Faculty of Science

Department of Biology

Unravelling the role of soil properties as predictors of local- and global-scale grassland productivity and soil microbial community composition patterns

Het ontrafelen van de rol van bodemfactoren als voorspellers van graslandproductiviteit op lokale en mondiale schaal en van de samenstelling van de microbiële gemeenschap in de bodem

Thesis submitted for the degree of Doctor of Science: Biology at the University of Antwerp to be defended by

Dajana Radujković

Supervisors: Prof. Dr. Sara Vicca & Prof. Dr. Erik Verbruggen

Antwerp, 2020

This research has been financially supported by the Research Foundation – Flanders (FWO): grant number 1505318N, by the European Research Council (ERC): grant number ERC-2013-SyG-610028-IMBALANCE-P and by Methusalem funding of the Research Council UA.

Table of contents

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisors Erik Verbruggen and Sara Vicca for their constant support and help throughout my doctoral studies and their time and effort invested in every project we worked on. Their enthusiasm for scientific work was a source of motivation for me and this thesis is largely a result of their creative ideas and our mutual conversations.

I am very grateful to Professor Ivan Janssens for giving me the opportunity to do this PhD. His kind words and interest in my work were always very appreciated. I also appreciate the help of James Weedon who was the first to introduce me to the exciting world of soil microbial ecology and thanks to him I was able to continue exploring it.

This work was only possible because of the help and effort of many collaborators from Herbaceous Diversity Network, Noordenveld experiment and Nutrient Network who were collecting and sending soil samples and providing data. Especially, I thank Peter Wilfahrt and Anke Jentsch for hosting me and organising the intensive sampling in Germany and Austria as well as Gretel van Rooyen and Leslie Brown for hosting me and organising the intensive sampling in South Africa. Moreover, I thank Tom van der Spiet for conducting nutrient analyses, Annelies Devos for her help with the microbial work, Kevin Van Sundert for his help during the initial stages of nutrient analyses, Matteo Campioli for his help with the databases and Erik Fransen for statistical advices. I am also grateful to the members of my PhD jury whose constructive comments and suggestions helped me to improve this thesis.

I felt very lucky to be a part of a great group such as PLECO and these past four years were particularly special thanks to all the people working with me: Silvia, Maral, Romain, Inge, Coline, Kevin, Luke, Arne, Cindy, Lingjuan and the entire PLECO group. Moreover, I would like to thank Professor Reinhart Ceulemans for his warm welcome to the group and Nadine and Laura for their big help with all the practicalities and particularly for sending numerous packages to different countries around the world.

A special thanks goes to my office mate Johan, who taught me a great deal about lab work and helped me whenever I had a problem but also for our inspiring conversations in the office. I am very grateful to Irene, my office mate, my housemate and a great friend with whom I shared so much throughout these four years and who always believed in me and helped me remember where I was when I would get confused. I thank Jonas for all the walks next to the lake and the conversations in the hallway that led to very exciting projects. Irene, Jonas and Miguel, our lunch-time conversations and Supplies activities will remain some of the nicest memories from this period.

I also thank Tim, my first friend in Antwerp, for the revision of the Dutch translation of the summary. I am grateful to Sassie for supporting me during the difficult times of my PhD with a smile on her face that always made me feel more positive. My whole experience in Antwerp and everything that went with it might not have happened if it hadn't been for the determination of my friend Perić, who was there to help at some crucial moments.

Miguel, thank you for your help with soil analyses, thesis formatting and conceptual figures, but much more than that, I am grateful for your presence and your invaluable support when it was needed the most…

Finally, I thank my family for all their love and unconditional support. Volim vas!

