH10   | 40,066           | 4.5                      | 82                   | 146.6                       |
| SuL10   | 38,400           | 3.4                      | 85                   | 110.6                       |
| AuH10   | 28,289           | 4.6                      | 86                   | 110.9                       |
| AuL10   | 98,914           | 2.1                      | 83                   | 170.3                       |

The resulting plasmid libraries contained 10,379 – 98,914 clones, which represents successful library construction considering the high insert frequency (> 80 %). Comparable results in clone numbers have been reported for metagenomic plasmid libraries (Berlemont et al., 2013; Steele et al., 2009). As plasmids have a small insert carrying capacity (< 10 kb), average insert size in the constructed libraries followed this trend, ranging from 1.6 – 5.5. kb (Henne et al., 2000). The likelihood of isolating intact genes, entire operons or even gene clusters increases with insert size and increases the potential diversity of the library (Kakirde et al., 2010). Nacke et al. (2011) constructed a plasmid library from forest soil, with a large insert size (9.4 kb) and high insert frequency (96 %); confirming that insert size did not affect insert frequency. Nevertheless, it is possible to obtain full genes encoding smaller proteins (such as lipolytic, amyolytic or antibiotic resistance genes) with the plasmid libraries constructed here (Yun et al., 2004).

#### 3.4.2 Screening for lignocellulolytic enzymes

Metagenomic libraries were subjected to activity-based screening for lignocellulolytic activity, which is expected to be abundant in forest soils, by using selective substrates on agar plates.

##### 3.4.2.1 Lipolytic activity

Lipolytic activity was detected via the formation of rings around colonies on turbid tributyrin agar (Figure 3. 3), indicating substrate hydrolysis by potential lipases or esterases. Plasmid preparations

from clones harboring activity were analyzed by restriction analysis and sequenced further to determine the nature of the activity-inducing insert.



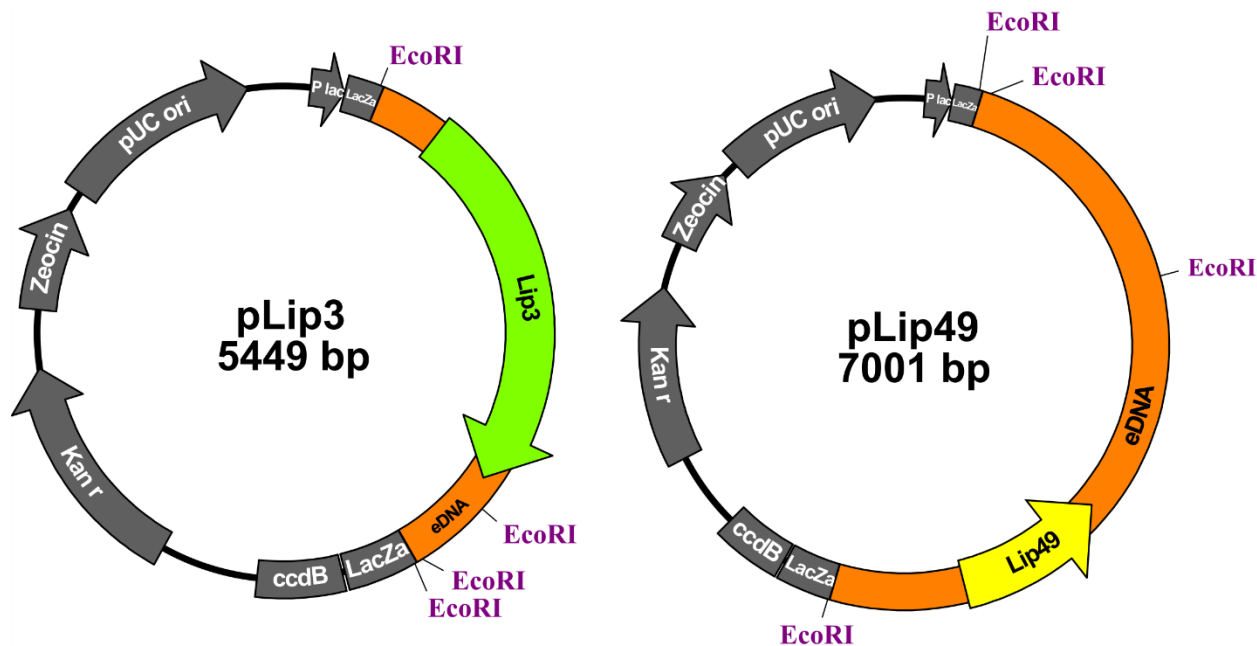
**Figure 3. 3 Potential lipolytic activity resulting in zones of clearance around *E. coli* DH5 $\alpha$  clones (indicated by a black arrow), conferred by metagenomic insert.**

Functional screening recovered 41 bacterial clones with potential lipolytic activity (Table 3.2). Insert size for these clones ranged from 1,350 – 7,250 bp, and demonstrates the range of lipase-encoding genes, which falls in this range (Sudan and Vakhlu, 2013). Interestingly, most positive clones were detected in libraries from the summer (32 clones) plasmid libraries, compared to spring (5 clones) and autumn (4 clones) libraries. The effect of season, and indeed other environmental factors, has not been shown to affect activity-based screening; nevertheless, it is possible that since environmental patterns affect bacterial function, such factors also contribute to cloning bias in metagenomic libraries.

**Table 3. 2 Table of potential lipolytic clones from a forest soil metagenomic library. Restriction fragments were generated with EcoRI. Gene fragments are named after bacterial clones from which they were isolated. Bacterial clones are named after the screening plates on which they were first detected and has no nomenclature significance.**

| Library | Fragment | Restriction fragments (bp) |      |      |      |      | Total size (bp) |
|---------|----------|----------------------------|------|------|------|------|-----------------|
| SuL10   | 3        | 750                        | 2800 |      |      |      | 3550            |
|         | 5        | 750                        | 2800 |      |      |      | 3550            |
|         | 18       | 750                        | 2800 |      |      |      | 3550            |
|         | 27       | 300                        | 375  | 600  | 1700 | 3000 | 5975            |
|         | 49       | 1100                       | 2200 |      |      |      | 3300            |
|         | 51       | 750                        | 2800 |      |      |      | 3550            |
|         | 53       | 300                        | 375  | 600  | 1700 | 3000 | 5975            |
|         | 72       | 1000                       | 2800 | 3000 |      |      | 6800            |
|         | 81       | 300                        | 375  | 600  | 2000 | 3000 | 6275            |
|         | 85       | 250                        | 1500 | 3200 |      |      | 4950            |
|         | 86       | 750                        | 2800 |      |      |      | 3550            |
|         | 54_1     | 750                        | 2800 |      |      |      | 3550            |
|         | 54_2     | 750                        | 2800 |      |      |      | 3550            |
|         | 54_3     | 2100                       |      |      |      |      | 2100            |
|         | 61_1     | 1350                       |      |      |      |      | 1350            |
|         | 61_2     | 250                        | 1500 | 3200 |      |      | 4950            |
|         | 61_2     | 1350                       |      |      |      |      | 1350            |
|         | 61_3     | 750                        | 3000 |      |      |      | 3750            |
|         | SuH10    | 1                          | 1750 | 3000 |      |      |                 |
| 2       |          | 200                        | 1150 | 3000 |      |      | 4350            |
| 3       |          | 250                        | 1700 |      |      |      | 1950            |
| 4       |          | 200                        | 1150 | 3000 |      |      | 4350            |
| 5       |          | 1000                       | 2000 | 3200 |      |      | 6200            |
| 6       |          | 1600                       |      |      |      |      | 1600            |
| 7       |          | 250                        | 1000 | 3000 |      |      | 4250            |
| 8       |          | 750                        | 1500 | 4000 |      |      | 6250            |
| 10      |          | 600                        | 800  | 900  | 1500 |      | 3800            |
| 11      |          | 4500                       |      |      |      |      | 4500            |
| 17      |          | 250                        | 500  | 4000 |      |      | 4750            |
| 23      |          | 200                        | 1150 | 3000 |      |      | 4350            |
| 24      |          | 1000                       | 1500 | 1750 |      |      | 4250            |
| 25      |          | 200                        | 1150 | 3000 |      |      | 4350            |
| SprH10  |          | 1                          | 250  | 3000 | 4000 |      |                 |
| SprL10  | 1        | 2250                       |      |      |      |      | 2250            |
|         | 2        | 2250                       |      |      |      |      | 2250            |
|         | 3        | 2250                       |      |      |      |      | 2250            |
|         | 5        | 750                        | 900  | 1750 |      |      | 3400            |
| AuH10   | 1        | 1200                       | 4500 |      |      |      | 5700            |
|         | 2        | 1900                       |      |      |      |      | 1900            |
|         | 4        | 600                        | 2700 |      |      |      | 3300            |
|         | 5        | 1000                       | 5000 |      |      |      | 6000            |

Two lipolytic inserts from the SuL10 library were completely sequenced. Subsequent analysis revealed an ORF in each insert, putatively named Lip3 and Lip49 (Figure 3. 4). Respectively, the genes encode 445 aa (1338 bp) and 230 aa (693 bp) polypeptides of approximately 48 kDa and 24 kDa. Both genes lack upstream RBS, indicating that protein translation is, to an extent, facilitated by mRNA secondary structure (Babitzke and O'Connor, 2017; Scharff et al., 2011). No signal peptide sequences in insert sequences were observed and may reveal Lip3 and Lip49 to be intracellular proteins, which has been reported in esterases and lipases (Javed et al., 2018); however, these could be present in upstream translation sequences, which are absent.



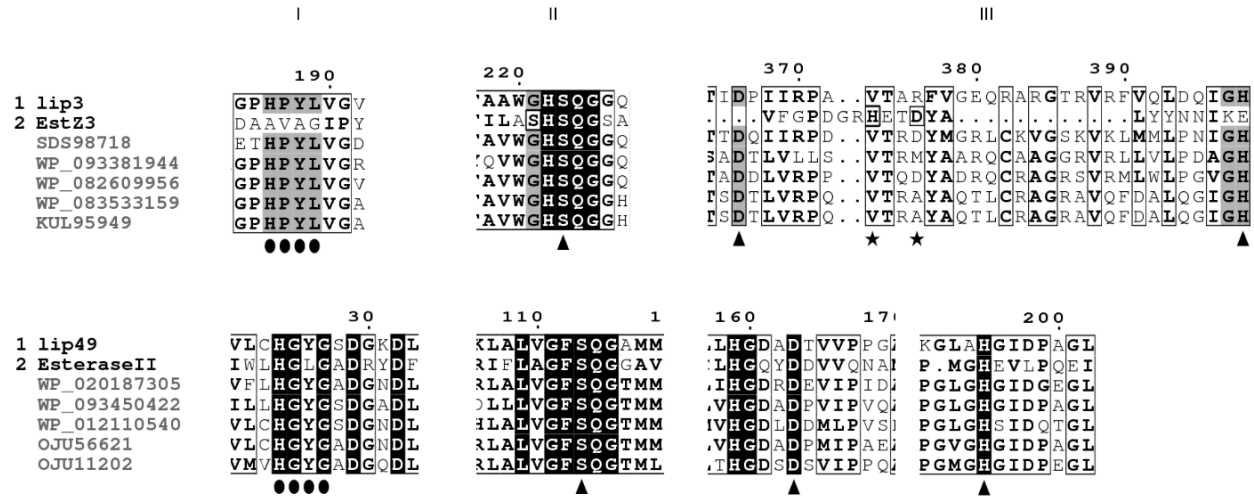
**Figure 3. 4. Lipolytic inserts for Lip3 and Lip49 from a metagenomic library derived from forest soil. Elements (gray) of pCR-TOPO-XL plasmid form the backbone. Inserts from eDNA (orange) are shown with predicted ORFs for Lip3 (green) and Lip49 (yellow). Direction of ORF indicates the reading frame of the protein. Restriction sites for EcoRI, used for size determination, are shown.**

Searches conducted with BLAST reveal that proteins closely related to Lip3 and Lip49 belong to the esterase and lipase families in the  $\alpha/\beta$  hydrolase superfamily. The highest sequence similarity ranged between 54 % and 63 % (for Lip3 and Lip49, respectively), indicating that the lipases may encode novel lipolytic proteins (Table 3.3). Closest related organisms generally belong the *Rhizobiales* order of *Proteobacteria*. Exceptions were *Variovorax* and *Sphingomonas*, which belong to *Burkholderiales* and *Sphingomonadales*. Given that these taxa are commonly associated with plant and soil environments (Laforest-Lapointe et al., 2016; Romanowicz et al., 2016), it follows that Lip3 and Lip49 are expressed by similar bacteria from soil. *Alphaproteobacteria* was one of the dominant taxa in our taxonomic datasets, therefore it is not unexpected that functional genes also derive from this group.

**Table 3. 3 BLAST results showing genes and corresponding organisms closely related to putative genes Lip3 and Lip49, derived from a forest soil metagenomic library.**

| Gene  | Related protein<br>(Accession no.)       | Organism                            | Query<br>Cover (%) | E-value   | Identity<br>(%) |
|-------|--|-------------------------------------|--------------------|-----------|-----------------|
| Lip3  | Lipase<br>(WP_082609956.1)               | <i>Bosea</i> sp. Root381            | 80                 | 5.19e-107 | 54              |
|       | Lipase<br>(WP_083533159.1)               | <i>Bosea</i> sp. WAO                | 86                 | 4.31e-114 | 53              |
|       | Lipase<br>(KUL95949.1)                   | <i>Bosea</i> sp. WAO                | 87                 | 2.62e-114 | 52              |
|       | Secretory lipase<br>(SDS98718.1)         | <i>Bradyrhizobium canariense</i>    | 80                 | 6.24e-108 | 52              |
|       | Lipase<br>(WP_093381944.1)               | <i>Variovorax</i> sp. OV329         | 80                 | 2.16e-96  | 52              |
| Lip49 | Phospholipase<br>(OJU56621.1)            | Unclassified<br>Alphaproteobacteria | 98                 | 1.25e-91  | 63              |
|       | Phospholipase<br>(OJU11202.1)            | Unclassified<br>Alphaproteobacteria | 94                 | 3.97e-84  | 60              |
|       | Esterase<br>(WP_093450422.1)             | <i>Sphingomonas</i> YR710           | 90                 | 7.12e-72  | 60              |
|       | Phospholipase<br>(WP_012110540.1)        | <i>Parvibaculum lavamentivorans</i> | 92                 | 5.52e-82  | 59              |
|       | Hypothetical protein<br>(WP_020187305.1) | <i>Methylopila</i> sp. 73B          | 93                 | 1.93e-78  | 59              |

Lip3 and Lip49 share key conserved sequences with their corresponding close relatives (Figure 3. 5). The catalytic serine (Ser<sup>228</sup>) in Lip3 occurs in the conserved motif GHSQG, while Ser 113 in Lip49 is found within GFSQG. The catalytic triad also consist of conserved aspartate (Asp) and histidine (His) residues for Lip3 (Asp<sup>366</sup>, His<sup>398</sup>) and Lip49 (Asp<sup>113</sup>, His<sup>195</sup>), respectively. Lip3 could be a true lipase from Family I, according to the original classification by Arpigny and Jaeger (1999) as the GHSQG motif was identified in this family. However, as the number of esterase families has grown, the motif was since identified in newer non-canonical esterase families (Family VII and Family 14) (Elend et al., 2006; Zarafeta et al., 2016). Lip49 has the GFSQG motif which is found in classical Family VI esterases, which catalyse short chain carbon substrates; however, little is known about their catalytic spectrum (Arpigny and Jaeger, 1999; Nacke et al., 2011). Further experiments to determine the substrate affinity of proteins encoded by Lip3 and Lip49 are required to confirm whether the proteins belong in their respective families.



**Figure 3. 5. Multiple sequence alignment of Lip3 and Lip49 showing conserved sequence blocks with closely related esterases. Conserved residues have a black background. For Lip3, conserved sequences not present in EstZ3 have a gray background. Potential residues for oxyanion holes are marked with filled circles. Catalytic triad residues are marked by filled triangles. Catalytic aspartate and histidine residues in EstZ3 are marked with filled stars. In addition to the closest related protein, a representative of each family of esterases was included, based on the classification scheme outlined by Arpigny and Jaeger (1999).**

### 3.4.2.2 Further screening for lignocellulolytic function

Activity-based screening conducted for laccase, xylanase and cellulases did not yield positive clones. The limitations encountered are generally unrelated to the lack of the corresponding genes in the soil environment, as function-based screening is not an indicator of abundance (Sudan and Vakhlu, 2013). Several, interdependent factors such as cloning host/vector compatibility, protein expression quality and quality of source environmental DNA can introduce heavy bias in the process (Gaida et al., 2015; Uchiyama and Miyazaki, 2009). Prospects for improving gene-discovery from the forest environment would involve construction of larger insert libraries in fosmids which have a higher carrying capacity (Nacke et al., 2012), and experimenting with other bacterial hosts (Cheng et al., 2014; Lam et al., 2015). Despite the absence of a diverse range of lignocellulolytic genes in this study, a function-based approach was successful in the isolation of several potential lipolytic clones and identification of novel lipase and carboxylesterase genes.



# Chapter 4

---

## A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum

Amélie Dukunde<sup>1</sup>, Dominik Schneider<sup>1</sup>, Mingji Lu<sup>1</sup>, Silija Brady<sup>1,2</sup>, Rolf Daniel<sup>1</sup>

Part of this chapter is published in:

Dukunde, A., Schneider, D., Brady, S., Daniel, R. (2017). A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum. *Biotechnol Lett* 39: 577 – 587

Author contributions:

Conceived and designed the experiments: **AD**, SB, RD

Performed the experiments: **AD**

Analysed the results: **AD**, ML

Wrote the publication: **AD**, DS, RD

# A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum

Amélie Dukunde · Dominik Schneider · Mingji Lu ·  
Silja Brady · Rolf Daniel 

Received: 8 December 2016 / Accepted: 21 December 2016  
© Springer Science+Business Media Dordrecht 2016

## Abstract

**Objectives** To investigate the properties of a novel metagenome-derived member of the hormone-sensitive lipase family of lipolytic enzymes.

**Results** A forest soil metagenome-derived gene encoding an esterase (Est06) belonging to the hormone-sensitive lipase family of lipolytic enzymes was subcloned, heterologously expressed and characterized. Est06 is a polypeptide of 295 amino acids with a molecular mass of 31 kDa. The deduced protein sequence shares 61% similarity with a hypothetical protein from the marine symbiont Candidatus *Entotheonella* sp. TSY1. Purified Est06 exhibited high affinity for acyl esters with short-chain fatty acids, and showed optimum activity with *p*-nitrophenyl valerate (C5). Maximum enzymatic activity was at 50 °C and

pH 7. Est06 exhibited high stability at moderate temperatures by retaining all of its catalytic activity below 30 °C over 13 days. Additionally, Est06 displayed high stability between pH 5 and 9. Esterase activity was not inhibited by metal ions or detergents, although organic solvents decreased activity.

**Conclusions** The combination of Est06 properties place it among novel biocatalysts that have potential for industrial use including low temperature applications.

**Keywords** Carboxylesterases · Esterase · Family IV esterase · Hormone-sensitive lipase family · Soil metagenome

**Electronic supplementary material** The online version of this article (doi:10.1007/s10529-016-2282-1) contains supplementary material, which is available to authorized users.

A. Dukunde · D. Schneider · M. Lu · S. Brady ·  
R. Daniel (✉)  
Department of Genomic and Applied Microbiology and  
Göttingen Genomics Laboratory, Institute of  
Microbiology and Genetics, Georg-August University  
Göttingen, Grisebachstr. 8, 37077 Göttingen, Germany  
e-mail: rdaniel@gwdg.de

**Present Address:**  
S. Brady  
Bayer AGET-TD-UP Biochemistry & Biocatalysis, Building  
B310, 207, 51368 Leverkusen, Germany

## Introduction

Metagenomics have provided a wealth of previously inaccessible genetic material from uncultivable organisms across a wide range of environments (de Castro et al. 2014). In addition to expanding the genetic diversity of known organisms, metagenomics has advanced the discovery of novel proteins, that can be used i.e., as catalysts for laundry, pharmaceutical or bioremediation applications (Selvin et al. 2012; Facchin et al. 2013). Esterases or carboxylesterases (EC 3.1.1.1), which catalyze synthesis and hydrolysis of short-chain glycerol esters, are an important group of industrially relevant biocatalysts. Several esterases

have been isolated and identified by metagenomic approaches using environmental DNA from diverse environments ranging from bovine rumen (Rodríguez et al. 2015) to cryopegs in permafrost (Novototskaya-Vlasova et al. 2012). Esterases belong to a large group of lipolytic enzymes, that includes true lipases. A combination of cultivation dependent and independent approaches have expanded the original esterase classification of Arpigny and Jaeger (1999) from seven families to 15 families, the most recent being family XV (Charbonneau and Beauregard 2013).