List of publications

Radujković D, Diggelen R, Bobbink R, Weijters M, Harris J, Pawlett M, Vicca S, Verbruggen E. 2020. Initial soil community drives heathland fungal community trajectory over multiple years through altered plant–soil interactions. *New Phytologist* 225: 2140–2151. DOI: 10.1111/nph.16226

Radujković D, Verbruggen E, Sigurdsson BD, Leblans NIW, Janssens IA, Vicca S, Weedon JT. 2018. Prolonged exposure does not increase soil microbial community compositional response to warming along geothermal gradients. *FEMS Microbiology Ecology* 94. DOI: 10.1093/femsec/fix174

De Jonge C, **Radujković D**, Sigurdsson BD, Weedon JT, Janssens I, Peterse F. 2019. Lipid biomarker temperature proxy responds to abrupt shift in the bacterial community composition in geothermally heated soils. *Organic Geochemistry* 137: 103897. https://doi.org/10.1016/j.orggeochem.2019.07.006

Van Sundert K, **Radujković D**, Cools N, De Vos B, Etzold S, Fernández-Martínez M, Janssens IA, Merilä P, Peñuelas J, Sardans J, et al. 2019. Towards comparable assessment of the soil nutrient status across scales – Review and development of nutrient metrics. *Global Change Biology* 26: 392–409. DOI: 10.1111/gcb.14802

Walker TWN, Janssens IA, Weedon JT, Sigurdsson BD, Richter A, Peñuelas J, Leblans NIW, Bahn M, Bartrons M, De Jonge C, Fuchslueger L, Gargallo-Garriga A, Gunnarsdóttir GE, Marañón-Jiménez S, Oddsdóttir ES, Ostonen I,Poeplau C, Prommer J, **Radujković D**, Sardans J, Sigurðsson P, Soong JL, Vicca S, Wallander H, Ilieva-Makulec K,Verbruggen E. 2020. A systemic overreaction to years versus decades of warming in a subarctic grassland ecosystem. *Nature Ecology and Evolution* 4: 101–108. https://doi.org/10.1038/s41559-019-1055-3

Halbritter AH, De Boeck HJ, Eycott AE, Reinsch S, Robinson DA, Vicca S, Berauer B, Christiansen CT, Estiarte M, Grünzweig JM, et al. 2020. The handbook for standardized field and laboratory measurements in terrestrial climate change experiments and observational studies (ClimEx). *Methods in Ecology and Evolution* 11: 22–37. https://doi.org/10.1111/2041-210X.13331

Lembrechts J, Broeders L, De Gruyter J, **Radujković D**, Ramirez-Rojas I, Lenoir J, Verbruggen E. 2020. A framework to bridge scales in distribution modeling of soil microbiota. *FEMS Microbiology Ecology* 96. https://doi.org/10.1093/femsec/fiaa051

Summary

Plants and soil microorganisms are the main components of every terrestrial ecosystem. They drive the cycle of carbon in nature and they form complex, often species-specific interactions with each other shaping both aboveground and belowground communities. Soil is a medium that connects these two worlds and mediates all interactions between them.

Grasslands are one of the largest terrestrial systems that sustain high levels of biodiversity and play an important role in global carbon sequestration. In this work, two coordinated distributed experiments in grasslands (Nutrient Network – NutNet and Herbaceous Diversity Network – HerbDivNet) were used to examine the role of soil properties as predictors of plant productivity and microbial community composition. The NutNet experiment included 72 worldwide distributed grasslands and HerDivNet included 21 grassland sites around the world, most of which contained a local-scale plant productivity gradient. Besides experiments in grasslands, a large-scale heathland restoration experiment was used to investigate the importance of plant-soil interactions for the development of soil fungal community composition.

Plant biomass production has long been known to depend on the soil they grow on and the climate they live in. While the effects of climate have been widely studied, which of the multitude of individual soil properties and nutrients are the best predictors of global grassland productivity patterns has not been comprehensively assessed, nor has it been studied how much of the total variation they can explain.

In the first part of the thesis, the NutNet and the HerbDivNet experiments were used to conduct two separate studies investigating the importance of a large set of soil properties and nutrients as predictors of global grassland productivity patterns. The results of our studies demonstrated that the most important soil predictors of plant productivity included soil organic matter, soil texture, cation exchange capacity and bulk density. These soil properties generally determine soil nutrients availability and water holding capacity. Regarding particular soil nutrients, Zn emerged as the most important predictor. Moreover, in the study that contained globally replicated gradients in productivity, we found that the same predictors identified at the global scale (soil organic matter, soil texture, bulk density) were consistently good predictors of local-scale variation in grassland productivity across different climates. Overall, both studies revealed that soil properties are better predictors of global grassland productivity than the commonly used climatic factors (soil properties explained 32% of the variation in biomass vs 24% explained by climate in the NutNet study and 45% vs 32%, respectively in the HerbDivNet study).

The second part of the thesis used a heathland restoration experiment and the HerbDivNet experiment in grasslands to investigate the drivers of soil microbial community composition.

The study conducted in the heathland restoration experiment simultaneously investigated the importance of three mechanisms structuring microbial community assembly (timing of colonisation, environmental filters and biotic interactions). The results demonstrated that, when the bare soil is inoculated with heathland fungal and plant communities from the beginning of the system development, they form strong links with each other leading to the convergence of fungal community composition under different environmental conditions. We, therefore, argue that the early stage presence of heathland soil and plant communities and the interactions they form can reinforce the development of a heathland system and alleviate the environmental filter. When the associated heathland soil communities are not present from the beginning, the environment exerts a strong influence on fungal community assembly.

Furthermore, we used a network of globally distributed local-scale gradients in plant productivity from the HerbDivNet experiment in grasslands to examine i) if the same abiotic or biotic factors can predict both global and local-scale patterns in bacterial and fungal community composition, and ii) if community composition differs consistently with local plant productivity (low vs high) across different sites. We found that microbial community composition can be predicted by similar factors on the global and the local scale; with bacteria predominantly associated with soil properties (such as base saturation and pH) and fungi predominantly associated with plant community composition. Moreover, microbial community composition differed consistently at two plant productivity levels across different grassland sites. These findings suggest that there are universal forces that shape microbial community composition across different contexts.

To conclude, this work demonstrates that soil properties are crucial predictors of plant productivity in global grasslands and that soil shapes microbial community composition both directly through providing particular abiotic conditions and indirectly through mediating plant-microbial interactions. Moreover, it is shown here that some of the factors and interactions predicting both grassland plant productivity and microbial community composition are universal across contrasting climates suggesting that similar universality might also hold in other ecosystem types. These findings have implications for understanding and predicting global grassland carbon storage potential, management of grassland biodiversity, understanding current and predicting future local- and global-scale grassland microbial community composition patterns and restoration of heathland ecosystems.

Samenvatting

Planten en bodemmicro-organismen zijn de belangrijkste componenten van elk terrestrisch ecosysteem. Zij sturen de koolstofcyclus in de natuur aan en vormen complexe, vaak soortspecifieke interacties met elkaar die zowel boven- als ondergronds gemeenschappen vormen. De bodem is een medium dat deze twee werelden met elkaar verbindt en alle interacties tussen hen bemiddelt.

Graslanden zijn een van de grootste terrestrische systemen die een hoog niveau van biodiversiteit in stand houden en een belangrijke rol spelen in de wereldwijde koolstofvastlegging. In dit werk zijn twee gecoördineerde gedistribueerde experimenten in graslanden (Nutrient Network - NutNet en Herbaceous Diversity Network - HerbDivNet) gebruikt om de rol van bodemeigenschappen als voorspellers van de productiviteit van planten en de samenstelling van de microbiële gemeenschap te onderzoeken. Het NutNet experiment omvatte 72 wereldwijd verspreide graslanden en HerDivNet omvatte 21 graslandplaatsen over de hele wereld, waarvan de meeste een lokale gradient in plantproductiviteit bevatten.

Naast experimenten in graslanden werd een grootschalig heideherstel-experiment gebruikt om het belang van plant-bodem interacties voor de ontwikkeling van de samenstelling van de bodemschimmelgemeenschap te onderzoeken.