A distinguishing feature of esterases is the presence of a GX SXG consensus sequence that contains the catalytic serine residue. This serine forms the nucleophilic part of a catalytic triad, together with aspartic acid or glutamine and histidine (Bornscheuer 2002; López-López et al. 2014). A substrate binds to the active site at the serine residue to form a tetrahedral intermediate that is stabilised by Asp/Glu and His residues. An acyl-enzyme complex is formed by elimination of alcohol, which is then hydrolysed by a water molecule to generate the final tetrahedral intermediate and free enzyme (Bornscheuer 2002). The diversity in primary amino acid sequences of esterases still results in a conserved tertiary structure that typically, consists of a *N*-terminal  $\alpha$ -helical domain and a central barrel, catalytic domain of parallel  $\beta$ -strands known as the  $\alpha/\beta$  hydrolase fold (Nardini and Dijkstra 1999; Lenfant et al. 2013).

As a group, esterases exhibit broad substrate range, temperature and pH spectra, and tolerance to organic solvents, metal ions and salts to varying degrees. Individually, they are highly selective enzymes, capable of chemo-, regio- and stereo-selective activity (López-López et al. 2014; Ma et al. 2014). Therefore, esterases and related lipolytic enzymes are important biocatalysts for organic synthesis, stereospecific drug production (Levisson et al. 2007), bioremediation of waste from oil refineries (Martínez-Martínez et al. 2014), detergent preparations (Zhang et al. 2014) and flavour development in the food industry (Mohamed et al. 2013).

In this study, a metagenome-derived gene encoding an esterase belonging to the hormone-sensitive lipase (HSL) family and the corresponding protein have been subcloned, heterologously produced and characterized. The gene was derived from a metagenomic plasmid (pLE06), which conferred lipolytic activity. The plasmid originated from a forest soil metagenome (Nacke et al. 2011). Lipolytic enzymes of the HSL

group share sequence and structural homology to the catalytic domain of the mammalian hormone-sensitive lipase (Østerlund 2001). Subfamilies in this group possess either a GDSA or GTSA as a catalytic motif (Li et al. 2014), and were among the first esterases assigned to the  $\alpha/\beta$  hydrolase superfamily (Lenfant et al. 2013).

## Materials and methods

### Bacterial strains and plasmids

The lipolytic recombinant strain *E. coli*/pLE06 harbors the metagenomic plasmid pLE06 including the putative esterase gene *est06*. The clone was derived from a functional screen of a metagenomic plasmid library, which was constructed from forest soil of Schorfheide-Chorin (Germany). Soil sampling, metagenomic library construction and identification of lipolytic clones was performed by Nacke et al. (2011). The plasmid pLE06 was used as template for amplification of *est06*. The corresponding nucleotide sequence is available at the National Center for Biotechnology Information (NCBI) gene bank under the accession number HQ156905. *Escherichia coli* strain BL21(DE3) and pET101/D (Invitrogen) were used as expression host and vector, respectively. Detailed maps of plasmids are provided in Supplementary Fig. 1.

### Identification and analysis of *est06* sequence

Amino acid sequences of esterases homologous to the deduced gene product of *est06* were retrieved from the NCBI database and aligned using MUSCLE (Edgar 2004). Secondary and tertiary structure prediction was performed with I-TASSER (Zhang 2008; Roy et al. 2010; Yang and Zhang 2015) by modelling Est06 along the structure of E40 (Protein Data Bank code: 4XVC\_A; Li et al. 2015), an esterase that shares 55% amino acid sequence similarity with Est06. Annotation of aligned sequences was performed with EsPript 3.0 (Robert and Gouet 2014). Additional parameters were calculated with the ProtParam tool at ExPasy (Gasteiger et al. 2005).

### Subcloning, overexpression and purification of Est06

The primer pair 5'-CACCATGTCGCAACAACAACTCC-3' and 5'-CGCGGCGTTGGCGCGGACGA

A-3' were used to amplify the 888 bp fragment of *est06* from pLE06. The *est06* PCR product was purified by using agarose gel electrophoresis and the Qiaquick gel extraction kit, according to the manufacturer's instructions (Qiagen). Subsequently, the purified product was cloned into expression vector pET101/D (Invitrogen) as recommended by the manufacturer. In this way, also sequences encoding a V5 epitope and a His<sub>6</sub>-tag were added to the 3'-end of the gene. The resulting recombinant vector (pEST06) was used to transform chemically-competent *E. coli* BL21 (DE3) (Invitrogen). Transformants were grown on lysogeny broth (LB) agar plates containing 15 g agar l<sup>-1</sup>, 100 mg ampicillin l<sup>-1</sup>.

For expression of *est06* gene, single colonies were grown in 30 ml LB medium supplemented with 100 mg ampicillin l<sup>-1</sup> overnight with shaking at 30 °C. Expression was initiated by inoculation of a 1 l-culture with 2% (v/v) of the overnight culture. At an OD<sub>600</sub> of 0.4, expression was induced by the addition of 0.5 mM IPTG. Cells were harvested by centrifugation at 7000×g for 10 min at 4 °C after 6 h of incubation and washed twice in 50 mM sodium phosphate buffer (pH 7). The cells were suspended in LEW buffer (Macherey-Nagel, Düren, Germany) and lysed on ice by sonication with a UPS200S homogeniser (Hielscher Ultrasonics GmbH, Teltow, Germany).

Crude cell extract was cleared by centrifugation at 7000×g for 1 h. Purification of Est06 was initially carried out with the Protino-NI-TED packed columns as recommended by the manufacturer (Macherey-Nagel, Düren, Germany). Resulting elution fractions were pooled and dialysed overnight against 20 mM Tris/HCl buffer (pH 8) at 4 °C, and finally concentrated with the Vivaspin 10,000 Da exclusion columns (Sartorius AG). Further purification was performed by anion exchange chromatography by using a SOURCE 15Q 4.6/100 PE tricorn exchanger column and Äkta FPLC system (GE Healthcare Europe GmbH, München) over an elution gradient of 1 M NaCl/Tris/HCl (pH8). Active fractions were spectrophotometrically detected with *p*-nitrophenyl butyrate as substrate and, subsequently, pooled and concentrated with a Vivaspin column and 50 mM sodium phosphate (pH 7) as described above. The purity of Est06 was confirmed with SDS-PAGE. Protein concentrations were determined by the Bradford method.

### Standard enzyme assays

The activity of Est06 was quantified by measuring initial rates of hydrolysis of *p*-nitrophenyl esters (*p*NP) and subsequent production of *p*-nitrophenol at 405 nm. Standard reactions were conducted at 25 °C in 1 ml containing 50 mM sodium phosphate buffer (pH 7) and 1 mM *p*-nitrophenyl valerate (C5) as substrate. Unless otherwise stated, the reaction was initiated by adding Est06 (2 U) to the assay reaction and monitored at 405 nm. The extinction coefficient of *p*-nitrophenol under the above-mentioned conditions was experimentally determined to be 8726 M<sup>-1</sup> cm<sup>-1</sup> (Hotta et al. 2002). One unit of enzyme activity was defined as the amount of enzyme required to release 1 μmol *p*-nitrophenol per min from *p*-nitrophenyl ester.

### Characterisation of Est06

Esterase substrate specificity was assayed with triacylglycerides and *p*-nitrophenyl esters of fatty acids with different chain length. Plate assays for triacylglycerides were carried out by dropping 20 μl purified Est06 (7 U) on to LB agar plates supplemented with either 1% (w/v) tributyrin (C4), tricaproin (C6) or tricaprylin (C8), as well as 0.1% (w/v) tricaprin (C10), trilaurin (C12) or tripalmitin (C16). Zones of clearing on agar plates indicated lipolytic activity. Substrate specificity of Est06 for *p*-nitrophenyl esters was measured with 1 mM *p*NP acetate (C2), *p*NP butyrate (C4), *p*NP valerate (C5), *p*NP caproate (C6), *p*NP-caprylate (C8), *p*NP caprate (C10), *p*NP laurate (C12), *p*NP myristate (C14), or *p*NP palmitate (C16) under standard reaction conditions.

Optimum temperature was determined between 10 and 80 °C by using temperature-adjusted 50 mM sodium phosphate buffer (pH 7). Temperature stability of the enzyme was determined by incubating Est06 at 10, 25, 30, 40, 50 and 60 °C and measuring residual activity. As standard reaction conditions were defined for 25 °C, activity of Est06 at 25 °C was taken as 100%. The optimum pH of Est06 activity was assayed between pH 1 and 11, using the following overlapping buffer systems: KCl/HCl (pH 1–2.2), glycine/HCl (pH 2.2–3.6), sodium citrate (pH 3–6), sodium phosphate (pH 6–8), Tris/HCl (pH 7.6–9), CHES (pH 8.8–10), sodium carbonate (pH 9.6–10.6) and CAPS (pH

10–11). Due to pH-dependent absorption of *p*-nitrophenol, activity was monitored at the isosbestic point of *p*-nitrophenol (348 nm), the wavelength at which absorption is independent of pH (Hriscu et al. 2013). The pH stability of Est06 was assayed by incubating aliquots of enzyme in 20 µl buffer between pH 5 and 10 on a shaking incubator at 25 °C for 4 h. Aliquots of Est06 were pre-incubated between pH 5 and 10 in the respective buffers for 3 min. Residual enzyme activity was evaluated under standard reaction conditions and expressed as percentage of the activity of control reactions.

#### Effect of additives on Est06 activity

The effect of additives on Est06 activity was measured by pre-incubating the enzyme in assay buffer in 100 µl for 3 min at 25 °C on a shaking incubator. Est06 stability was determined by incubating the enzyme in assay buffer as above-mentioned for 1 h at 25 °C on a shaking incubator. Residual activity was calculated as a percentage of the activity of the control, which was not supplemented with the respective additive.

The effect of metal ions was studied by pre-incubating Est06 with metal ions at 1 and 5 mM in the assay buffer. The effect of salt on Est06 activity was measured directly by adding NaCl to the standard reaction assay from 0.05 to 3 M. Tolerance to salt was assayed by incubating the enzyme in 100 µl aliquots of assay buffers with NaCl from 0.05 to 3 M. The effect of detergents and inhibitors on Est06 activity was studied by pre-incubating Est06 with TritonX-100, Tween 20, Tween 80, SDS, 2-mercaptoethanol, dithiothreitol (DTT) or phenylmethylsulfonyl fluoride (PMSF) at 1 and 5 mM. The effect of organic solvents on Est06 activity was assayed by incubating the enzyme with 25 and 50% water-miscible organic solvents: acetone, 1-propanol, 2-propanol, ethanol, or dimethyl sulfoxide (DMSO) for 1 h.

## Results

### Identification and sequence analysis of a novel esterase-encoding gene

A lipolytic *E. coli* clone harbouring the recombinant plasmid pLE06 was recovered by functional screening of a forest soil metagenomic library (Nacke et al.

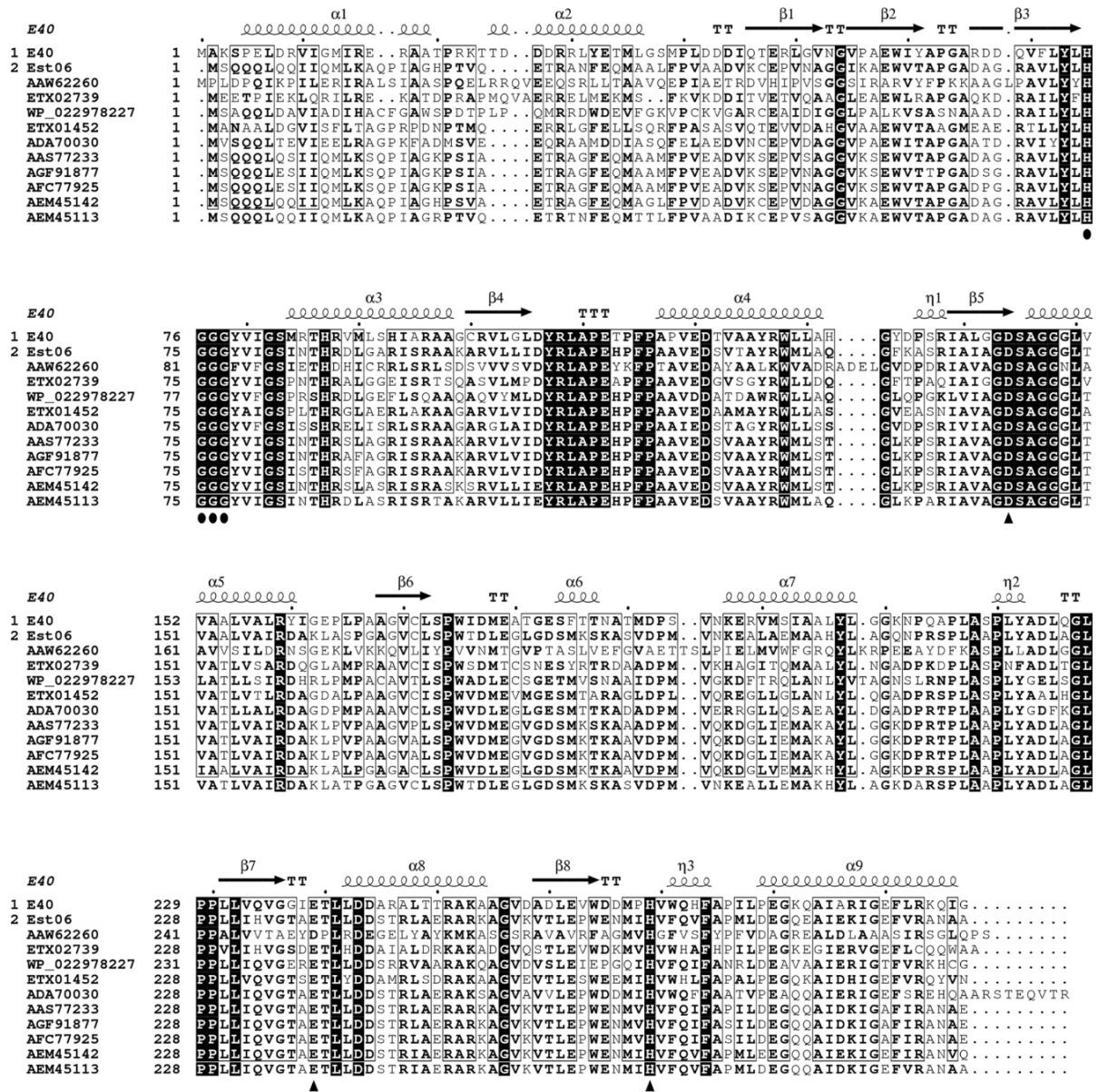
2011). Sequence analysis of the insert of the plasmid inserts revealed that it encodes a putative gene (*est06*) for a lipolytic enzyme, which is responsible for the lipolytic phenotype of the corresponding recombinant *E. coli* strain. The *est06* gene (888 bp) encodes a polypeptide chain of 295 residues. Sequence similarity searches showed that Est06 shares amino acid sequence similarity of more than 50% with several hypothetical or uncharacterised esterases and lipases from uncultured prokaryotes (Supplementary Table 1), including an esterase (ETX01452.1) from Candidatus *Entotheonella* sp. (61%). Multiple sequence alignments of Est06 with other esterases revealed that Est06 belongs to the HSL group of lipolytic proteins (Fig. 1). These include a HGGG motif (residues 74 to 77) and a catalytic triad composed of Ser<sup>144</sup>, Glu<sup>238</sup>, and His<sup>268</sup>. Tertiary structure prediction confirmed the presence of two  $\alpha$ -helices,  $\alpha$ 1 and  $\alpha$ 2, that constitute the *N*-terminal cap domain followed by the  $\alpha/\beta$  fold of the catalytic domain (Supplementary Fig. 2). The hydrolase fold comprises eight  $\beta$ -strands that form parallel structures, with Ser<sup>144</sup> located between  $\beta$ 5 and  $\alpha$ 5, Glu<sup>238</sup> between  $\beta$ 7 and  $\alpha$ 7 and His<sup>268</sup> between  $\beta$ 8 and  $\alpha$ 9. Two cysteine residues, Cys<sup>46</sup> and Cys<sup>170</sup>, which may participate in disulphide bridge formation are also present.

### Expression of the *est06* gene and purification of the corresponding protein

Est06 was expressed in *E. coli* BL21 (DE3) for 6 h at 30 °C. The enzyme was purified to homogeneity by a combination of affinity and anion exchange chromatography, yielding 248 µg pure enzyme from 2 l culture, with a specific activity of (5.76 U/mg). The protein has a calculated molecular mass of approx. 31 kDa and a theoretical pI of 5.99. Separation of Est06 enzyme preparation by SDS-PAGE showed a single band of approx. 35 kDa (Fig. 2), which is in accordance with calculated mass including the V5 epitope and His<sub>6</sub>-tag added during cloning of the gene in the expression vector.

### Substrate specificity of Est06

The ability of Est06 to hydrolyse different triacylglycerides and *p*-nitrophenyl esters of varying chain length was determined. A qualitative measure of triacylglyceride hydrolysis revealed the formation of

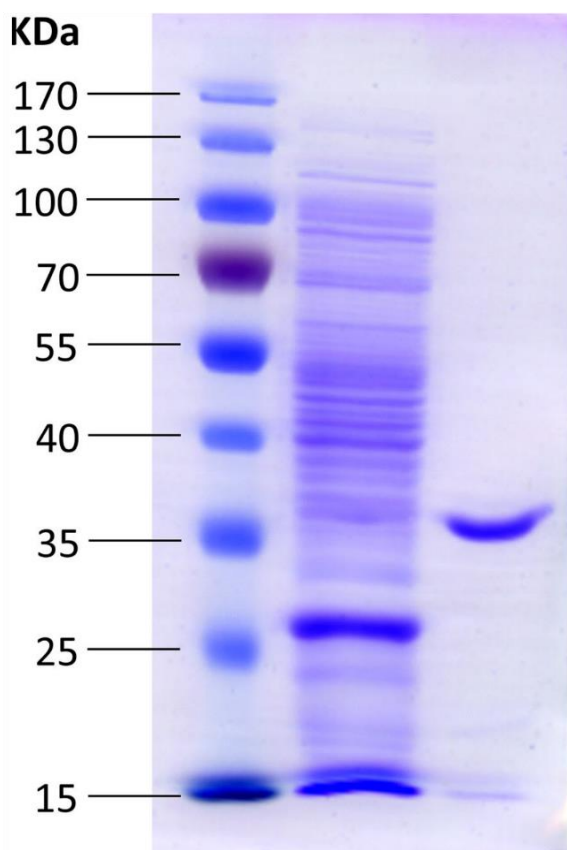


**Fig. 1** Multiple sequence alignment of Est06 and proteins related to the HSL family. Protein accession numbers are shown in parentheses. Catalytic triad residues are indicated by (filled triangle); residues of the oxanion hole are marked with (filled circle), turns are marked with (T) and 3<sub>10</sub> turns are labelled with (η). Secondary structures are annotated on top of sequences and

are modelled along E40, an HSL esterase that shares 55% sequence similarity with Est06. Regions with α-helices are shown by squiggles; regions with β-strands are marked by an arrow. Conserved residues are marked with a black background; similar residues are highlighted in bold, while sequences within the same group are marked with a border

halos on tributyrin-containing agar plates, which is indicative of lipolytic activity. However, none of the longer chain triacylglycerides were hydrolysed (data not shown), showing that Est06 could only hydrolyse short-chain triacylglycerides. Quantitative assays with

p-nitrophenyl esters showed that Est06 exhibited a marked substrate preference for esters with short-chain fatty acids (Fig. 3). The maximum of enzyme activity was recorded with pNP valerate (C5) as substrate. Negligible activities were detected with



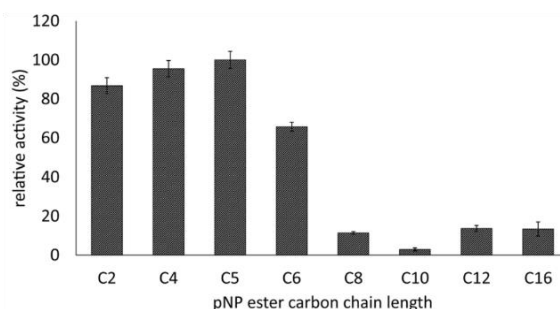
**Fig. 2** SDS-PAGE of purified, recombinant Est06. Lane 1 Marker PageRuler Prestained Protein Ladder (Fermentas), Lane 2 *E. coli* crude cellular lysate, Lane 3 purified Est06 after anion exchange chromatography

substrates of longer chain fatty acids (>C6). The inability to hydrolyse long-chain glycerol esters or *p*-nitrophenyl esters demonstrated that Est06 is an esterase and not a lipase.