Het is al lang bekend dat de productie van plantaardige biomassa afhankelijk is van de bodem waarop ze groeien en het klimaat waarin ze leven. Hoewel de effecten van het klimaat op grote schaal zijn bestudeerd, is zelden nagegaan welke van de vele individuele bodemeigenschappen en voedingsstoffen de beste voorspellers zijn voor de productiviteitspatronen van het grasland op aarde, en is ook niet onderzocht hoeveel van de totale variatie in productiviteit zij kunnen verklaren.

In het eerste deel van het proefschrift werden de NutNet en de HerbDivNet experimenten gebruikt om twee afzonderlijke studies uit te voeren die het belang van een grote set bodemeigenschappen en voedingsstoffen als voorspellers van wereldwijde graslandproductiviteitspatronen onderzochten. De resultaten toonden aan dat de belangrijkste voorspellers voor de productiviteit van planten onder meer bestaan uit organisch materiaal in de bodem, bodemtextuur, kationenuitwisselingscapaciteit en bulkdichtheid. Deze bodemeigenschappen bepalen over het algemeen de hoeveelheid beschikbare voedingsstoffen in de bodem en de capaciteit om water vast te houden. Wat specifieke voedingstoffen in de bodem betreft, kwam Zn naar voren als de belangrijkste voorspeller. Bovendien vonden we in de studie die globaal gerepliceerde productiviteitsgradiënten bevatte, dat dezelfde voorspellers die op wereldschaal werden geïdentificeerd (organische stof in de bodem, bodemtextuur, bulkdichtheid) consistent goede voorspellers waren voor de variatie in de productiviteit van grasland op lokale schaal in verschillende klimaten. Over het geheel genomen toonden beide

studies aan dat de bodemeigenschappen betere voorspellers zijn voor de wereldwijde productiviteit van grasland dan de algemeen gebruikte klimaatfactoren (de bodemeigenschappen verklaren 32% van de variatie in biomassa vs 24% in de NutNet-studie en 45% vs 32% in de HerbDivNet-studie).

Het tweede deel van het proefschrift maakte gebruik van een heideherstel-experiment en het HergDivNet-experiment in graslanden om de drijvende krachten achter de samenstelling van de microbiële gemeenschap in de bodem te onderzoeken.

De studie die in het heideherstel-experiment werd uitgevoerd, onderzocht tegelijkertijd het belang van drie mechanismen die de samenstelling van de microbiële gemeenschap structureren (timing van kolonisatie, milieufilters en biotische interacties). De resultaten toonden aan dat wanneer de kale bodem vanaf het begin van de systeemontwikkeling wordt geënt met geassocieerde heideschimmel- en plantengemeenschappen, deze sterke banden met elkaar vormen die leiden tot de convergentie van de samenstelling van de schimmelgemeenschap onder verschillende omgevingscondities. We stellen daarom dat de vroege aanwezigheid van heidegronden en plantengemeenschappen en de interacties die zij vormen, de ontwikkeling van een heidesysteem kunnen versterken en het milieufilter kunnen verlichten. Wanneer de bijbehorende bodemgemeenschappen niet vanaf het begin aanwezig zijn, oefent het milieu een sterke invloed uit op de samenstelling van de schimmelgemeenschap.

Bovendien hebben we een netwerk van wereldwijd verspreide lokale gradiënten in plantproductiviteit van het HerbDivNet grasland gebruikt om te onderzoeken i) of dezelfde abiotische of biotische factoren zowel wereldwijde als lokale patronen in bacteriële en schimmelgemeenschapsamenstelling kunnen voorspellen, en ii) of de gemeenschapsamenstelling consistent verschilt met de lokale plantproductiviteit (laag vs. hoog) over verschillende locaties. We vonden dat de samenstelling van de microbiële gemeenschap kan worden voorspeld door vergelijkbare factoren op wereldwijde en lokale schaal. Bacteriën werden voornamelijk geassocieerd met bodemeigenschappen (zoals basisverzadiging en pH) en schimmels werden voornamelijk geassocieerd met de samenstelling van de plantengemeenschap. Bovendien verschilde de samenstelling van de microbiële gemeenschap consistent op twee niveaus van plantenproductiviteit op verschillende graslandplaatsen. Deze bevindingen suggereren dat er universele krachten zijn die de samenstelling van de microbiële gemeenschap in verschillende contexten vormgeven.