#### Optimum temperature and thermal stability of Est06

The maximum activity of purified Est06 toward *p*NP-valerate (C5) was at 50 °C under standard reaction conditions. Activity increased continually from 10 to 50 °C. At higher temperatures, activity decreased rapidly and was not detectable at 80 °C (Fig. 4a).

While Est06 is able to function above 50 °C, stability above this temperature was very low as activity was not recorded after 1 h at 50 °C and 60 °C. However, Est06 had remarkably high stability for extended incubation time (13 days end of experiment)



**Fig. 3** Substrate specificity of Est06. Activity was determined for *p*NP esters of varying chain length at 25 °C and pH 7. Relative activity was expressed as a percentage of *p*NP valerate (C5) enzyme activity (2.1 U/mg). Values are given as the mean of three experiments  $\pm$  standard deviations

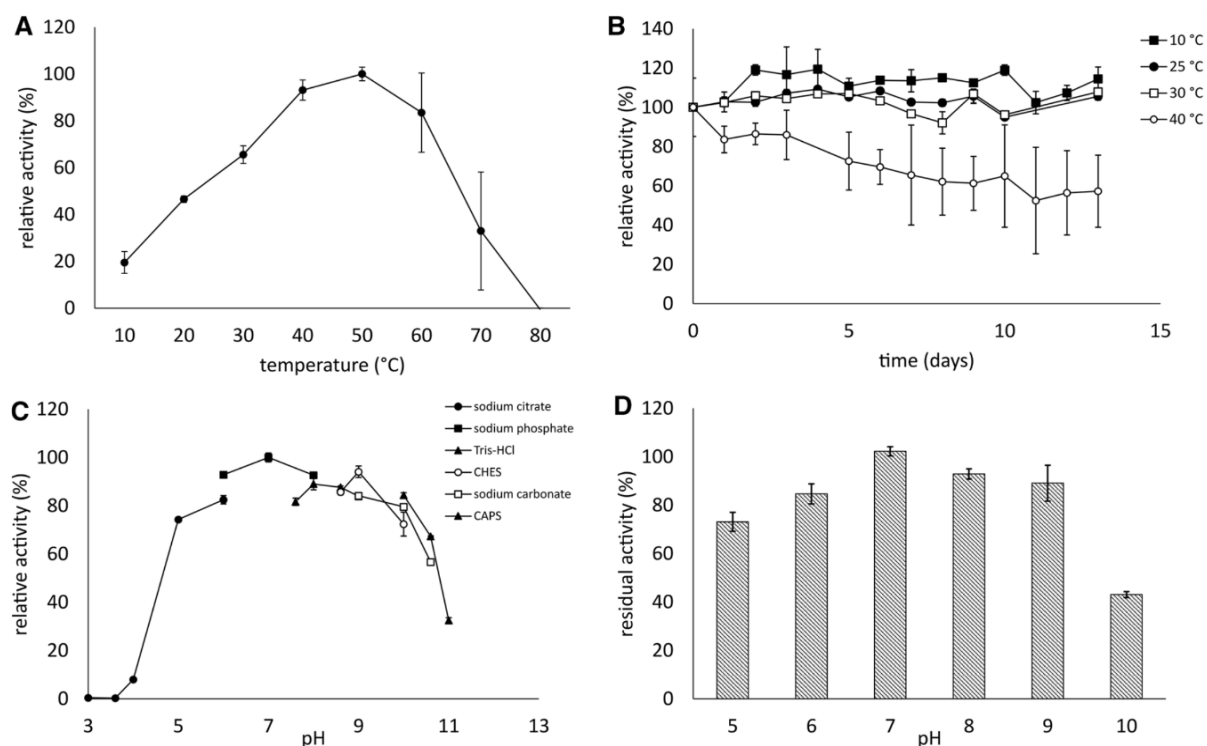
at below 50 °C (Fig. 4b). Est06 retained all its activity at 10, 25 and 30 °C over the entire period. After six days of incubation at 40 °C, esterase activity remained above 70%. Considerable loss in Est06 activity (less than 65% residual activity) was observed only after the 10<sup>th</sup> day of incubation.

#### Optimum pH and pH stability of Est06

The optimum pH of Est06 was evaluated at 25 °C (see Fig. 4c). Activity was detected from pH 3.6–11. More than 70% activity was retained between pH 5 and pH 10. Maximum activity was at pH 7 and remained as high as 80% up to pH 10. Est06 retained more than 70% of its activity between pH 5 and 9 after 4 h incubation at 25 °C (Fig. 4d). Thus, Est06 is not only active over a wide range of pH values but is also stable over a broad pH spectrum and active under acidic as well as alkaline conditions.

#### Effect of additives on Est06 activity

Incubation of Est06 with different metal ions resulted in no significant loss of activity with the exception of  $\text{Fe}^{2+}$  (65% residual activity) and  $\text{Co}^{2+}$  (55% residual activity) (Fig. 5a). None of the metal ions significantly stimulated activity. The inability of EDTA to inhibit esterase activity is an indication that Est06 does not require metal ion activators or prosthetic groups for activity or stability (Mohamed et al. 2013; Peng et al. 2011). Addition of NaCl produced no stimulatory effect on Est06, and activity decreased linearly with increasing NaCl concentration (Fig. 5b). After incubation for 1 h,



**Fig. 4** Biochemical characterisation of lipolytic enzyme Est06. **a** Effect of temperature on Est06 activity, **b** temperature stability profile of Est06, **c** effect of pH on Est06 enzyme activity, **d** stability of Est06 after incubation at 4 h at pH 5–10. Specific

activities corresponding to 100% relative activity are: 3.7, 4.9, 1.1 and 1.4 U/mg for graphs A, B, C and D, respectively. Values are given as the mean of three experiments  $\pm$  standard deviations

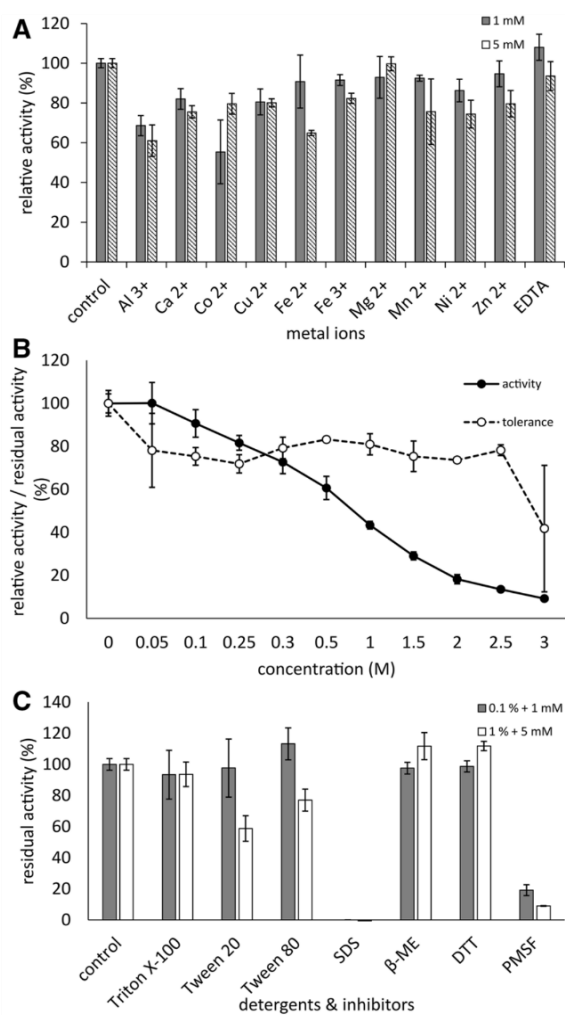
Est06 exhibited moderate stability with NaCl up to 2.5 M with the residual activity being over 75%. The results demonstrated that Est06 is not dependent on salt for activation or stability.

At low detergent concentrations (0.1%), Est06 retained most of its activity (above 90%), with the exception of SDS and PMSF, which resulted in loss of activity (Fig. 5c). At higher concentrations (1%), 2-mercaptoethanol and DTT exhibited a stimulatory effect on esterase activity (110%), while SDS and PMSF continued to inhibit Est06. Enzyme inhibition caused by SDS or PMSF is indicative of serine hydrolases (Peng et al. 2011). Tolerance of Est06 towards water-miscible organic solvents was generally low (Supplementary Table 2). An increase in enzyme activity (130%) was observed at 25% (v/v) DMSO although at a higher concentration (50%), DMSO decreased enzyme activity to 72%. All other tested organic solvents exhibited strong inhibition of Est06 activity, particularly at high solvent concentrations.

## Discussion

Metagenomics is a powerful tool for the discovery of novel biocatalysts from various environments. In this study, a novel esterase from a forest soil metagenome was cloned, expressed and characterised. Est06 was identified through its ability to hydrolyse tributyrin on solid media and was initially classified as a putative esterase (Nacke et al. 2011). Further sequence analysis revealed that Est06 belongs to the HSL (Family IV) group of lipolytic proteins. The esterase shares similarity with a hypothetical protein from *Candidatus Entotheonella* sp. TSY1, which is a bacterial symbiont from deltaproteobacteria (Wilson et al. 2014). Esterases belonging to Family IV share amino acid sequence similarity to mammalian HSL, which regulates lipid metabolism in adipose tissue (Østerlund 2001). This indicates that bacterial and mammalian esterases share a common evolutionary root (Rhee et al. 2005). Signature motifs in this group include an HGGG motif, which is present in the Est06





**Fig. 5** The effect of additives on Est06 activity and stability. Activity of Est06 under the influence of **a** metal ions, **b** salt and **d** detergents and inhibitors. Specific activity values expressed as percentages of the control reactions are: 4, 3.8, 5.2 and 3.2 U/mg for graphs A, B and C, respectively. Values are given as the mean of three experiments  $\pm$  standard deviations

protein sequence and participates in the formation of an oxyanion hole (Li et al. 2014). The oxyanion hole forms a pocket that aids in stabilising the negative charge of the tetrahedral intermediate during substrate binding (Lee et al. 2006). The catalytic triad in esterases/lipases is composed of Ser/His/Asp or, less commonly, Glu. Li et al. (2014) proposed two GX SXG motifs (GDSAG and GTSAG) within the HSL group and their separation into two distinct subfamilies. The active site serine in Est06 (Ser<sup>144</sup>) is located in the conserved GDSAG motif. The triad is

completed by glutamine (Glu<sup>238</sup>) in the EXLLD motif and histidine (His<sup>268</sup>). Prediction of esterase tertiary structures has been achieved by using proteins sharing low amino acid sequence similarity. Berlemont et al. (2013) predicted the structure of organic-solvent tolerant esterase (RBest1) by using a protein that shared 36% amino acid sequence similarity, while Novototskaya-Vlasova et al. (2012) used as little as 16% similarity for their modelling of the cold-active lipase EstPc. In this study, tertiary structure prediction of Est06 was performed by using esterase E40 (Li et al. 2015), which shares 55% amino acid sequence similarity. Results confirm the  $\alpha/\beta$  hydrolase fold reported for all members of lipase/esterase gene families (van Pouderoeyen et al. 2001).

Esterases of the HSL family exhibit diverse optima for temperature and pH. Est06 was derived from a temperate forest soil metagenome, and as such it exhibits maximum catalytic activity under mesophilic conditions (below 50 °C). In addition, Est06 displays unprecedented thermal tolerance at moderate temperatures. Kovacic et al. (2016) demonstrated the ability for esterases to evolve thermal adaptations that resemble their bacterial host habitats. Thermal tolerance was attributed to the presence of loops in the tertiary fold that increases or decreases flexibility with changes in temperature. In comparison to the four esterases described by Kovacic et al. (2016), Est06 did not lose activity below 30 °C, a feature that highlights Est06 as a candidate catalyst in low temperature bioprocesses. Byun et al. (2007) reported for HSL member EstE1 that thermostability is imparted by hydrophobic interaction of valine and phenylalanine residues on  $\beta$ 8, however these residues are absent in Est06 (Fig. 1). Additionally, factors such as hydrogen bonding and oligomerisation also play a role in esterase thermostability (De Simone et al. 2001; Rhee et al. 2006), which have not been evaluated in Est06. An adaptation to low working temperatures is an attractive trait for reactions in which high temperatures are not suitable, such as manufacture of thermolabile pharmaceutical products, food ingredients as well as production of cold-wash detergents (López-López et al. 2014). In addition, Est06 displays a broad pH spectrum for catalysis. Esterase activity was detected from as low as pH 3.6. The highest activities were observed from pH 5 to pH 10. Moreover, Est06 stability under different pH conditions was also high between pH 5 and pH 9. In comparison, a *Geobacillus*

sp. TF17 esterase with comparable pH stability was described by Ayna et al. (2013), but this enzyme exhibited maximal activity in a much narrower pH range (pH 7–8). The *Sulfolobus tokodaii* esterase reported by Suzuki et al. (2004) was highly active between pH 5 and 9. In contrast, an esterase from *Bacillus circulans* was maximally active only between pH 8 and 9, and was stable between pH 6 to 8 (Kademi et al. 2000). Acid-stable enzymes such as Est06 are desirable catalysts in bioremediation processes for contaminated waste water or biofuel production sites, and application in lignocellulosic waste degradation (Sharma et al. 2012).

Est06 does not require metal ion activators or prosthetic groups for activity or stability (Peng et al. 2011), as addition of EDTA did not result in loss of enzyme activity. The activity of two HSL esterases (SBLip 2 and SBLip 5.1) identified by Biver and Vandenbol (2013) was similarly unaffected by the addition of EDTA. Although mechanisms for metal ion resistance in esterases are not well understood, Est06 was resistant to  $\text{Cu}^{2+}$  and  $\text{Zn}^{2+}$ , metal ions usually with a strong inhibitory effect. Est06 retained more than 80% activity with both metal ions even at 5 mM. In contrast, an esterase (EstATII) characterised by Mohamed et al. (2013) was severely inhibited by  $\text{Cu}^{2+}$  and  $\text{Zn}^{2+}$ . Such an ability to withstand additions of metal ions is an important factor for enzymes used in the bioremediation of environmental waste (Brault et al. 2012).

Serine hydrolases such as Est06 are commonly prone to inactivation by inhibitors such as PMSF and SDS. PMSF binds irreversibly to the active site serine residue and makes it inaccessible to substrate molecules (Sharma and Radha Kishan 2011). SDS inhibits enzyme activity by forming complexes with proteins through non-specific binding, which results in unfolding of the protein. Est06 was generally more stable in non-ionic detergents compared to SDS and PMSF. At 1% detergent, Est06 showed a decrease in residual activity with Tween 20 and Tween 80. Similarly, EstATII showed comparable behaviour towards Tween 20 and Tween 80 but with a severe decrease in activity (Mohamed et al. 2013).

Redox reagents, 2-mercaptoethanol and DTT exhibited minor effects on Est06 activity. Resistance to these reagents is related to the number of cysteine residues present in the protein, which are capable of forming disulphide bridges. Levisson et al. (2007)

proposed that a single cysteine residue is unable to form disulphide bridges with nearby residues and as such, redox reagents such as 2-mercaptoethanol and DTT have no marked effect on activity. Therefore, the inability of 2-mercaptoethanol and DTT to inhibit or significantly activate Est06 confirmed that Cys<sup>46</sup> and Cys<sup>170</sup> form no disulphide bridges in Est06 tertiary structure. In contrast, Cys residues in esterase AprX-SK37 form 8 disulphide bridges, which subsequently contributed to strong inhibition by 2-mercaptoethanol and DTT (Phrommao et al. 2011). Neither activity nor stability of Est06 was dependent on NaCl. Halo-tolerance may be positively correlated with organic solvent tolerance (Berlemont et al. 2013). As salt causes a decrease in water molecules around the enzyme, these molecules would be replaced by organic solvents (Sellek and Chaudhuri 1999; Berlemont et al. 2013). A *Salimicrobium* sp. LY19 esterase supports these findings as it exhibited halo and organic solvent tolerant properties (Xin and Hui-Ying 2013). In comparison, low tolerance to salt by Est06 was accompanied by low organic solvent tolerance. Generally, enzymes have reduced activity in organic solvents than in water (Mattos and Ringe 2001). Est06 was inhibited by water-miscible organic solvents. At 25% total organic solvent concentration, there was a rough correlation between  $\log P_{ow}$ , which is a measure of solvent polarity, and residual activity. Thus, residual activity decreased with increasing  $\log P_{ow}$  (Supplementary Table 2). Esterase EstC2 showed similar inactivation by alcohols at 50% concentration (Kang et al. 2011). A 30% increase in activity was observed with 25% DMSO, although a higher concentration of DMSO (50%) had an inhibiting effect on Est06. Metin et al. (2006) attribute the activating effect of DMSO to the high rate of diffusion of substrate in its presence. Water-miscible solvents form a uniform phase with water and enable substrates quick access to the active site (Ogino and Ishikawa 2001). Mattos and Ringe (2001) counter argue that increased solvent polarity, as with DMSO, actually strips away the water layer around the enzyme and competes for hydrogen bonds in protein atoms, thus leading to inactivation. This was not the case with Est06 at low DMSO concentration but as concentration increased to 50%, a decrease in activity that could be caused by this inhibition phenomenon was recorded. In this study DMSO follows the amphiphilic effect described by Sellek and Chaudhuri (1999) in

which DMSO functions as an activator at low solvent concentrations and an inhibitor at higher concentrations.

In conclusion, a temperate forest soil metagenome yielded an esterase with significant stability at moderate temperatures and broad pH range. Est06 shows no reliance on additives for activity or stability and exhibited marked resistance to metal ions, detergents and low concentrations of organic solvents. The combination of these properties demonstrate the robust capabilities of Est06 as a novel biocatalysts in low temperature applications.

**Acknowledgements** We thank the DFG (German Science Foundation) for funding this work in the context of the graduate school 'Die Bedeutung der Biodiversität für Stoffkreisläufe und biotische Interaktionen in temperaten Laubwäldern (GRK 1086)'. We thank Dr. Heiko Nacke for providing the lipolytic clone pLE06 used in this work, as well as Florian Jung and Jörn Lindemann for providing technical assistance.

**Supporting information** Supplementary Table 1—Amino acid sequence similarities between Est06 and related lipolytic proteins.

Supplementary Table 2—The effect of water-miscible organic solvents on the stability of Est06.

Supplementary Figure 1—Maps of plasmids for cloning and expression of the lipolytic gene *est06*.

Supplementary Figure 2—Predicted tertiary structure of esterase Est06.

#### Compliance with ethical standards

**Conflict of interest** The authors declare no conflict of interest exists.

#### References

- Altschul SF, Madden TL, Schäffer AA et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Res* 25:3389–3402
- Arpigny JL, Jaeger K-E (1999) Bacterial lipolytic enzymes: classification and properties. *Biochem J* 343:177–183
- Ayna Ç, Kolcuoğlu Y, Öz F et al (2013) Purification and characterization of a pH and heat stable esterase from *Geobacillus* sp. TF17. *Turk J Biochem* 38:329–336
- Berlemont R, Spee O, Delsaute M et al (2013) Novel organic solvent-tolerant esterase isolated by metagenomics: insights into the lipase/esterase classification. *Rev Argent Microbiol* 45:3–12
- Biver S, Vandenbol M (2013) Characterization of three new carboxylic ester hydrolases isolated by functional screening of a forest soil metagenomic library. *J Ind Microbiol Biotechnol* 40:191–200
- Bornscheuer UT (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis. *FEMS Microbiol Rev* 26:73–81
- Brault G, Shareck F, Hurtubise Y et al (2012) Isolation and characterization of EstC, a new cold-active esterase from *Streptomyces coelicolor* A3(2). *PLoS ONE* 7:e32041
- Byun J-S, Rhee J-K, Kim ND et al (2007) Crystal structure of hyperthermophilic esterase EstE1 and the relationship between its dimerization and thermostability properties. *BMC Struct Biol* 7:47
- Charbonneau DM, Beauregard M (2013) Role of key salt bridges in thermostability of *G. thermodenitrificans* EstGtA2: distinctive patterns within the new bacterial lipolytic enzyme Family XV. *PLoS ONE* 8:1–18
- de Castro AP, Fernandes GdaR, Franco OL (2014) Insights into novel antimicrobial compounds and antibiotic resistance genes from soil metagenomes. *Front Microbiol* 5:1–9
- De Simone G, Menchise V, Manco G et al (2001) The crystal structure of a hyper-thermophilic carboxylesterase from the archaeon *Archaeoglobus fulgidus*. *J Mol Biol* 314:507–518
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Facchin S, Diniz Alves PD, De Faria Siqueira F et al (2013) Biodiversity and secretion of enzymes with potential utility in wastewater treatment. *Open J Ecol* 3:34–47
- Gasteiger E, Hoogland C, Gattiker A et al (2005) Protein identification and analysis tools on the ExpASY server. In: Walker JM (ed) *The proteomics protocols handbook*. Humana Press Inc., Totowa, pp 571–607
- Hotta Y, Ezaki S, Atomi H, Imanaka T (2002) Extremely stable and versatile carboxylesterase from a hyperthermophilic archaeon. *Appl Environ Microbiol* 68:3925–3931
- Hriscu M, Chiş L, Toşa M, Irimie FD (2013) pH-profiling of thermoactive lipases and esterases: caveats and further notes. *Eur J Lipid Sci Technol* 115:571–575
- Kademi A, Ait-Abdelkader N, Fakhreddine L, Baratti JC (2000) Characterization of a new thermostable esterase from the moderate thermophilic bacterium *Bacillus circulans*. *J Mol Catal B Enzym* 10:395–401
- Kang C-H, Oh K-H, Lee M-H et al (2011) A novel family VII esterase with industrial potential from compost metagenomic library. *Microb Cell Fact* 10:41
- Kovacic F, Mandrysch A, Poojari C et al (2016) Structural features determining thermal adaptation of esterases. *Protein Eng Des Sel* 29:65–76
- Lee M-H, Lee C-H, Oh T-K et al (2006) Isolation and characterization of a novel lipase from a metagenomic library of tidal flat sediments: evidence for a new family of bacterial lipases. *Appl Environ Microbiol* 72:7406–7409
- Lenfant N, Hotelier T, Velluet E et al (2013) ESTHER, the database of the  $\alpha/\beta$ -hydrolase fold superfamily of proteins: tools to explore diversity of functions. *Nucleic Acid Res* 41:D423–D429
- Levisson M, van der Oost J, Kengen SWM (2007) Characterization and structural modeling of a new type of thermostable esterase from *Thermotoga maritima*. *FEBS J* 274:2832–2842
- Li PY, Ji P, Li CY et al (2014) Structural basis for dimerization and catalysis of a novel esterase from the GTSAG motif

- subfamily of the bacterial hormone-sensitive lipase family. *J Biol Chem* 289:19031–19041
- Li P-Y, Chen X-L, Ji P et al (2015) Interdomain hydrophobic interactions modulate the thermostability of microbial esterases from the hormone-sensitive lipase family. *J Biol Chem* 290:1118–11198
- López-López O, Cerdán ME, González Siso MI (2014) New extremophilic lipases and esterases from metagenomics. *Curr Protein Pept Sci* 15:445–455
- Ma B-D, Kong X-D, Yu H-L et al (2014) Increased catalyst productivity in  $\alpha$ -hydroxy acids resolution by esterase mutation and substrate modification. *ACS Catal* 4:1026–1031
- Martínez-Martínez M, Lores I, Peña-García C et al (2014) Biochemical studies on a versatile esterase that is most catalytically active with polyaromatic esters. *Microb Biotechnol* 7:184–191
- Mattos C, Ringe D (2001) Proteins in organic solvents. *Curr Opin Struct Biol* 11:761–764
- Metin K, Ateslier ZBB, Basbulbul G, Biyik HH (2006) Characterization of esterase activity in *Geobacillus* sp. HBB-4. *J Basic Microbiol* 46:400–409
- Mohamed YM, Ghazy MA, Sayed A et al (2013) Isolation and characterization of a heavy metal-resistant, thermophilic esterase from a Red Sea brine pool. *Sci Rep* 3:3358
- Nacke H, Will C, Herzog S et al (2011) Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories. *FEMS Microbiol Ecol* 78:188–201
- Nardini M, Dijkstra BW (1999)  $\alpha/\beta$  hydrolase fold enzymes: the family keeps growing. *Curr Opin Struct Biol* 9:732–737
- Novototskaya-Vlasova K, Petrovskaya L, Yakimov S, Gilichinsky D (2012) Cloning, purification, and characterization of a cold-adapted esterase produced by *Psychrobacter cryohalolentis* K5T from Siberian cryopeg. *FEMS Microbiol Ecol* 82:367–375
- Ogino H, Ishikawa H (2001) Enzymes which are stable in the presence of organic solvents. *J Biosci Bioeng* 91:109–116
- Østerlund T (2001) Structure-function relationships of hormone-sensitive lipase. *Eur J Biochem* 268:1899–1907
- Peng Q, Zhang X, Shang M et al (2011) A novel esterase gene cloned from a metagenomic library from neritic sediments of the South China Sea. *Microb Cell Fact* 10:95
- Phrommao E, Yongsawatdigul J, Rodtong S, Yamabhai M (2011) A novel subtilase with NaCl-activated and oxidant-stable activity from *Virgibacillus* sp. SK37. *BMC Biotechnol* 11:65
- Rhee J-K, Ahn D-G, Kim Y-G, Oh J-W (2005) New thermophilic and thermostable esterase with sequence similarity to the hormone-sensitive lipase family, cloned from a metagenomic library. *Appl Environ Microbiol* 71:817–825
- Rhee J-K, Kim D-Y, Ahn D-G et al (2006) Analysis of the thermostability determinants of hyperthermophilic esterase EstE1 based on its predicted three-dimensional structure. *Appl Environ Microbiol* 72:3021–3025
- Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42:W320–W324
- Rodríguez MC, Loaces I, Amarelle V et al (2015) Est10: a novel alkaline esterase isolated from bovine rumen belonging to the new family XV of lipolytic enzymes. *PLoS ONE* 10:e0126651
- Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5:725–738
- Sellek GA, Chaudhuri JB (1999) Biocatalysis in organic media using enzymes from extremophiles. *Enzyme Microb Technol* 25:471–482
- Selvin J, Kennedy J, Lejon DPH et al (2012) Isolation identification and biochemical characterization of a novel halotolerant lipase from the metagenome of the marine sponge *Haliclona simulans*. *Microb Cell Fact* 11:72
- Sharma A, Radha Kishan KV (2011) Serine protease inhibitor mediated peptide bond re-synthesis in diverse protein molecules. *FEBS Lett* 585:3465–3470
- Sharma A, Kawarabayasi Y, Satyanarayana T (2012) Acidophilic bacteria and archaea: acid stable biocatalysts and their potential applications. *Extremophiles* 16:1–19
- Suzuki Y, Miyamoto K, Ohta H (2004) A novel thermostable esterase from the thermoacidophilic archaeon *Sulfolobus tokodaii* strain 7. *FEMS Microbiol Lett* 236:97–102
- van Pouderoyen G, Eggert T, Jaeger K-E, Dijkstra BW (2001) The crystal structure of *Bacillus subtilis* lipase: a minimal  $\alpha/\beta$  hydrolase fold enzyme. *J Mol Biol* 309:215–226
- Will C, Thürmer A, Wollherr A et al (2010) Horizon-specific bacterial community composition of german grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. *Appl Environ Microbiol* 76:6751–6759
- Wilson MC, Mori T, Rückert C et al (2014) An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* 506:58–62
- Xin L, Hui-Ying Y (2013) Purification and characterization of an extracellular esterase from a moderately halophilic bacterium. *BMC Biotechnol* 13:108
- Yang J, Zhang Y (2015) I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res* 43:W174–W181
- Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinform* 9:40
- Zhang X-Y, Fan X, Qiu Y-J et al (2014) Newly identified thermostable esterase from *Sulfobacillus acidophilus*: properties and performance in phthalate ester degradation. *Appl Environ Microbiol* 80:6870–6878

The following section contains supplementary information from Dukunde et al. (2017) *Biotechnol Lett* 39: 577 – 587.

## **Supplementary information**

### **A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum**

Amélie Dukunde, Dominik Schneider, Mingji Lu, Silja Brady <sup>1</sup>, Rolf Daniel \*

#### **Content**

**Supplementary Table 1** Amino acid sequence similarities between Est06 and related lipolytic proteins

**Supplementary Table 2** The effect of water-miscible organic solvents on the stability of Est06

**Supplementary Figure 1** Maps of plasmids for cloning and expression of the lipolytic gene *est06*.

**Supplementary Figure 2** Predicted tertiary structure of esterase Est06

**Supplementary Table 1.** Amino acid sequence similarities between Est06 and related lipolytic proteins

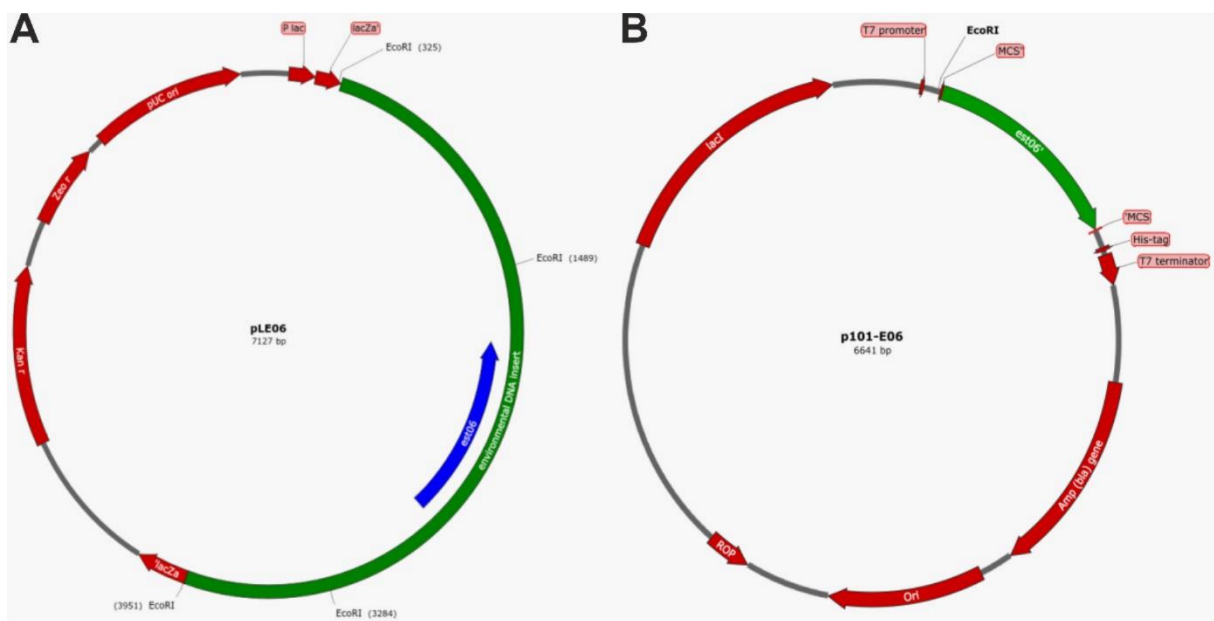
| Accession      | Description  | Identity (%) | Peptide length | Query cover (%) |
|----------------|--|--------------|----------------|-----------------|
| AEM45114.1     | Query protein Est06  | 100          | 296            | 100             |
| AEM45113.1     | Hypothetical protein from uncultured organism                                    | 90           | 296            | 95              |
| AAS77233.1     | Lipase or esterase from uncultured bacterium                                     | 82           | 296            | 90              |
| AGF91877.1     | Esterase or lipase protein from uncultured bacterium                             | 81           | 296            | 89              |
| AAX37296.1     | Esterase or lipase protein from uncultured bacterium                             | 80           | 297            | 89              |
| AFB82694.1     | Lipolytic enzyme from uncultured bacterium                                       | 70           | 214            | 81              |
| AFB82687.1     | Lipolytic enzyme from uncultured bacterium                                       | 65           | 306            | 78              |
| ETX01452.1     | Hypothetical protein (ETSY1_07290) from Candidatus <i>Entotheonella</i> sp. TSY1 | 61           | 296            | 75              |
| ALH24854.1     | Halotolerant esterase from uncultured bacterium                                  | 61           | 287            | 86              |
| KP696774       | Bacterial esterase E40   | 55           | 317            | 70              |
| GAN77874.1     | Lipase or esterase or arylesterase from <i>Acidisphaera rubrifaciens</i> HS-AP3  | 53           | 265            | 68              |
| WP_012455056.1 | $\alpha/\beta$ hydrolase from <i>Methylobacterium populi</i>                     | 50           | 296            | 96              |

**Supplementary Table 2.** The effect of water-miscible organic solvents on the stability of Est06

| Organic solvent    | Log P <sub>ow</sub> <sup>a</sup> | Residual activity in organic solvents |                  |
|--------------------|----------------------------------|---------------------------------------|------------------|
|                    |                                  | 25% <sup>b</sup>                      | 50% <sup>b</sup> |
| <b>Control</b>     | —                                | 100.0 ± 13.2                          | 100.0 ± 13.2     |
| <b>DMSO</b>        | -1.35                            | 129.7 ± 2.7                           | 72.3 ± 16.4      |
| <b>Methanol</b>    | -0.77                            | 71.2 ± 1.2                            | 0.0 ± 0.0        |
| <b>Ethanol</b>     | -0.31                            | 14.1 ± 7.7                            | -0.1 ± 0.1       |
| <b>Acetone</b>     | -0.24                            | 7.8 ± 2.4                             | 0.0 ± 0.1        |
| <b>Propan-2-ol</b> | 0.05                             | 36.6 ± 12.6                           | 0.2 ± 0.1        |
| <b>Propan-1-ol</b> | 0.25                             | -0.4 ± 0.6                            | 0.1 ± 0.1        |

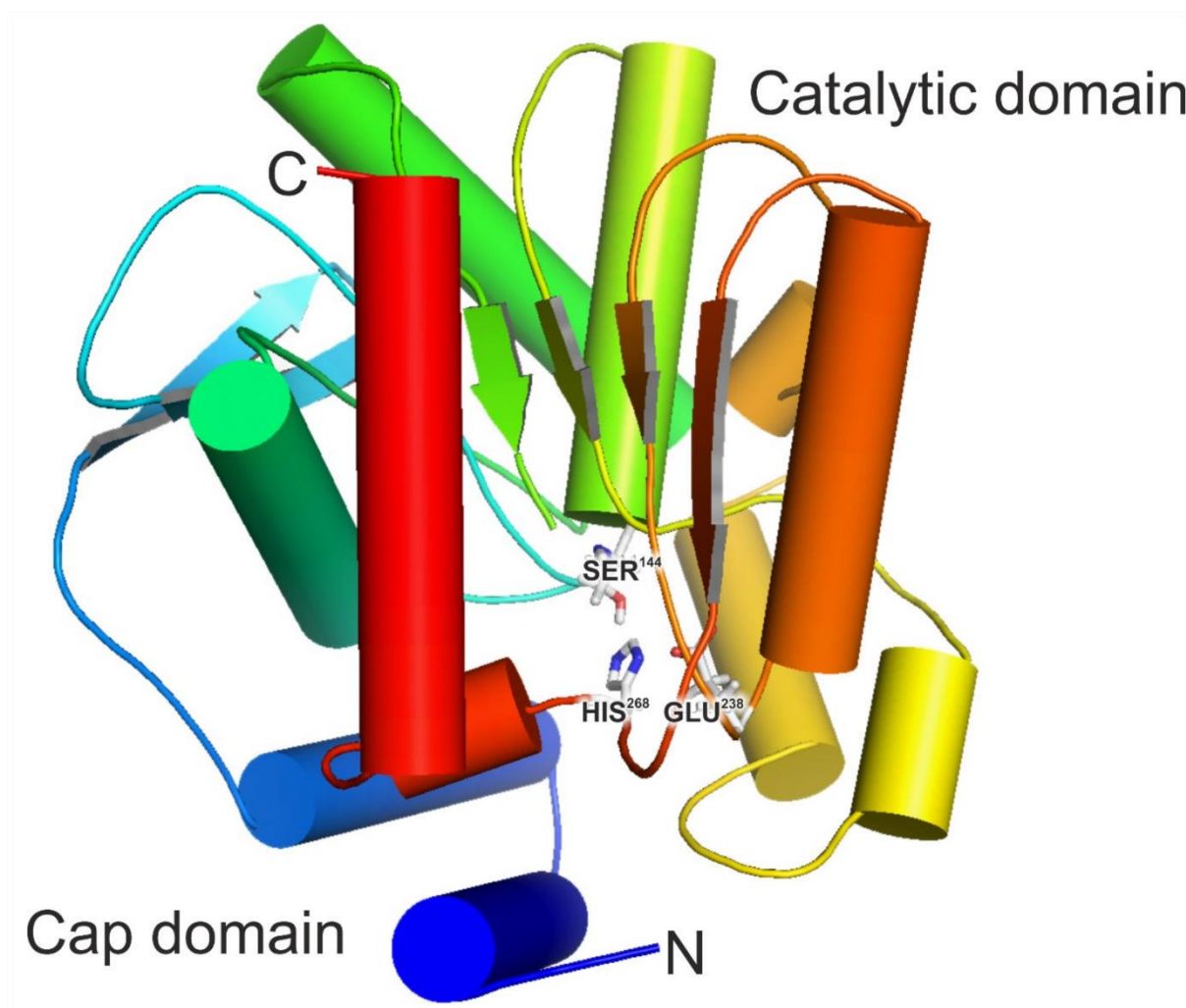
<sup>a</sup> The logarithmic partition coefficient, P, of the solvent between octanol and water describes quantitative measure of solvent polarity. Values are based those calculated by Leo et al. 1971 and compiled by Sangster, 1989.

<sup>b</sup> Final concentration of organic solvent in assay buffer under standard conditions



**Supplementary Figure 1.** Maps of plasmids vectors for cloning and expression of the lipolytic gene *est06*. (A) The plasmid pLE06 contains the original environmental DNA insert. (B) The expression vector p101-E06 used for the expression of esterase Est06. The plasmid backbone is coloured in red, while environmental DNA and *est06* gene are coloured green and blue, respectively





**Supplementary Figure 2.** Predicted tertiary structure of esterase Est06. Residues of the catalytic triad are highlighted within the active site. Cylindrical tubes indicate  $\alpha$ -helical regions and arrows indicate  $\beta$ -stranded regions. The letters N and C indicate the respective protein termini

### References

Leo A, Hansch C, Elkins D (1971) Partition coefficients and their uses. *Chem Rev* 6:525–616

Sangster, J (1989) Octanol-Water Partition Coefficients of Simple Organic Compounds. *J Phys Chem Ref Data* 18:1111-1227

# Chapter 5

---

## Summary

Forest ecosystems provide an important contribution to global carbon storage through processes mediated by trees. As the dominant forest vegetation, trees largely drive primary productivity and mediate the flow of aboveground and belowground carbon stocks. Subsequently, microbial communities rely on trees to carry out important ecosystem processes, such as nutrient cycling, through the decomposition of organic matter and metabolism of root exudates (Figure 1.3, page 12). Therefore, understanding how trees shape microbial community structure can improve the ability to predict ecosystem responses to environmental disturbance. Metagenomics remains a powerful tool for describing the ecology of microbial communities as it is possible to access taxonomic and functional information at the level of individual taxa. The following thesis examined the taxonomic structure and functional potential of soil bacterial communities in a temperate deciduous forest by employing a metagenomic approach.

In chapter 2, soil samples were collected from the A horizon of mono and mixed stands of beech, hornbeam, lime, and oak in spring, summer and autumn. Subsequently, amplicon-based analysis of 16S rRNA genes and transcripts revealed that the total (DNA-based) and potentially active (RNA-based) soil bacterial communities significantly responded to tree species identity (mono stands) and to a lesser extent, to tree species richness (mixed stands) (Figure 2.2, page 26). Members of *Rhizobiales* and *Rhodospirillales* (*Proteobacteria*), *Gaiellales*, *Frankiales*, and *Solirubrobacterales* (*Actinobacteria*) and *Bacteroidetes* were more abundant in nutrient-enriched lime and hornbeam mono stands. In contrast, *Acidobacteriales* and *Solibacterales* (*Acidobacteria*), and *Xanthomonadales* (*Proteobacteria*) exhibited a strong association for nutrient-reduced soils under beech and oak mono stands (Figure 2.4 and Figure S2.3, pages 31 and 41, respectively). Moreover, soil C/N ratio, pH and P content exhibited significant impact on soil bacterial communities and were attributed to direct and indirect effects of forest stands. Trees possess several species-dependent traits, including leaf litter quality and fine root biomass, which bring about changes in soil chemistry. Prediction of metabolic functions with Tax4Fun revealed that metabolic functions related to C fixation and degradation, and N metabolism responded significantly to seasonality, rather than tree species (Figure 2.6, page 33). Both processes were significantly abundant in spring, while C degradation gene abundances increased from summer to autumn, corresponding to periods of increased litterfall and decomposition.

Forests also generate several ecosystem services that are important to society and in chapter 3, activity-based functional screening of metagenomic libraries was conducted. Short insert plasmid libraries were constructed successfully using environmental DNA derived from forest soils screened for lignocellulolytic activity (Table 3.1, page 59). Two clones, Lip3 and Lip49, exhibited lipolytic activity on nutrient agar supplemented with tributyrin. Sequence analyses showed that Lip3 and Lip49 genes share 54 % and 63 % similarity, respectively, with closely related esterase genes (Table 3.3, page 63). The results indicate that Lip 3 and Lip49 encode potentially novel lipolytic proteins. Conserved

sequence blocks both genes include residues forming the catalytic triad (Ser, His and Asp) and possible oxyanion holes (Figure 3.5, page 64). Lip3 has a conserved GHSQG motif, which is commonly found in true lipases of Family I. However, Lip 49 has GFSQG motif and reveals it to belong to Family VI carboxylesterases.

In chapter 4, purification and characterisation of a metagenome-derived esterase, Est06, was conducted. Est06 is a novel 31 kDa carboxylesterase from Family IV, or hormone sensitive lipase (HSL) family (Fig. 1, page 70). As all bacterial HSL esterases, Est06 showed high affinity for acyl esters with short-chain fatty acids. Est06 exhibited optimum enzymatic activity at 50 °C and pH 7 with *p*-nitrophenyl valerate (C5) substrate. Interestingly, Est06 retained most of its activity below 30 °C over 13 days and showed high catalytic stability between pH 5 and pH 9 (Fig. 4, page 72). This is considerably higher stability than reported for other Family IV carboxyl esterases. Additionally, Est06 was not inhibited by metal ions (Fig. 5, page 73). These properties make Est06 an desirable candidate in low temperature industrial applications, such as detergent manufacture and bioremediation. The results of this work highlight soil bacterial community responses to forest stands and provide potential tools to recover bacterial-derived biocatalysts with industrial applications.

# References

References used in Dukunde et al. (2017) can be found from page 75 – 76.

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic Local Alignment Tool. *J. Mol. Biol.* 215, 403–410.
- Alves-Prado, H. F., Pavezzi, F. C., Leite, R. S. R., De Oliveira, V. M., Sette, L. D., and DaSilva, R. (2010). Screening and production study of microbial xylanase producers from Brazilian Cerrado. *Appl. Biochem. Biotechnol.* 161, 333–346. doi:10.1007/s12010-009-8823-5.
- Amano, M., Fearnside, P., Frangi, J., Frumhoff, P., Goldberg, D., Higuchi, N., et al. (2000). “Implications of Different Definitions and Generic Issues,” in *Land Use, Land-Use Change and Forestry*, eds. R. T. Watson, I. R. Noble, B. Bolin, N. Ravindranath, D. J. Dokken, and D. J. Verardo (Cambridge: Cambridge University Press), 375. Available at: [http://www.ipcc.ch/ipccreports/sres/land\\_use/index.php?idp=0](http://www.ipcc.ch/ipccreports/sres/land_use/index.php?idp=0).
- Angel, R., Pasternak, Z., Soares, M. I. M., Conrad, R., and Gillor, O. (2013). Active and total prokaryotic communities in dryland soils. *FEMS Microbiol. Ecol.* 86, 130–138. doi:10.1111/1574-6941.12155.
- Araújo, R., Casal, M., and Cavaco-Paulo, A. (2008). Application of enzymes for textile fibres processing. *Biocatal. Biotransformation* 26, 332–349. doi:10.1080/10242420802390457.
- Arias, M. E., Arenas, M. M., Rodríguez, J., Soliveri, J., Ball, A. S., Hernández, M., et al. (2003). Kraft pulp biobleaching and mediated oxidation of a nonphenolic substrate by laccase from *Streptomyces cyaneus* CECT 3335. *Appl. Environ. Microbiol.* 69, 1953–1958. doi:10.1128/AEM.69.4.1953-1958.2003.
- Arpigny, J. L., and Jaeger, K.-E. (1999). Bacterial lipolytic enzymes: classification and properties. *Biochem. J.* 343, 177. doi:10.1042/0264-6021:3430177.
- Abhauer, K. P., Wemheuer, B., Daniel, R., and Meinicke, P. (2015). Tax4Fun: Predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* 31, 2882–2884. doi:10.1093/bioinformatics/btv287.
- Augusto, L., De Schrijver, A., Vesterdal, L., Smolander, A., Prescott, C., and Ranger, J. (2015). Influences of evergreen gymnosperm and deciduous angiosperm tree species on the functioning of temperate and boreal forests. *Biol. Rev.* 90, 444–466. doi:10.1111/brv.12119.
- Augusto, L., Ranger, J., Binkley, D., and Rothe, A. (2002). Impact of several common tree species of European temperate forests on soil fertility. *Ann. For. Sci.* 59, 233–253. doi:10.1051/forest.
- Ausec, L., Zakrzewski, M., Goesmann, A., Schlüter, A., and Mandic-Mulec, I. (2011). Bioinformatic analysis reveals high diversity of bacterial genes for laccase-like enzymes. *PLoS One* 6. doi:10.1371/journal.pone.0025724.
- Babitzke, P., and O’Connor, M. (2017). Noncanonical translation initiation comes of age. *J. Bacteriol.* 199, 1–4. doi:10.1128/JB.00295-17.
- Bajpai, P. (2016). Pretreatment of Lignocellulosic Biomass for Biofuel Production. doi:10.1007/978-981-10-0687-6.
- Baker, B. J., Lazar, C. S., Teske, A. P., and Dick, G. J. (2015). Genomic resolution of linkages in carbon, nitrogen, and sulfur cycling among widespread estuary sediment bacteria. *Microbiome* 3, 14. doi:10.1186/s40168-015-0077-6.
- Baldrian, P. (2017). Forest microbiome: Diversity, complexity and dynamics. *FEMS Microbiol. Rev.* 41, 109–130. doi:10.1093/femsre/fuw040.

- Baldrian, P., Kolářík, M., Štursová, M., Kopecký, J., Valášková, V., Větrovský, T., et al. (2012). Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *ISME J.* 6, 248–258. doi:10.1038/ismej.2011.95.
- Balk, M., Keuskamp, J. A., and Laanbroek, H. J. (2015). Potential activity, size, and structure of sulfate-reducing microbial communities in an exposed, grazed and a sheltered, non-grazed mangrove stand at the Red Sea coast. *Front. Microbiol.* 6, 1–13. doi:10.3389/fmicb.2015.01478.
- Banik, J. J., and Brady, S. F. (2010). Recent application of metagenomic approaches toward the discovery of antimicrobials and other bioactive small molecules. *Curr. Opin. Microbiol.* 13, 603–609. doi:10.1016/j.mib.2010.08.012.
- Barabote, R. D., Xie, G., Leu, D. H., Normand, P., Necșulea, A., Daubin, V., et al. (2009). Complete genome of the cellulolytic thermophile *Acidothermus cellulolyticus* 11B provides insights into its ecophysiological and evolutionary adaptations. *Genome Res.* 19(6), 1033–1043. doi:10.1101/gr.084848.108.
- Barbier, S., Gosselin, F., and Balandier, P. (2008). Influence of tree species on understory vegetation diversity and mechanisms involved—A critical review for temperate and boreal forests. *For. Ecol. Manage.* 254, 1–15. doi:10.1016/j.foreco.2007.09.038.
- Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H-P., et al. (2016). Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiol. Mol. Biol. Rev.* 80, 1–43. doi:10.1128/MMBR.00019-15.
- Barone, R., De Santi, C., Palma Esposito, F., Tedesco, P., Galati, F., Visone, M., et al. (2014). Marine metagenomics, a valuable tool for enzymes and bioactive compounds discovery. *Front. Mar. Sci.* 1, 1–6. doi:10.3389/fmars.2014.00038.
- Bashir, Y., Pradeep Singh, S., and Kumar Konwar, B. (2014). Metagenomics: An Application Based Perspective. *Chinese J. Biol.* 2014, 1–7. doi:10.1155/2014/146030.
- Belova, S. E., Pankratova, T. A., Detkova, E. N., Kaparullina, E. N., and Dedysh, S. N. (2009). *Acidisoma tundrae* gen. nov., sp. nov. and *Acidisoma sibiricum* sp. nov., two acidophilic, psychrotolerant members of the Alphaproteobacteria from acidic northern wetlands. *Int. J. Syst. Evol. Microbiol.* 59, 2283–2290. doi:10.1099/ijs.0.009209-0.
- Berg, J., Tymoczko, J., and Stryer, L. (2012). “Der Fettsäurestoffwechsel,” in *Biochemie*, eds. J. Berg, J. Tymoczko, and L. Stryer (Berlin: Springer Spektrum), 645–681.
- Berger, T. W., and Berger, P. (2014). Does mixing of beech (*Fagus sylvatica*) and spruce (*Picea abies*) litter hasten decomposition? *Plant Soil* 377, 217–234. doi:10.1007/s11104-013-2001-9.
- Bergkemper, F., Welzl, G., Lang, F., Krüger, J., Schloter, M., and Schulz, S. (2016). The importance of C, N and P as driver for bacterial community structure in German beech dominated forest soils. *J. Plant Nutr. Soil Sci.* 179, 472–480. doi:10.1002/jpln.201600077.
- Berlemont, R., Spee, O., Delsaute, M., Lara, Y., Schuldes, J., Simon, C., et al. (2013). Novel organic solvent-tolerant esterase isolated by metagenomics: insights into the lipase/esterase classification. *Rev. Argent. Microbiol.* 45, 3–12.
- Blazewicz, S. J., Barnard, R. L., Daly, R. A., and Firestone, M. K. (2013). Evaluating rRNA as an indicator of microbial activity in environmental communities: Limitations and uses. *ISME J.* 7, 2061–2068. doi:10.1038/ismej.2013.102.
- Bonan, G. B. (2008). Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Fiorests. *Science* 320, 1444–1449. doi:10.1126/science.1155121.
- Bonito, G., Reynolds, H., Robeson, M. S., Nelson, J., Hodkinson, B. P., Tuskan, G., et al. (2014). Plant host and soil origin influence fungal and bacterial assemblages in the roots of woody plants. *Mol. Ecol.* 23, 3356–3370. doi:10.1111/mec.12821.

- Bornscheuer, U. T. (2002). Microbial carboxyl esterases: Classification, properties and application in biocatalysis. *FEMS Microbiol. Rev.* 26, 73–81. doi:10.1016/S0168-6445(01)00075-4.
- Bragg, L., Stone, G., Imelfort, M., Hugenholtz, P., and Tyson, G. W. (2012). Fast, accurate error-correction of amplicon pyrosequences using Acacia. *Nat. Methods* 9, 425–426. doi:10.1038/nmeth.1990.
- Brunet, J., Falkengren-Grerup, U., and Tyler, G. (1997). Pattern and dynamics of the ground vegetation in south Swedish Carpinus betulus forests: Importance of soil chemistry and management. *Ecography (Cop.)* 20, 513–520. doi:10.1111/j.1600-0587.1997.tb00420.x.
- Burke, C., Steinberg, P., Rusch, D. B., Kjelleberg, S., and Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. *Proc. Natl. Acad. Sci. USA* 108, 14288–14293. doi:10.1073/pnas.1101591108/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1101591108.
- Cain, M. I., Bowman, W. D., and Hacker, S. D. (2011). “Chapter 22: Nutrient supply and cycling,” in *Ecology, Third Edition*, eds. M. I. Cain, W. D. Bowman, and S. D. Hacker (Sunderland (Massachusetts): Sinauer Associates).
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nat. Publ. Gr.* 7, 335–336. doi:10.1038/nmeth0510-335.
- Cardenas, E., Kranabetter, J. M., Hope, G., Maas, K. R., Hallam, S., and Mohn, W. W. (2015). Forest harvesting reduces the soil metagenomic potential for biomass decomposition. *ISME J.* 9, 2465–2476. doi:10.1038/ismej.2015.57.
- Cardenas, E., Orellana, L. H., Konstantinidis, K. T., and Mohn, W. W. (2018). Effects of timber harvesting on the genetic potential for carbon and nitrogen cycling in five North American forest ecozones. *Sci. Rep.* 8, 1–13. doi:10.1038/s41598-018-21197-0.
- Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., et al. (2012). Biodiversity loss and its impact on humanity. *Nature* 486, 59–67. doi:10.1038/nature11148.
- Carvalho, F. M., Souza, R. C., Barcellos, F. G., Hungria, M., and Vasconcelos, A. T. R. (2010). Genomic and evolutionary comparisons of diazotrophic and pathogenic bacteria of the order Rhizobiales. *BMC Microbiol.* 10, 37. doi:10.1186/1471-2180-10-37.
- Castelle, C. J., and Banfield, J. F. (2018). Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. *Cell* 172, 1181–1197. doi:10.1016/j.cell.2018.02.016.
- Cavard, X., Bergeron, Y., Chen, H. Y. H., Paré, D., Laganière, J., and Brassard, B. (2011). Competition and facilitation between tree species change with stand development. *Oikos* 120, 1683–1695. doi:10.1111/j.1600-0706.2011.19294.x.
- Cesarz, S., Fender, A. C., Beyer, F., Valtanen, K., Pfeiffer, B., Gansert, D., et al. (2013a). Roots from beech (*Fagus sylvatica* L.) and ash (*Fraxinus excelsior* L.) differentially affect soil microorganisms and carbon dynamics. *Soil Biol. Biochem.* 61, 23–32. doi:10.1016/j.soilbio.2013.02.003.
- Cesarz, S., Ruess, L., Jacob, M., Jacob, A., Schaefer, M., and Scheu, S. (2013b). Tree species diversity versus tree species identity: Driving forces in structuring forest food webs as indicated by soil nematodes. *Soil Biol. Biochem.* 62, 36–45. doi:10.1016/j.soilbio.2013.02.020.
- Chakdar, H., Kumar, M., Pandiyan, K., Singh, A., Nanjappan, K., Kashyap, P. L., et al. (2016). Bacterial xylanases: biology to biotechnology. *3 Biotech* 6, 1–15. doi:10.1007/s13205-016-0457-z.
- Chang, L., Ding, M., Bao, L., Chen, Y., Zhou, J., and Lu, H. (2011). Characterization of a bifunctional xylanase/endoglucanase from yak rumen microorganisms. *Appl. Microbiol.*

- Biotechnol.* 90, 1933–1942. doi:10.1007/s00253-011-3182-x.
- Chen, H. (2014). "Chemical Composition and Structure of Natural Cellulose" in *Biotechnology of lignocellulose: Theory and practice*. doi:10.1007/978-94-007-6898-7.
- Chen, X., Liu, M., Hu, F., Mao, X., and Li, H. (2007). Contributions of soil micro-fauna (protozoa and nematodes) to rhizosphere ecological functions. *Acta Ecol. Sin.* 27, 3132–3143. doi:10.1016/S1872-2032(07)60068-7.
- Cheng, J., Pinnell, L., Engel, K., Neufeld, J. D., and Charles, T. C. (2014). Versatile broad-host-range cosmids for construction of high quality metagenomic libraries. *J. Microbiol. Methods* 99, 27–34. doi:10.1016/j.mimet.2014.01.015.
- Cherobaeva, A. S., Kizilova, A. K., Stepanov, A. L., and Kravchenko, I. K. (2011). Molecular analysis of the diversity of nitrifying bacteria in the soils of the forest and steppe zones of European Russia. *Microbiology* 80, 395–402. doi:10.1134/S0026261711030064.
- Chodak, M., Klimek, B., and Niklińska, M. (2016). Composition and activity of soil microbial communities in different types of temperate forests. *Biol. Fertil. Soils* 52, 1093–1104. doi:10.1007/s00374-016-1144-2.
- Ciesla, W. M. (2002). "An overview of temperate broadleaf forests," in *Non-wood forest products from temperate broad-leaved trees* (Rome: Food and Agriculture Organization of the United Nations).
- Cline, L. C., Zak, D. R., Upchurch, R. A., Freedman, Z. B., and Peschel, A. R. (2017). Soil microbial communities and elk foraging intensity: implications for soil biogeochemical cycling in the sagebrush steppe. *Ecol. Lett.* 20, 202–211. doi:10.1111/ele.12722.
- Cline, M. S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., et al. (2007). Integration of biological networks and gene expression data using cytoscape. *Nat. Protoc.* 2, 2366–2382. doi:10.1038/nprot.2007.324.Integration.
- Clivot, H., Pagnout, C., Aran, D., Devin, S., Bauda, P., Poupin, P., et al. (2012). Changes in soil bacterial communities following liming of acidified forests. *Appl. Soil Ecol.* 59, 116–123. doi:10.1016/j.apsoil.2011.09.010.
- Colin, Y., Nicolitch, O., Turpault, M. P., and Uroz, S. (2017). Mineral types and tree species determine the functional and taxonomic structures of forest soil bacterial communities. *Appl. Environ. Microbiol.* 83, 1–23. doi:10.1128/AEM.02684-16.
- Condit, R., Pitman, N. C. A., Jr, E. G. L., Chave, J. J., Terborgh, J. W., Foster, R. B., et al. (2002). Beta-diversity in tropical forest trees. *Science* 295, 666–9. doi:10.1126/science.1066854.
- Coughlan, L. M., Cotter, P. D., Hill, C., and Alvarez-Ordóñez, A. (2015). Biotechnological applications of functional metagenomics in the food and pharmaceutical industries. *Front. Microbiol.* 6, 1–22. doi:10.3389/fmicb.2015.00672.
- Creamer, R. E. E., Hannula, S. E. E., Leeuwen, J. P. V. P. V., Stone, D., Rutgers, M., Schmelz, R. M. M., et al. (2016). Ecological network analysis reveals the inter-connection between soil biodiversity and ecosystem function as affected by land use across Europe. *Appl. Soil Ecol.* 97, 112–124. doi:10.1016/j.apsoil.2015.08.006.
- Crowther, T. W., Glick, H. B., Covey, K. R., Bettigole, C., Maynard, D. S., Thomas, S. M., et al. (2015). Mapping tree density at a global scale. *Nature* 525, 201–205. doi:10.1038/nature14967.
- Dawid, W. (2000). Biology and global distribution of myxobacteria in soils. *FEMS Microbiol. Rev.* 24, 403–427. doi:10.1016/S0168-6445(00)00032-2.
- Dawud, S. M., Raulund-Rasmussen, K., Domisch, T., Finér, L., Jaroszewicz, B., and Vesterdal, L. (2016). Is Tree Species Diversity or Species Identity the More Important Driver of Soil Carbon Stocks, C/N Ratio, and pH? *Ecosystems* 19, 645–660. doi:10.1007/s10021-016-9958-1.
- De Cáceres, M. (2013). How to use the indicspecies package (ver. 1.7.1). *R Proj.*, 1–29. Available at: <https://cran.r->

project.org/web/packages/indicspecies/vignettes/indicspeciesTutorial.pdf.

- De Jaegere, T., Hein, S., and Claessens, H. (2016). A review of the characteristics of small-leaved lime (*Tilia cordata* Mill.) and their implications for silviculture in a changing climate. *Forests* 7, 1–21. doi:10.3390/f7030056.
- DeAngelis, K. M., Pold, G., Topçuoğlu, B. D., van Diepen, L. T. A., Varney, R. M., Blanchard, J. L., et al. (2015). Long-term forest soil warming alters microbial communities in temperate forest soils. *Front. Microbiol.* 6, 1–13. doi:10.3389/fmicb.2015.00104.
- Decocq, G., Aubert, M., Dupont, F., Alard, D., Saguez, R., Wattez-Franger, A., et al. (2004). Plant diversity in a managed temperate deciduous forest: understory response to two silvicultural systems. *J. Appl. Ecol.* 41, 1065–1079.
- Demain, A. L. (2000). Microbial Biotechnology. *Trends Biotechnol.* 18, 26–31. doi:10.1007/978-981-10-7140-9.
- Deshmukh, Y., and Sao, S. (2015). “Degradation of Lignin through Carbon Utilization by the Microbial Ligninolytic Enzymes for Environmental Management.” 27–31.
- Devetter, M., Háněl, L., Řeháková, K., and Doležal, J. (2017). Diversity and feeding strategies of soil microfauna along elevation gradients in Himalayan cold deserts. *PLoS One* 12, e0187646. doi:10.1371/journal.pone.0187646.
- Dilly, O., Bach, H. J., Buscot, F., Eschenbach, C., Kutsch, W. L., Middelhoff, U., et al. (2000). Characteristics and energetic strategies of the rhizosphere in ecosystems of the Bornhöved lake district. *Appl. Soil Ecol.* 15, 201–210. doi:10.1016/S0929-1393(00)00096-2.
- Dukunde, A., Schneider, D., Lu, M., Brady, S., and Daniel, R. (2017). A novel, versatile family IV carboxylesterase exhibits high stability and activity in a broad pH spectrum. *Biotechnol. Lett.* 39, 577–587. doi:10.1007/s10529-016-2282-1.
- Dwivedi, U. N., Singh, P., Pandey, V. P., and Kumar, A. (2011). Structure-function relationship among bacterial, fungal and plant laccases. *J. Mol. Catal. B Enzym.* 68, 117–128. doi:10.1016/j.molcatb.2010.11.002.
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi:10.1093/nar/gkh340.
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi:10.1093/bioinformatics/btq461.
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi:10.1093/bioinformatics/btr381.
- Eilers, K. G., Debenport, S., Anderson, S., and Fierer, N. (2012). Digging deeper to find unique microbial communities: The strong effect of depth on the structure of bacterial and archaeal communities in soil. *Soil Biol. Biochem.* 50, 58–65. doi:10.1016/j.soilbio.2012.03.011.
- Eisenhauer, N., Lanoue, A., Strecker, T., Scheu, S., Steinauer, K., Thakur, M. P., et al. (2017). Root biomass and exudates link plant diversity with soil bacterial and fungal biomass. *Sci. Rep.* 7, 1–8. doi:10.1038/srep44641.
- Elend, C., Schmeisser, C., Leggewie, C., Babiak, P., Steele, H. L., Reymond, J., et al. (2006). Isolation and Biochemical Characterization of Two Novel Metagenome-Derived Esterases. *Appl. Environ. Microbiol.* 72, 3637–45. doi:10.1128/AEM.72.5.3637.
- Falkengren-Grerup, U., Brunet, J., and Diekmann, M. (1998). Nitrogen mineralisation in deciduous forest soils in south Sweden in gradients of soil acidity and deposition. *Environ. Pollut.* 102, 415–420. doi:10.1016/S0269-7491(98)80062-6.
- Fang, W., Fang, Z., Zhou, P., Chang, F., Hong, Y., Zhang, X., et al. (2012). Evidence for Lignin Oxidation by the Giant Panda Fecal Microbiome. *PLoS One* 7, 1–10.



doi:10.1371/journal.pone.0050312.

- FAO (2006). Global Forest Resources Assessment 2005. Rome Available at: <http://www.fao.org/docrep/008/a0400e/a0400e00.htm>.
- FAO (2016). *Global Forest Resources Assessment 2015*. Second edi. Rome: Food and Agriculture Organization of the United Nations doi:10.1002/2014GB005021.
- Farag, I. F., Youssef, N. H., and Elshahed, M. S. (2017). Global distribution patterns and pangenomic diversity of the candidate phylum “Latescibacteria” (WS3). *Appl. Environ. Microbiol.* 83, 1–21. doi:10.1128/AEM.00521-17.
- Felske, A., Wolterink, A., Lis, R. Van, Antoon, D., Akkermans, L., and Lis, R. V. a N. (1998). Phylogeny of the Main Bacterial 16S rRNA Sequences in Drentse A Grassland Soils (The Netherlands). *Appl. Environ. Microbiol.* 64, 871–879.
- Ferrenberg, S., O’neill, S. P., Knelman, J. E., Todd, B., Duggan, S., Bradley, D., et al. (2013). Changes in assembly processes in soil bacterial communities following a wildfire disturbance. *ISME J.* 7, 1102–1111. doi:10.1038/ismej.2013.11.
- Fierer, N., Grandy, A. S., Six, J., and Paul, E. A. (2009). Searching for unifying principles in soil ecology. *Soil Biol. Biochem.* 41, 2249–2256. doi:10.1016/j.soilbio.2009.06.009.
- Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., et al. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc. Natl. Acad. Sci.* 109, 21390–21395. doi:10.1073/pnas.1215210110.
- Finzi, A. C., Austin, A. T., Cleland, E. E., Frey, S. D., Houlton, B. Z., and Wallenstein, M. D. (2011). Responses and feedbacks of coupled biogeochemical cycles to climate change: examples from terrestrial ecosystems. *Front. Ecol. Environ.* 9, 61–67. doi:10.1890/100001.
- Flores, G. E., Hunter, R. C., Liu, Y., Mets, A., Schouten, S., and Reysenbach, A. L. (2012). *Hippea jasoniae* sp. nov. and *Hippea alviniae* sp. nov., thermoacidophilic members of the class Deltaproteobacteria isolated from deep-sea hydrothermal vent deposits. *Int. J. Syst. Evol. Microbiol.* 62, 1252–1258. doi:10.1099/ijs.0.033001-0.
- Fridley, J. D. (2012). Extended leaf phenology and the autumn niche in deciduous forest invasions. *Nature* 485, 359–362. doi:10.1038/nature11056.
- Fudou, R., Iizuka, T., Sato, S., Ando, T., Shimba, N., and Yamanaka, S. (2001). Haliangicin, a novel antifungal metabolite produced by a marine myxobacterium. 2. Isolation and structural elucidation. *J. Antibiot. (Tokyo)*. 54, 153–156. doi:10.7164/antibiotics.54.153.
- Fudou, R., Jojima, Y., Iizuka, T., and Yamanaka, S. (2002). *Haliangium ochraceum* gen. nov., sp. nov. and *Haliangium tepidum* sp. nov.: novel moderately halophilic myxobacteria isolated from coastal saline environments. *J. Gen. Appl. Microbiol.* 48, 109–116. doi:10.2323/jgam.48.109.
- Gaida, S. M., Sandoval, N. R., Nicolaou, S. A., Chen, Y., Venkataramanan, K. P., and Papoutsakis, E. T. (2015). Expression of heterologous sigma factors enables functional screening of metagenomic and heterologous genomic libraries. *Nat. Commun.* 6, 1–10. doi:10.1038/ncomms8045.
- Gandolfo, M. A., Hermsen, E. J., Zamalao, M. C., Nixon, K. C., González, C. C., Wilf, P., et al. (2011). Oldest known eucalyptus macrofossils are from south america. *PLoS One* 6, e21084. doi:10.1371/journal.pone.0021084.
- Garcia, R. A. M., Pereira, M. R., Maester, T. C., and de Macedo Lemos, E. G. (2015). Investigation, Expression, and Molecular Modeling of ORF2, a Metagenomic Lipolytic Enzyme. *Appl. Biochem. Biotechnol.* 175, 3875–3887. doi:10.1007/s12010-015-1556-8.
- Gasteiger, E., C., Hoogland, A., Gattiker, S., Duvaud, M. R., Wilkins, R. D., And, A., et al. (2005). Protein Identification and Analysis Tools on the

- ExPASy Server. *Proteomics Protoc. Handb.*, 571–608.
- Gergs, R., and Rothhaupt, K-O. (2015). Invasive species as driving factors for the structure of benthic communities in Lake Constance, Germany. *Hydrobiologia* 746, 245–254. doi:10.1007/s10750-014-1931-4.
- Giles, M., Morley, N., Baggs, E. M., and Daniell, T. J. (2012). Soil nitrate reducing processes - Drivers, mechanisms for spatial variation, and significance for nitrous oxide production. *Front. Microbiol.* 3, 1–16. doi:10.3389/fmicb.2012.00407.
- Gill, A. S., Lee, A., and McGuire, K. L. (2017). Phylogenetic and functional diversity of total (DNA) and expressed (RNA) bacterial communities in urban green infrastructure bioswale soils. *Appl. Environ. Microbiol.* 83. doi:10.1128/AEM.00287-17.
- Gilliam, F. S. (2007). The Ecological Significance of the Herbaceous Layer in Temperate Forest Ecosystems. *Bioscience* 57, 845–858. doi:10.1641/B571007.
- Goodale, C. L., Fredriksen, G., Weiss, M. S., McCalley, C. K., Sparks, J. P., and Thomas, S. A. (2015). Soil processes drive seasonal variation in retention of in a deciduous forest catchment N tracers. *Ecology* 96, 2653–2668.
- Green, S. J., Prakash, O., Jasrotia, P., Overholt, W. A., Cardenas, E., Hubbard, D., et al. (2012). Denitrifying bacteria from the genus *Rhodanobacter* dominate bacterial communities in the highly contaminated subsurface of a nuclear legacy waste site. *Appl. Environ. Microbiol.* 78, 1039–1047. doi:10.1128/AEM.06435-11.
- Guckland, A. (2009). Nutrient stocks, acidity, processes of N transformation and net uptake of methane in soils of a temperate deciduous forest with different abundance of beech (*Fagus sylvatica* L.). 1–121.
- Guckland, A., Jacob, M., Flessa, H., Thomas, F. M., and Leuschne, C. (2009). Acidity, nutrient stocks, and organic-matter content in soils of a temperate deciduous forest with different abundance of European beech (*Fagus sylvatica* L.). *J. Plant Nutr. Soil Sci.* 172, 500–511. doi:10.1002/jpln.200800072.
- Haichar, F. E. Z., Achouak, W., Christen, R., Heulin, T., Marol, C., Marais, M. F., et al. (2007). Identification of cellulolytic bacteria in soil by stable isotope probing. *Environ. Microbiol.* 9, 625–634. doi:10.1111/j.1462-2920.2006.01182.x.
- Hakemian, A. S., and Rosenzweig, A. C. (2007). The Biochemistry of Methane Oxidation. *Annu. Rev. Biochem.* 76, 223–241. doi:10.1146/annurev.biochem.76.061505.175355.
- Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., and Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5, R245–R249. doi:10.1016/S1074-5521(98)90108-9.
- Hartmann, M., Frey, B., Mayer, J., Mäder, P., and Widmer, F. (2015). Distinct soil microbial diversity under long-term organic and conventional farming. *ISME J.* 9, 1177–1194. doi:10.1038/ismej.2014.210.
- Hausmann, B., Knorr, K. H., Schreck, K., Tringe, S. G., Glavina Del Rio, T., Loy, A., et al. (2016). Consortia of low-abundance bacteria drive sulfate reduction-dependent degradation of fermentation products in peat soil microcosms. *ISME J.* 10, 2365–2375. doi:10.1038/ismej.2016.42.
- Henne, A., Schmitz, R. a, Bömeke, M., Daniel, R., and Gottschalk, G. (2000). Screening of Environmental DNA Libraries for the Presence of Genes Conferring Lipolytic Activity on *Escherichia coli* Screening of Environmental DNA Libraries for the Presence of Genes Conferring Lipolytic Activity on *Escherichia coli*. *Appl. Environ. Microbiol.* 66, 3113–3116. doi:10.1128/AEM.66.7.3113-3116.2000.Updated.
- Herzog, S., Wemheuer, F., Wemheuer, B., and

- Daniel, R. (2015). Effects of fertilization and sampling time on composition and diversity of entire and active bacterial communities in German grassland soils. *PLoS One* 10, e0145575. doi:10.1371/journal.pone.0145575.
- Hjort, K., Bergström, M., Adesina, M. F., Jansson, J. K., Smalla, K., and Sjöling, S. (2010). Chitinase genes revealed and compared in bacterial isolates, DNA extracts and a metagenomic library from a phytopathogen-suppressive soil. *FEMS Microbiol. Ecol.* 71, 197–207. doi:10.1111/j.1574-6941.2009.00801.x.
- Hoppe, B., Krüger, D., Kahl, T., Arnstadt, T., Buscot, F., Bauhus, J., et al. (2015). A pyrosequencing insight into sprawling bacterial diversity and community dynamics in decaying deadwood logs of *Fagus sylvatica* and *Picea abies*. *Sci. Rep.* 5, 23–25. doi:10.1038/srep09456.
- Huber, K. J., Geppert, A. M., Wanner, G., Fösel, B. U., Wüst, P. K., and Overmann, J. (2016). The first representative of the globally widespread subdivision 6 Acidobacteria, *Vicinamibacter silvestris* gen. nov., sp. nov., isolated from subtropical savannah soil. *Int. J. Syst. Evol. Microbiol.* 66, 2971–2979. doi:10.1099/ijsem.0.001131.
- Ilmberger, N., Meske, D., Juergensen, J., Schulte, M., Barthen, P., Rabausch, U., et al. (2012). Metagenomic cellulases highly tolerant towards the presence of ionic liquids - Linking thermostability and halotolerance. *Appl. Microbiol. Biotechnol.* 95, 135–146. doi:10.1007/s00253-011-3732-2.
- Jacob, A., Hertel, D., and Leuschner, C. (2013). On the significance of belowground overyielding in temperate mixed forests: Separating species identity and species diversity effects. *Oikos* 122, 463–473. doi:10.1111/j.1600-0706.2012.20476.x.
- Jacob, M., Viedenz, K., Polle, A., and Thomas, F. M. (2010). Leaf litter decomposition in temperate deciduous forest stands with a decreasing fraction of beech (*Fagus sylvatica*). *Oecologia* 164, 1083–1094. doi:10.1007/s00442-010-1699-9.
- Jacob, M., Weland, N., Platner, C., Schaefer, M., Leuschner, C., and Thomas, F. M. (2009). Nutrient release from decomposing leaf litter of temperate deciduous forest trees along a gradient of increasing tree species diversity. *Soil Biol. Biochem.* 41, 2122–2130. doi:10.1016/j.soilbio.2009.07.024.
- Jacquioud, S., Demanèche, S., Franqueville, L., Ausec, L., Xu, Z., Delmont, T. O., et al. (2014). Characterization of new bacterial catabolic genes and mobile genetic elements by high throughput genetic screening of a soil metagenomic library. *J. Biotechnol.* 190, 18–29. doi:10.1016/j.jbiotec.2014.03.036.
- Javed, S., Azeem, F., Hussain, S., Rasul, I., Siddique, M. H., Riaz, M., et al. (2018). Bacterial lipases: A review on purification and characterization. *Prog. Biophys. Mol. Biol.* 132, 23–34. doi:10.1016/j.pbiomolbio.2017.07.014.
- Jeanbille, M., Buée, M., Bach, C., Cébron, A., Frey-Klett, P., Turpault, M. P., et al. (2016). Soil Parameters Drive the Structure, Diversity and Metabolic Potentials of the Bacterial Communities Across Temperate Beech Forest Soil Sequences. *Microb. Ecol.* 71, 482–493. doi:10.1007/s00248-015-0669-5.
- Johnson, D. B., Stallwood, B., Kimura, S., and Hallberg, K. B. (2006). Isolation and characterization of *Acidicaldus organivorus*, gen. nov., sp. nov.: A novel sulfur-oxidizing, ferric iron-reducing thermo-acidophilic heterotrophic Proteobacterium. *Arch. Microbiol.* 185, 212–221. doi:10.1007/s00203-006-0087-7.
- Jones, R. T., Robeson, M. S., Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME J.* 3, 442–453. doi:10.1038/ismej.2008.127.
- Kaeberlein, T. (2002). Isolating “Uncultivable” Microorganisms in Pure Culture in a Simulated Natural Environment. *Science* 296, 1127–1129. doi:10.1126/science.1070633.

- Kaiser, C., Koranda, M., Kitzler, B., Fuchslueger, L., Schnecker, J., Schweiger, P., et al. (2010). Belowground carbon allocation by trees drives seasonal patterns of extracellular enzyme activities by altering microbial community composition in a beech forest soil. *New Phytol.* 187, 843–858. doi:10.1111/j.1469-8137.2010.03321.x.
- Kaiser, K., Wemheuer, B., Korolkow, V., Wemheuer, F., Nacke, H., Schöning, I., et al. (2016). Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. *Sci. Rep.* 6, 1–12. doi:10.1038/srep33696.
- Kakirde, K. S., Parsley, L. C., and Liles, M. R. (2010). Size does matter: Application-driven approaches for soil metagenomics. *Soil Biol. Biochem.* 42, 1911–1923. doi:10.1016/j.soilbio.2010.07.021.
- Khelifa, R., Paquette, A., Messier, C., Reich, P. B., and Munson, A. D. (2017). Do temperate tree species diversity and identity influence soil microbial community function and composition? *Ecol. Evol.* 7, 7965–7974. doi:10.1002/ece3.3313.
- Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., van Veen, J. A., and Kuramae, E. E. (2016a). The ecology of Acidobacteria: Moving beyond genes and genomes. *Front. Microbiol.* 7, 1–16. doi:10.3389/fmicb.2016.00744.
- Kielak, A. M., Scheublin, T. R., Mendes, L. W., van Veen, J. A., and Kuramae, E. E. (2016b). Bacterial community succession in pine-wood decomposition. *Front. Microbiol.* 7, 1–12. doi:10.3389/fmicb.2016.00231.
- Kim, J. S., Lee, K. C., Kim, D. S., Ko, S. H., Jung, M. Y., Rhee, S. K., et al. (2015). Pyrosequencing analysis of a bacterial community associated with lava-formed soil from the Gotjawal forest in Jeju, Korea. *Microbiologyopen* 4, 301–312. doi:10.1002/mbo3.238.
- Kim, S. J., Ahn, J. H., Lee, T. H., Weon, H. Y., Hong, S. B., Seok, S. J., et al. (2013). *Reyranella soli* sp. nov., isolated from forest soil, and emended description of the genus *Reyranella* Pagnier et al. 2011. *Int. J. Syst. Evol. Microbiol.* 63, 3164–3167. doi:10.1099/ijs.0.045922-0.
- Kimoto, K. I., Aizawa, T., Urai, M., Ve, N. B., Suzuki, K. I., Nakajima, M., et al. (2010). *Acidocella aluminiidurans* sp. nov., an aluminiumtolerant bacterium isolated from *Panicum repens* grown in a highly acidic swamp in actual acid sulfate soil area of Vietnam. *Int. J. Syst. Evol. Microbiol.* 60, 764–768. doi:10.1099/ijs.0.011569-0.
- Klimek, B., Chodak, M., Jaźwa, M., Solak, A., Tarasek, A., and Niklińska, M. (2016). The relationship between soil bacteria substrate utilisation patterns and the vegetation structure in temperate forests. *Eur. J. For. Res.* 135, 179–189. doi:10.1007/s10342-015-0929-4.
- Kolton, M., Sela, N., Elad, Y., and Cytryn, E. (2013). Comparative Genomic Analysis Indicates that Niche Adaptation of Terrestrial Flavobacteria Is Strongly Linked to Plant Glycan Metabolism. *PLoS One* 8, 1–11. doi:10.1371/journal.pone.0076704.
- Koyama, A., Wallenstein, M. D., Simpson, R. T., and Moore, J. C. (2014). Soil bacterial community composition altered by increased nutrient availability in Arctic tundra soils. *Front. Microbiol.* 5, 1–16. doi:10.3389/fmicb.2014.00516.
- Kubisch, P., Hertel, D., and Leuschner, C. (2015). Do ectomycorrhizal and arbuscular mycorrhizal temperate tree species systematically differ in root order-related fine root morphology and biomass? *Front. Plant Sci.* 6, 1–12. doi:10.3389/fpls.2015.00064.
- Kumar, R., Singh, S., and Singh, O. V. (2008). Bioconversion of lignocellulosic biomass: Biochemical and molecular perspectives. *J. Ind. Microbiol. Biotechnol.* 35, 377–391. doi:10.1007/s10295-008-0327-8.
- Laforest-Lapointe, I., Messier, C., and Kembel, S. W. (2016). Tree phyllosphere bacterial communities: exploring the magnitude of intra- and inter-individual variation among host species. *PeerJ* 4, e2367. doi:10.7717/peerj.2367.

- Lam, K. N., Cheng, J., Engel, K., Neufeld, J. D., and Charles, T. C. (2015). Current and future resources for functional metagenomics. *Front. Microbiol.* 6, 1–8. doi:10.3389/fmicb.2015.01196.
- Lang, C. (2008). *Diversität der Ektomykorrhizen in verschieden artenreichen Laubbaumbeständen im Nationalpark Hainich (Thüringen)*. Available at: <http://univerlag.uni-goettingen.de/handle/3/isbn-978-3-940344-31-1>.
- Lang, C., and Polle, A. (2011). Ectomycorrhizal fungal diversity, tree diversity and root nutrient relations in a mixed Central European forest. *Tree Physiol.* 31, 531–538. doi:10.1093/treephys/tpr042.
- Lang, F., Bauhus, J., Frossard, E., George, E., Kaiser, K., Kaupenjohann, M., et al. (2016). Phosphorus in forest ecosystems: New insights from an ecosystem nutrition perspective. *J. Plant Nutr. Soil Sci.* 2015, 129–135. doi:10.1002/jpln.201500541.
- Langenbruch, C., Helfrich, M., and Flessa, H. (2012). Effects of beech (*Fagus sylvatica*), ash (*Fraxinus excelsior*) and lime (*Tilia spec.*) on soil chemical properties in a mixed deciduous forest. *Plant Soil* 352, 389–403. doi:10.1007/s11104-011-1004-7.
- Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.* 75, 5111–5120. doi:10.1128/AEM.00335-09.
- Lauber, C. L., Strickland, M. S., Bradford, M. A., and Fierer, N. (2008). The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biol. Biochem.* 40, 2407–2415. doi:10.1016/j.soilbio.2008.05.021.
- Li, S. guang, Zhao, L., Han, K., Li, P. fei, Li, Z. feng, Hu, W., et al. (2014a). Diversity of epothilone producers among *Sorangium* strains in producer-positive soil habitats. *Microb. Biotechnol.* 7, 130–141. doi:10.1111/1751-7915.12103.
- Li, Y., Chen, L., Wen, H., Zhou, T., Zhang, T., and Gao, X. (2014b). 454 Pyrosequencing Analysis of Bacterial Diversity revealed by a Comparative Study of Soils from mining Subsidence and Reclamation Areas. *J. Microbiol. Biotechnol.* 24, 313–323.
- Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., et al. (2015a). A new antibiotic kills pathogens without detectable resistance. *Nature* 517, 455–459. doi:10.1038/nature14098.
- Ling, Y. C., Bush, R., Grice, K., Tulipani, S., Berwick, L., and Moreau, J. W. (2015b). Distribution of iron- and sulfate-reducing bacteria across a coastal acid sulfate soil (CASS) environment: Implications for passive bioremediation by tidal inundation. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00624.
- Liu, Z., Lozupone, C., Hamady, M., Bushman, F. D., and Knight, R. (2007). Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res.* 35. doi:10.1093/nar/gkm541.
- Lladó, S., López-Mondéjar, R., and Baldrian, P. (2017). Forest Soil Bacteria: Diversity, Involvement in Ecosystem Processes, and Response to Global Change. *Microbiol. Mol. Biol. Rev.* 81, e00063-16. doi:10.1128/MMBR.00063-16.
- Lladó, S., Žifčáková, L., Větrovský, T., Eichlerová, I., and Baldrian, P. (2016). Functional screening of abundant bacteria from acidic forest soil indicates the metabolic potential of Acidobacteria subdivision 1 for polysaccharide decomposition. *Biol. Fertil. Soils* 52, 251–260. doi:10.1007/s00374-015-1072-6.
- Lopez-Lopez, O., Cerdan, M., and Siso, M. (2014). New Extremophilic Lipases and Esterases from Metagenomics. *Curr. Protein Pept. Sci.* 15, 445–455. doi:10.2174/1389203715666140228153801.
- López-Mondéjar, R., Voříšková, J., Větrovský, T., and Baldrian, P. (2015). The bacterial

- community inhabiting temperate deciduous forests is vertically stratified and undergoes seasonal dynamics. *Soil Biol. Biochem.* 87, 43–50. doi:10.1016/j.soilbio.2015.04.008.
- López-Mondéjar, R., Zühlke, D., Becher, D., Riedel, K., and Baldrian, P. (2016). Cellulose and hemicellulose decomposition by forest soil bacteria proceeds by the action of structurally variable enzymatic systems. *Sci. Rep.* 6, 1–12. doi:10.1038/srep25279.
- Lorenz, P., and Eck, J. (2005). Outlook: Metagenomics and industrial applications. *Nat. Rev. Microbiol.* 3, 510–516. doi:10.1038/nrmicro1161.
- Lorenz, P., Liebeton, K., Niehaus, F., and Eck, J. (2002). Screening for novel enzymes for biocatalytic processes: Accessing the metagenome as a resource of novel functional sequence space. *Curr. Opin. Biotechnol.* 13, 572–577. doi:10.1016/S0958-1669(02)00345-2.
- Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., and Knight, R. (2011). UniFrac: An effective distance metric for microbial community comparison. *ISME J.* 5, 169–172. doi:10.1038/ismej.2010.133.
- Lu, H., Mohren, G. M. J., den Ouden, J., Goudiaby, V., and Sterck, F. J. (2016). Overyielding of temperate mixed forests occurs in evergreen–deciduous but not in deciduous–deciduous species mixtures over time in the Netherlands. *For. Ecol. Manage.* 376, 321–332. doi:10.1016/j.foreco.2016.06.032.
- Lüneberg, K., Schneider, D., Siebe, C., and Daniel, R. (2018). Drylands soil bacterial community is affected by land use change and different irrigation practices in the Mezquital Valley, Mexico. *Sci. Rep.* 8, 1–15. doi:10.1038/s41598-018-19743-x.
- MacDicken, K. G. (2015). Global Forest Resources Assessment 2015: What, why and how? *For. Ecol. Manage.* 352, 3–8. doi:10.1016/j.foreco.2015.02.006.
- MacDicken, K. G., Reams, G., and Freitas, J. De (2015). Forest Ecology and Management Introduction to the Changes in Global Forest Resources from 1990 to. *For. Ecol. Manage.* 352, 1–2. doi:10.1016/j.foreco.2015.06.018.
- Malhi, Y., Baldocchi, D. D., and Jarvis, P. G. (1999). The carbon balance of tropical, temperate and boreal forests. *Plant. Cell Environ.* 22, 715–740. doi:10.1046/j.1365-3040.1999.00453.x.
- Mamlouk, D., and Gullo, M. (2013). Acetic Acid Bacteria: Physiology and Carbon Sources Oxidation. *Indian J. Microbiol.* 53, 377–384. doi:10.1007/s12088-013-0414-z.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12
- Mendes, L. W., Kuramae, E. E., Navarrete, A. A., Van Veen, J. A., and Tsai, S. M. (2014). Taxonomical and functional microbial community selection in soybean rhizosphere. *ISME J.* 8, 1577–1587. doi:10.1038/ismej.2014.17.
- Mendes, R., Garbeva, P., and Raaijmakers, J. M. (2013). The rhizosphere microbiome: Significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol. Rev.* 37, 634–663. doi:10.1111/1574-6976.12028.
- Minussi, R. C., Pastore, G. M., and Durán, N. (2002). Potential applications of laccase in the food industry. *Trends Food Sci. Technol.* 13, 205–216. doi:10.1016/S0924-2244(02)00155-3.
- Mirete, S., Morgante, V., and González-Pastor, J. E. (2016). Functional metagenomics of extreme environments. *Curr. Opin. Biotechnol.* 38, 143–149. doi:10.1016/j.copbio.2016.01.017.
- Miyashita, N. T. (2015). Contrasting soil bacterial community structure between the phyla Acidobacteria and Proteobacteria in tropical Southeast Asian and temperate Japanese forests. *Genes Genet. Syst.* 90, 61–77. doi:10.1266/ggs.90.61.
- Mohagheghi, A., Grohmann, K., Himmel, M., Leighton, L., and Updegraff, D. M. (1986). Isolation and Characterization of *Acidothermus cellulolyticus* gen. nov., sp. nov., a New Genus

- of Thermophilic, Acidophilic, Cellulolytic Bacteria. *Int. J. Syst. Bacteriol.* 36, 435–443.
- Mölder, A., Bernhardt-Römermann, M., and Schmidt, W. (2006). Forest ecosystem research in Hainich National Park (Thuringia) First results on flora and vegetation in stands with contrasting tree species diversity. *Waldökologie online* 3, 83–99.
- Moran, M. A., Satinsky, B., Gifford, S. M., Luo, H., Rivers, A., Chan, L. K., et al. (2013). Sizing up metatranscriptomics. *ISME J.* 7, 237–243. doi:10.1038/ismej.2012.94.
- Morawe, M., Hoeke, H., Wissenbach, D. K., Lentendu, G., Wubet, T., Kröber, E., et al. (2017). Acidotolerant Bacteria and Fungi as a Sink of Methanol-Derived Carbon in a Deciduous Forest Soil. *Front. Microbiol.* 8, 1–18. doi:10.3389/fmicb.2017.01361.
- Mtui, G. Y. S. (2012). Lignocellulolytic enzymes from tropical fungi: Types, substrates and applications. *Sci. Res. Essays* 7, 1544–1555. doi:10.5897/SRE11.1812.
- Nacke, H. (2011). Identification and characterization of microbial key functions in soils of the German Biodiversity Exploratories representing different land use and management types. 132.
- Nacke, H., Engelhaupt, M., Brady, S., Fischer, C., Tautz, J., and Daniel, R. (2012). Identification and characterization of novel cellulolytic and hemicellulolytic genes and enzymes derived from German grassland soil metagenomes. *Biotechnol. Lett.* 34, 663–675. doi:10.1007/s10529-011-0830-2.
- Nacke, H., Goldmann, K., Schöning, I., Pfeiffer, B., Kaiser, K., Villamizar, G. A. C., et al. (2016). Fine spatial scale variation of soil microbial communities under European beech and Norway spruce. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.02067.
- Nacke, H., Will, C., Herzog, S., Nowka, B., Engelhaupt, M., and Daniel, R. (2011). Identification of novel lipolytic genes and gene families by screening of metagenomic libraries derived from soil samples of the German Biodiversity Exploratories. *FEMS Microbiol. Ecol.* 78, 188–201. doi:10.1111/j.1574-6941.2011.01088.x.
- Naether, A., Foesel, B. U., Naegele, V., Wüst, P. K., Weinert, J., Bonkowski, M., et al. (2012). Environmental factors affect acidobacterial communities below the subgroup level in grassland and forest soils. *Appl. Environ. Microbiol.* 78, 7398–7406. doi:10.1128/AEM.01325-12.
- Naushad, S., Adeolu, M., Wong, S., Sohail, M., Schellhorn, H. E., and Gupta, R. S. (2015). A phylogenomic and molecular marker based taxonomic framework for the order Xanthomonadales: proposal to transfer the families Algiphilaceae and Solimonadaceae to the order Nevskiales ord. nov. and to create a new family within the order Xanthomonadales, the family Rhodanobacteraceae fam. nov., containing the genus Rhodanobacter and its closest relatives. *Antonie Van Leeuwenhoek* 107, 467–485. doi:10.1007/s10482-014-0344-8.
- Nemergut, D., Shade, A., and Violle, C. (2014). When, where and how does microbial community composition matter? *Front. Microbiol.* 5, 2012–2014. doi:10.3389/fmicb.2014.00497.
- Nichols, D., Cahoon, N., Trakhtenberg, E. M., Pham, L., Mehta, A., Belanger, A., et al. (2010). Use of ichip for high-throughput in situ cultivation of "uncultivable microbial species. *Appl. Environ. Microbiol.* 76, 2445–2450. doi:10.1128/AEM.01754-09.
- Norton, J. M., Alzerreca, J. J., Suwa, Y., and Klotz, M. G. (2002). Diversity of ammonia monooxygenase operon in autotrophic ammonia-oxidizing bacteria. *Arch. Microbiol.* 177, 139–149. doi:10.1007/s00203-001-0369-z.
- Obeng, E. M., Adam, S. N. N., Budiman, C., Ongkudon, C. M., Maas, R., and Jose, J. (2017). Lignocellulases: a review of emerging and developing enzymes, systems, and practices. *Bioresour. Bioprocess.* 4, 16. doi:10.1186/s40643-017-0146-8.

- Oh, Y. M., Kim, M., Lee-Cruz, L., Lai-Hoe, A., Go, R., Ainuddin, N., et al. (2012). Distinctive Bacterial Communities in the Rhizoplane of Four Tropical Tree Species. *Microb. Ecol.* 64, 1018–1027. doi:10.1007/s00248-012-0082-2.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., Hara, R. B. O., et al. (2015). *vegan: Community Ecology Package*. R Package Version 2.2-1. Available online at: <https://cran.r-project.org/we>. doi:ISBN 0-387-95457-0.
- Pan, Y., Birdsey, R. A., Phillips, O. L., and Jackson, R. B. (2013). The Structure, Distribution, and Biomass of the World's Forests. *Annu. Rev. Ecol. Evol. Syst.* 44, 593–622. doi:10.1146/annurev-ecolsys-110512-135914.
- Pardo, I., Vicente, A. I., Mate, D. M., Alcalde, M., and Camarero, S. (2012). Development of chimeric laccases by directed evolution. *Biotechnol. Bioeng.* 109, 2978–2986. doi:10.1002/bit.24588.
- Park, S. H., Ong, R. G., and Sticklen, M. (2016). Strategies for the production of cell wall-deconstructing enzymes in lignocellulosic biomass and their utilization for biofuel production. *Plant Biotechnol. J.* 14, 1329–1344. doi:10.1111/pbi.12505.
- Parviainen, J., and Frank, G. (2003). Protected forests in Europe approaches-harmonising the definitions for international comparison and forest policy making. *J. Environ. Manage.* 67, 27–36. doi:10.1016/S0301-4797(02)00185-8.
- Peralta, R. M., Ahn, C., and Gillevet, P. M. (2013). Characterization of soil bacterial community structure and physicochemical properties in created and natural wetlands. *Sci. Total Environ.* 443, 725–732. doi:10.1016/j.scitotenv.2012.11.052.
- Petersen, T. N., Brunak, S., Von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi:10.1038/nmeth.1701.
- Pfeiffer, B., Fender, A. C., Lasota, S., Hertel, D., Jungkunst, H. F., and Daniel, R. (2013). Leaf litter is the main driver for changes in bacterial community structures in the rhizosphere of ash and beech. *Appl. Soil Ecol.* 72, 150–160. doi:10.1016/j.apsoil.2013.06.008.
- Pochon, X., Zaiko, A., Fletcher, L. M., Laroche, O., and Wood, S. A. (2017). Wanted dead or alive? Using metabarcoding of environmental DNA and RNA to distinguish living assemblages for biosecurity applications. *PLoS One* 12, 1–19. doi:10.1371/journal.pone.0187636.
- Prakash, O., Green, S. J., Jasrotia, P., Overholt, W. A., Canion, A., Watson, D. B., et al. (2012). *Rhodanobacter denitrificans* sp. nov., isolated from nitrate-rich zones of a contaminated aquifer. *Int. J. Syst. Evol. Microbiol.* 62, 2457–2462. doi:10.1099/ijs.0.035840-0.
- Prescott, C. E., and Grayston, S. J. (2013). Tree species influence on microbial communities in litter and soil: Current knowledge and research needs. *For. Ecol. Manage.* 309, 19–27. doi:10.1016/j.foreco.2013.02.034.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* 41, 590–596. doi:10.1093/nar/gks1219.
- Rabausch, U., Juergensen, J., Ilmberger, N., Böhnke, S., Fischer, S., Schubach, B., et al. (2013). Functional screening of metagenome and genome libraries for detection of novel flavonoid-modifying enzymes. *Appl. Environ. Microbiol.* 79, 4551–4563. doi:10.1128/AEM.01077-13.
- Rasche, F., Knapp, D., Kaiser, C., Koranda, M., Kitzler, B., Zechmeister-Boltenstern, S., et al. (2011). Seasonality and resource availability control bacterial and archaeal communities in soils of a temperate beech forest. *ISME J.* 5, 389–402. doi:10.1038/ismej.2010.138.
- Rashamuse, K. J., Visser, D. F., Hennessy, F., Kemp, J., Roux-Van Der Merwe, M. P., Badenhorst, J., et al. (2013). Characterisation of two



- bifunctional cellulase-Xylanase enzymes isolated from a bovine rumen metagenome library. *Curr. Microbiol.* 66, 145–151. doi:10.1007/s00284-012-0251-z.
- Reich, P. B., and Frelich, L. (2002). “Temperate Deciduous Forests,” in *Encyclopedia of Global Environmental Change*, eds. H. a. Mooney and J. G. Canadell (Chichester: John Wiley & Sons), 565–569.
- Ricklefs, E., Winkler, N., Koschorreck, K., and Urlacher, V. B. (2014). Expanding the laccase-toolbox: A laccase from *Corynebacterium glutamicum* with phenol coupling and cuprous oxidase activity. *J. Biotechnol.* 191, 46–53. doi:10.1016/j.jbiotec.2014.05.031.
- Riesenfeld, C. S., Schloss, P. D., and Handelsman, J. (2004). Metagenomics: Genomic Analysis of Microbial Communities. *Annu. Rev. Genet.* 38, 525–552. doi:10.1146/annurev.genet.38.072902.091216.
- Rime, T., Hartmann, M., and Frey, B. (2016). Potential sources of microbial colonizers in an initial soil ecosystem after retreat of an alpine glacier. *ISME J.* 10, 1625–1641. doi:10.1038/ismej.2015.238.
- Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, 320–324. doi:10.1093/nar/gku316.
- Rocca, J. D., Hall, E. K., Lennon, J. T., Evans, S. E., Waldrop, M. P., Cotner, J. B., et al. (2015). Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. *ISME J.* 9, 1693–1699. doi:10.1038/ismej.2014.252.
- Romanowicz, K. J., Freedman, Z. B., Upchurch, R. A., Argiroff, W. A., and Zak, D. R. (2016). Active microorganisms in forest soils differ from the total community yet are shaped by the same environmental factors: the influence of pH and soil moisture. *FEMS Microbiol. Ecol.* 92, 1–9. doi:10.1093/femsec/fw149.
- Saini, A., Aggarwal, N. K., Sharma, A., and Yadav, A. (2015). Actinomycetes: A Source of Lignocellulolytic Enzymes. *Enzyme Res.* 2015, 279381. doi:10.1155/2015/279381.
- Sait, M., Davis, K. E. R., and Janssen, P. H. (2006). Effect of pH on isolation and distribution of members of subdivision 1 of the phylum Acidobacteria occurring in soil. *Appl. Environ. Microbiol.* 72, 1852–1857. doi:10.1128/AEM.72.3.1852-1857.2006.
- Salehi, A., Amiri, G., Burslem, D., and Swaine, M. (2007). Relationships between tree species competition, soil properties and topographic factors in a temperate deciduous forest in northern Iran. *Asian J. Plant Sci.* 6, 455–462.
- Sambrook, J., Fritsch, E., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual, 2nd edition*. Second. , eds. F. Ausubel, R. Brent, R. Kingston, D. Moore, J. Seideman, J. Smith, et al. New York: Cold Spring Harbor Laboratory Press.
- Santhanam, N., Vivanco, J. M., Decker, S. R., and Reardon, K. F. (2011). Expression of industrially relevant laccases: Prokaryotic style. *Trends Biotechnol.* 29, 480–489. doi:10.1016/j.tibtech.2011.04.005.
- Scharff, L. B., Childs, L., Walther, D., and Bock, R. (2011). Local absence of secondary structure permits translation of mRNAs that lack ribosome-binding sites. *PLoS Genet.* 7. doi:10.1371/journal.pgen.1002155.
- Scheibe, A., Steffens, C., Seven, J., Jacob, A., Hertel, D., Leuschner, C., et al. (2015). Effects of tree identity dominate over tree diversity on the soil microbial community structure. *Soil Biol. Biochem.* 81, 219–227. doi:10.1016/j.soilbio.2014.11.020.
- Schlatter, D. C., Bakker, M. G., Bradeen, J. M., and Kinkel, L. L. (2015). Plant community richness and microbial interactions structure bacterial communities in soil. *Ecology* 96, 134–142. doi:10.1890/13-1648.1.
- Schlesinger, W. H., Cole, J. J., Finzi, A. C., and Holland, E. A. (2011). Introduction to coupled biogeochemical cycles. *Front. Ecol. Environ.* 9, 5–8. doi:10.1890/090235.

- Schloss, P. D., Girard, R. A., Martin, T., Edwards, J., and Thrash, J. C. (2016). Status of the archaeal and bacterial census: An update. *MBio* 7, 1–10. doi:10.1128/mBio.00201-16.
- Schloss, P. D., and Handelsman, J. (2003). Biotechnological prospects from metagenomics. *Curr. Opin. Biotechnol.* 14, 303–310. doi:10.1016/S0958-1669(03)00067-3.
- Schmidt, M., Veldkamp, E., and Corre, M. D. (2015). Tree species diversity effects on productivity, soil nutrient availability and nutrient response efficiency in a temperate deciduous forest. *For. Ecol. Manage.* 338, 114–123. doi:10.1016/j.foreco.2014.11.021.
- Schneider, D., Engelhaupt, M., Allen, K., Kurniawan, S., Krashevskaya, V., Heinemann, M., et al. (2015a). Impact of lowland rainforest transformation on diversity and composition of soil prokaryotic communities in Sumatra (Indonesia). *Front. Microbiol.* 6, 1–12. doi:10.3389/fmicb.2015.01339.
- Schneider, D., Reimer, A., Hahlbrock, A., Arp, G., and Daniel, R. (2015b). Metagenomic and Metatranscriptomic Analyses of Bacterial Communities Derived From a Calcifying Karst Water Creek Biofilm and Tufa. *Geomicrobiol. J.* 32, 316–331. doi:10.1080/01490451.2014.907376.
- Schulze, E. D., Aas, G., Grimm, G. W., Gossner, M. M., Walentowski, H., Ammer, C., et al. (2016). A review on plant diversity and forest management of European beech forests. *Eur. J. For. Res.* 135, 51–67. doi:10.1007/s10342-015-0922-y.
- Schulze, E. D. D. (2017). Effects of forest management on biodiversity in temperate deciduous forests: An overview based on Central European beech forests. *J. Nat. Conserv.*, 0–1. doi:10.1016/j.jnc.2017.08.001.
- Senécal, J. F., Doyon, F., and Messier, C. (2018). Tree death not resulting in gap creation: An investigation of canopy dynamics of northern temperate deciduous forests. *Remote Sens.* 10, 1–17. doi:10.3390/rs10010121.
- Seven, J. (2014). Mycorrhizal fungi in deciduous forests of differing tree species diversity and their role for nutrient transfer. Available at: <http://hdl.handle.net/11858/00-1735-0000-0022-5F2E-3>.
- Seven, J., and Polle, A. (2014). Subcellular nutrient element localization and enrichment in ecto- and arbuscular mycorrhizas of field-grown beech and ash trees indicate functional differences. *PLoS One* 9, 1–18. doi:10.1371/journal.pone.0114672.
- Shade, A., Jones, S. E., Caporaso, J. G., Shade, A., Jones, S. E., Caporaso, J. G., et al. (2014). Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity. *Am. Soc. Microbiol.* 5, 1–9. doi:10.1128/mBio.01371-14.Editor.
- Shange, R. S., Ankumah, R. O., Ibekwe, A. M., Zabawa, R., and Dowd, S. E. (2012). Distinct soil bacterial communities revealed under a diversely managed agroecosystem. *PLoS One* 7. doi:10.1371/journal.pone.0040338.
- Shannon, P., Markiel, A., Owen Ozier, 2, Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, 2498–2504. doi:10.1101/gr.1239303.metabolite.
- Sharma, S., and Kanwar, S. S. (2014). Organic solvent tolerant lipases and applications. *Sci. World J.* 2014. doi:10.1155/2014/625258.
- Shen, Z., Wang, D., Ruan, Y., Xue, C., Zhang, J., Li, R., et al. (2014). Deep 16S rRNA pyrosequencing reveals a bacterial community associated with banana Fusarium wilt disease suppression induced by bio-organic fertilizer application. *PLoS One* 9, 1–10. doi:10.1371/journal.pone.0098420.
- Shi, S., Richardson, A. E., O' Callaghan, M., DeAngelis, K. M., Jones, E. E., Stewart, A., et al. (2011). Effects of selected root exudate components on soil bacterial communities. *FEMS Microbiol. Ecol.* 77, 600–610. doi:10.1111/j.1574-6941.2011.01150.x.

- Siles, J. A., and Margesin, R. (2017). Seasonal soil microbial responses are limited to changes in functionality at two Alpine forest sites differing in altitude and vegetation. *Sci. Rep.* 7, 1–16. doi:10.1038/s41598-017-02363-2.
- Šnajdr, J., Dobiášová, P., Urbanová, M., Petránková, M., Cajthaml, T., Frouz, J., et al. (2013). Dominant trees affect microbial community composition and activity in post-mining afforested soils. *Soil Biol. Biochem.* 56, 105–115. doi:10.1016/j.soilbio.2012.05.004.
- Söderberg, K. H., Probanza, A., Jumpponen, A., and Bååth, E. (2004). The microbial community in the rhizosphere determined by community-level physiological profiles (CLPP) and direct soil- and cfu-PLFA techniques. *Appl. Soil Ecol.* 25, 135–145. doi:10.1016/j.apsoil.2003.08.005.
- Sohng, J., Han, A. R., Jeong, M. A., Park, Y., Park, B. B., and Park, P. S. (2014). Seasonal pattern of decomposition and N, P, and C dynamics in leaf litter in a mongolian oak forest and a Korean pine plantation. *Forests* 5, 2561–2580. doi:10.3390/f5102561.
- Sommer, J. (2016). Species-Specific Effects on the Carbon and Nitrogen Cycle in Temperate Deciduous Forest.
- Steele, H. L., Jaeger, K. E., Daniel, R., and Streit, W. R. (2009). Advances in recovery of novel biocatalysts from metagenomes. *J. Mol. Microbiol. Biotechnol.* 16, 25–37. doi:10.1159/000142892.
- Štursová, M., Žifčáková, L., Leigh, M. B., Burgess, R., and Baldrian, P. (2012). Cellulose utilization in forest litter and soil: Identification of bacterial and fungal decomposers. *FEMS Microbiol. Ecol.* 80, 735–746. doi:10.1111/j.1574-6941.2012.01343.x.
- Sudan, A. K., and Vakhlu, J. (2013). Isolation of a thioesterase gene from the metagenome of a mountain peak, Apharwat, in the northwestern Himalayas. *3 Biotech* 3, 19–27. doi:10.1007/s13205-012-0065-5.
- Thompson, I. D., Okabe, K., Tylianakis, J. M., Kumar, P., Brockerhoff, E. G., Schellhorn, N. A., et al. (2011). Forest Biodiversity and the Delivery of Ecosystem Goods and Services: Translating Science into Policy. *Bioscience* 61, 972–981. doi:10.1525/bio.2011.61.12.7.
- Thoms, C., Gattinger, A., Jacob, M., Thomas, F. M., and Gleixner, G. (2010). Direct and indirect effects of tree diversity drive soil microbial diversity in temperate deciduous forest. *Soil Biol. Biochem.* 42, 1558–1565. doi:10.1016/j.soilbio.2010.05.030.
- Thoms, C., and Gleixner, G. (2013). Seasonal differences in tree species' influence on soil microbial communities. *Soil Biol. Biochem.* 66, 239–248. doi:10.1016/j.soilbio.2013.05.018.
- Tkacz, A., and Poole, P. (2015). Role of root microbiota in plant productivity. *J. Exp. Bot.* 66, 2167–2175. doi:10.1093/jxb/erv157.
- Torsvik, V., and Øvreås, L. (2002). Microbial diversity and function in soil: from genes to ecosystems. *Curr. Opin. Microbiol.* 5, 240–245. doi:10.1016/S1369-5274(02)00324-7.
- Toushik, S. H., Lee, K. T., Lee, J. S., and Kim, K. S. (2017). Functional Applications of Lignocellulolytic Enzymes in the Fruit and Vegetable Processing Industries. *J. Food Sci.* 82, 585–593. doi:10.1111/1750-3841.13636.
- Townsend, A. R., Cleveland, C. C., Houlton, B. Z., Alden, C. B., and White, J. W. C. (2011). Multi-element regulation of the tropical forest carbon cycle. *Front. Ecol. Environ.* 9, 9–17. doi:10.1890/100047.
- Tripathi, B. M., Kim, M., Lai-Hoe, A., Shukor, N. A. A., Rahim, R. A., Go, R., et al. (2013). PH dominates variation in tropical soil archaeal diversity and community structure. *FEMS Microbiol. Ecol.* 86, 303–311. doi:10.1111/1574-6941.12163.
- Uchiyama, T., and Miyazaki, K. (2009). Functional metagenomics for enzyme discovery: challenges to efficient screening. *Curr. Opin. Biotechnol.* 20, 616–622. doi:10.1016/j.copbio.2009.09.010.
- Urbanová, M., Šnajdr, J., and Baldrian, P. (2015). Composition of fungal and bacterial communities in forest litter and soil is largely

- determined by dominant trees. *Soil Biol. Biochem.* 84, 53–64. doi:10.1016/j.soilbio.2015.02.011.
- Uroz, S., Buée, M., Deveau, A., Mieszkin, S., and Martin, F. (2016). Ecology of the forest microbiome: Highlights of temperate and boreal ecosystems. *Soil Biol. Biochem.* 103, 471–488. doi:10.1016/j.soilbio.2016.09.006.
- Uroz, S., Courty, P. E. P., Pierrat, J. C. J., Peter, M., Buée, M., Turpault, M. P., et al. (2013). Functional Profiling and Distribution of the Forest Soil Bacterial Communities Along the Soil Mycorrhizosphere Continuum. *Microb. Ecol.* 66, 404–415. doi:10.1007/s00248-013-0199-y.
- Uroz, S., Oger, P., Lepleux, C., Collignon, C., Frey-Klett, P., and Turpault, M. P. (2011). Bacterial weathering and its contribution to nutrient cycling in temperate forest ecosystems. *Res. Microbiol.* 162, 821–831. doi:10.1016/j.resmic.2011.01.013.
- van den Heuvel, R. N., van der Biezen, E., Jetten, M. S. M., Hefting, M. M., and Kartal, B. (2010). Denitrification at pH 4 by a soil-derived Rhodanobacter-dominated community. *Environ. Microbiol.* 12, 3264–3271. doi:10.1111/j.1462-2920.2010.02301.x.
- Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F., et al. (2007). Genomics of Actinobacteria: Tracing the Evolutionary History of an Ancient Phylum. *Microbiol. Mol. Biol. Rev.* 71, 495–548. doi:10.1128/MMBR.00005-07.
- Veresoglou, S. D., Halley, J. M., and Rillig, M. C. (2015). Extinction risk of soil biota. *Nat. Commun.* 6, 1–10. doi:10.1038/ncomms9862.
- Virk, A. P., Sharma, P., and Capalash, N. (2012). Use of laccase in pulp and paper industry. *Biotechnol. Prog.* 28, 21–32. doi:10.1002/btpr.727.
- Wang, W., Wang, H., Feng, Y., Wang, L., Xiao, X., Xi, Y., et al. (2016). Consistent responses of the microbial community structure to organic farming along the middle and lower reaches of the Yangtze River. *Sci. Rep.* 6, 1–11. doi:10.1038/srep35046.
- Wang, Y., and Qian, P. Y. (2009). Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One* 4. doi:10.1371/journal.pone.0007401.
- Wemheuer, B., Wemheuer, F., and Daniel, R. (2012). RNA-based assessment of diversity and composition of active archaeal communities in the German bight. *Archaea* 2012. doi:10.1155/2012/695826.
- Wemheuer, B., Wemheuer, F., Meier, D., Billerbeck, S., Giebel, H-A., Simon, M., et al. (2017). Linking Compositional and Functional Predictions to Decipher the Biogeochemical Significance in DFAA Turnover of Abundant Bacterioplankton Lineages in the North Sea. *Microorganisms* 5, 68. doi:10.3390/microorganisms5040068.
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag Available at: <http://ggplot2.org>.
- Wirth, S., and Ulrich, A. (2002). Cellulose-degrading potentials and phylogenetic classification of carboxymethylcellulose-decomposing bacteria isolated from soil. *Syst. Appl. Microbiol.* 25, 584–591.
- Wood, S. A., Gilbert, J. A., Leff, J. W., Fierer, N., D'Angelo, H., Bateman, C., et al. (2017). Consequences of tropical forest conversion to oil palm on soil bacterial community and network structure. *Soil Biol. Biochem.* 112, 258–268. doi:10.1016/j.soilbio.2017.05.019.
- Yousuf, B., Kumar, R., Mishra, A., and Jha, B. (2014). Unravelling the carbon and sulphur metabolism in coastal soil ecosystems using comparative cultivation-independent genome-level characterisation of microbial communities. *PLoS One* 9. doi:10.1371/journal.pone.0107025.
- Yun, J., Kang, S., Park, S., Yoon, H., Kim, M., Heu, S., et al. (2004). Characterization of a novel amylolytic enzyme encoded by a gene from a soil derived metagenomic library. *Appl. Environ. Microbiol.* 70, 7229–7235. doi:10.1128/AEM.70.12.7229.
- Zapras, A., Liu, Y. J., Liu, S. J., Drake, H. L., and

- Horn, M. A. (2010). Abundance of novel and diverse *tfdA*-like genes, encoding putative phenoxyalkanoic acid herbicide-degrading dioxygenases, in soil. *Appl. Environ. Microbiol.* 76, 119–128. doi:10.1128/AEM.01727-09.
- Zarafeta, D., Moschidi, D., Ladoukakis, E., Gavrilov, S., Chrysina, E. D., Chatziioannou, A., et al. (2016). Metagenomic mining for thermostable esterolytic enzymes uncovers a new family of bacterial esterases. *Sci. Rep.* 6, 1–16. doi:10.1038/srep38886.
- Zhalnina, K., Louie, K. B., Hao, Z., Mansoori, N., Nunes da Rocha, U., Shi, S., et al. (2018). Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat. Microbiol.* in press. doi:10.1038/s41564-018-0129-3.
- Žifčáková, L., Větrovský, T., Howe, A., and Baldrian, P. (2016). Microbial activity in forest soil reflects the changes in ecosystem properties between summer and winter. *Environ. Microbiol.* 18, 288–301. doi:10.1111/1462-2920.13026.
- Žifčáková, L., Větrovský, T., Lombard, V., Henrissat, B., Howe, A., and Baldrian, P. (2017). Feed in summer, rest in winter: microbial carbon utilization in forest topsoil. *Microbiome* 5, 122. doi:10.1186/s40168-017-0340-0.
- Zwetsloot, M. J., Kessler, A., and Bauerle, T. L. (2018). Phenolic root exudate and tissue compounds vary widely among temperate forest tree species and have contrasting effects on soil microbial respiration. *New Phytol.* doi:10.1111/nph.15041.

# Acknowledgements

---

*“Utazi ubwenge ashima ubwe” – Rwandese proverb*  
One who knows no other wisdom always praises one’s own

*“Ubwenge buraraburwa” – Rwandese proverb*  
Intelligence, like fire, is obtained from one’s neighbor

The lines above are among the most often repeated proverbs in Rwanda. One cautions against complacency in one’s knowledge and one encourages to counter this by seeking out help and wisdom from others, and points to the importance of collective efforts by many in improving one’s self. Considering my journey, they have been the most important lessons I’ve learned during this project, and I’d like to thank all those who contributed to my education.

My first thanks and sincere gratitude to Prof. Dr. Rolf Daniel for accepting me in his department and opening the world of genomics with a great project, as well as the tools and environment to explore it. And by the gods, the patience, too!

I’d like to thank PD. Dr. Michael Hoppert as my co-advisor in the project, his invaluable input in my thesis committees and generally keeping track of my progress throughout the project.

My thanks also go to PD. Dr. Fabian Commichau, PD. Dr. Kai Heimel, PD. DR. Ina Meier, and Prof. Dr. Stefanie Pöggeler for readily accepting to be part of the examination committee.

I’d like to thank Dominik Schneider for helping me to wade through the murky waters of data analysis; without him, this thesis would be an Excel-bar graph fest. Apart from that, he was a great mentor, co-supervisor and just all-round top friend that I have come to value. Proud to be a Dominikan!

Special thanks to Robert Hertel for reading through my thesis, and Jacqueline Hollensteiner for her help in clearing administrative issues. Most importantly, though, they’ve been a great a balm to frayed nerves. They come highly recommended!

Many thanks to Silja Brady, Sarah Herzog, Heiko Nacke, Birgit Pfeiffer, Stefanie Rockstroh, Carola Simon, Bernd Wemheuer and Christiane Will. They were my first introduction to the department, and Germany; from registering at the Rathaus to working in the lab... Thank you for the warm welcome, indeed.

To my lab mates, past and present, particularly Mingji Lu (my PhD sister!), Vera Korolkow, Genis Castillo Villamizar, Soumya Biswas, Rahadian Pratama, Richard Egelkamp, Martin Engelhaupt, Kristin Kaiser, Dirk Berkelman, Avril von Hoyningen-Huene, Simone Hanekamp, Ines Friedrich, Tobias Schilling, and Peggy Klempert. And to my very own students Marcel Wittwer, Jörn Lindeman and Florian Jung who lightened the load. Oh, so many stellar moments! You’re my second family, I really couldn’t imagine a better group. And much like family, I know most of them are only reading this section, while others have made a valiant stab at the Introduction...

To Petra Ehrenreich, Daniela Dreykluft and Nicole Dörjer, thank you for your assistance. The same for Mechthild Bömeke and Melanie Heinemann for their technical support in the lab, the Werkstatt staff and Christiane Wohlfeil. These people are the glue that hold the workplace together. Thank you!

I extend my thanks to members of Grako 1086, particularly Marcus Schmidt, who set up the sampling site and acted as forest warden to a lab rat! On that note, I'd like to add Heiko Nacke and Kristin Kaiser who were my unpaid chauffeurs. Thanks again to Richard Egelkamp, Martin Engelhaupt, Marcel Wittwer and Carl who volunteered to help with sampling and made quick work of it!

Then my sincerest gratitude to having the following friends: Cumbiera Intellectuales, Ladies Who Lunch, Tanya 'Brunchelina' Turipamwe Stroh, the Mexican Trio, the Crazy Sisterhood, the Kovsie 5, the 2 Rhodents, Sobremensa, Andres Agudelo-Toro and Miriam Ibenthal, Stephen Asabere (*It is known!*), Michael Ainoosson, Hellen Adeboye, Daisy Caraté, Rebecca Critchley, Agathe Wehrli, Monika Hammerstein and Ulrike Mälzig. With your constant support and encouragement, it was absolutely impossible to move anywhere but forward.

My biggest thanks go to The Clan, my family: Gilbert Habimana, Gemma Akayezu, Pierre Celestin Barayagwiza, Rémy Ngamije, Ange Mucyo and Brice Mihigo. They have been my constant cheerleaders at each milestone and, in parts, advisors, financiers, entertainers, anchors...etc. Fam, this is literally my Saigo no Getsuga Tenshō. Thank you!

Thank you all for bringing me to the finish line,

Amélie Dukunde

This thesis is dedicated to my father,  
who took me to school for the first time.  
And then, unexpectedly, I never wanted to leave!  
Ha ha ha!

# Appendices

---

## List of non-standard abbreviations

|           |   |
|-----------|---|
| ANOVA     | Analysis of variance                                    |
| ANOSIM    | Analysis of similarities                                |
| PERMANOVA | Permutational multivariate analysis of variance         |
| OTU       | Operational taxonomic unit                              |
| NMDS      | Non-metric multidimensional scaling                     |
| cDNA/eDNA | Complementary DNA/environmental DNA                     |
| RBS       | Ribosomal binding site                                  |
| CAZymes   | Carbohydrate active enzymes                             |
| ORF       | Open reading frame                                      |
| ABTS      | 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) |

## Nucleobases

|   |          |
|---|----------|
| A | Adenine  |
| C | Cytosine |
| G | Guanine  |
| T | Thymine  |

## Amino acids

|         |               |         |                |
|---------|---------------|---------|----------------|
| Ala (A) | Alanine       | Asp (D) | Aspartic acid  |
| Asn (N) | Asparagine    | Gln (Q) | Glutamine      |
| Cys (C) | Cysteine      | Gly (G) | Glycine        |
| Glu (E) | Glutamic acid | Ile (I) | Isoleucine     |
| His (H) | Histidine     | Lys (K) | Lysine         |
| Leu (L) | Leucine       | Phe (F) | Phenylalanine  |
| Met (M) | Methionine    | Ser (S) | Serine         |
| Pro (P) | Proline       | Trp (W) | Tryptophan     |
| Thr (T) | Threonine     | Val (V) | Valine         |
| Tyr (Y) | Tyrosine      | X       | Any amino acid |
| Arg (R) | Arginine      |         |                |



### **Academic curriculum vitae**

For the electronic publication of this thesis, the CV has been omitted for security purposes.

## Thesis Declaration

---

### Declaration of plagiarism

I hereby confirm that I have written this doctoral thesis independently. I have not used other sources or facilities others than the ones mentioned in the chapters. Moreover, I have not used unauthorized assistance and have not submitted this thesis previously in any form for another degree at any institution or university.

Amélie Dukunde

Göttingen, May 2018