Tot slot toont dit werk aan dat bodemeigenschappen cruciale voorspellers zijn voor de productiviteit van planten ingraslanden wereldwijd en dat de bodem de samenstelling van de microbiële gemeenschap vormgeeft, zowel direct door bepaalde abiotische omstandigheden te creëren als indirect door plant-microbiële interacties te bemiddelen. Bovendien wordt hier aangetoond dat sommige factoren en interacties die zowel de productiviteit van graslandplanten als de samenstelling van de microbiële gemeenschap voorspellen, universeel zijn in contrasterende klimaten, wat suggereert dat een soortgelijke universaliteit ook in andere

ecosysteemtypen zou kunnen gelden. Deze bevindingen hebben implicaties voor het begrijpen en voorspellen van het wereldwijde koolstofopslagpotentieel van graslanden, het beheer van de biodiversiteit van graslanden, het begrijpen van de huidige en het voorspellen van toekomstige patronen van de samenstelling van de microbiële gemeenschap van graslanden op lokale en wereldwijde schaal en het herstel van heide-ecosystemen.

General introduction

CHAPTER I

CHAPTER I

General introduction

Plants and soil microorganisms are at the core of the functioning of every terrestrial ecosystem. While plants are the main producers of organic matter, soil microorganisms are the main decomposers that transform organic material back to minerals necessary for plant growth. Beyond this, plant and microbes are inseparably intertwined forming complex interactions, ranging from mutualism to parasitism and competition, that shape both above and below-ground communities. Soil is a medium that connects these two worlds and mediates all interactions between them.

In this work, we use coordinated distributed experiments in grasslands worldwide to examine the role of soil properties as predictors of plant productivity and microbial community composition. We focus on grasslands as one of the largest terrestrial systems occupying up to 40% of the land (Blair *et al.*, 2014) which sustain large levels of biodiversity (Nerlekar & Veldman, 2020). Given that they account for up to one-third of the net primary production on land (Vitousek, 2015) they also play an important role in global carbon sequestration (Squires *et al.*, 2018).

Besides grasslands, we performed a study in a developing heathland (as a relatively simple model system) to investigate the importance of plant-soil interactions for co-development of soil fungal communities and their associated plant communities.

1.1 Part I: Grassland productivity

1.1.1 Patterns and drivers of aboveground grasslands productivity – an overview

Plant productivity is a key characteristic of grassland ecosystems. It determines their capacity to take up and store carbon (White *et al.*, 2000), controls which plant species can (co-)exist in plant assemblages and shape their diversity (Fraser *et al.*, 2015) and likewise influences the diversity of numerous animals that depend on plants for food and habitat (Huston, 1995; Squires *et al.*, 2018) as well as the diversity and community composition of soil microorganisms (Zak *et al.*, 2003; Chen *et al.*, 2015; Waldrop *et al.*, 2017). Understanding grassland productivity patterns has, thus, long been one of the central topics in ecological

research (e.g. Rosenzweig, 1968; Sala *et al.*, 1988, 2012, Tilman *et al.*, 1997, 2001, 2009; Grace *et al.*, 2007; LeBauer & Treseder, 2008; Oehri *et al.*, 2017).

Climate is thought to be an overarching factor that determines the broad-scale patterns of productivity, not only for grasslands but also for other ecosystems (Rosenzweig, 1968; Huxman *et al.*, 2004; Sala *et al.*, 2012; Hovenden *et al.*, 2014, 2019). For instance, different terrestrial biomes on Earth which are intrinsically related to the amount of plant biomass they can produce (which increases from deserts and tundra to grasslands, temperate forests and finally tropical forests) are largely a result of climatic conditions in which these biomes occur (Archibold, 1995) (Figure 1.1). Large-scale productivity patterns in grasslands are found to be particularly strongly related to precipitation patterns; hence, water availability is thought to be one of the main limiting factors for plant growth in grasslands (Sala *et al.*, 1988).

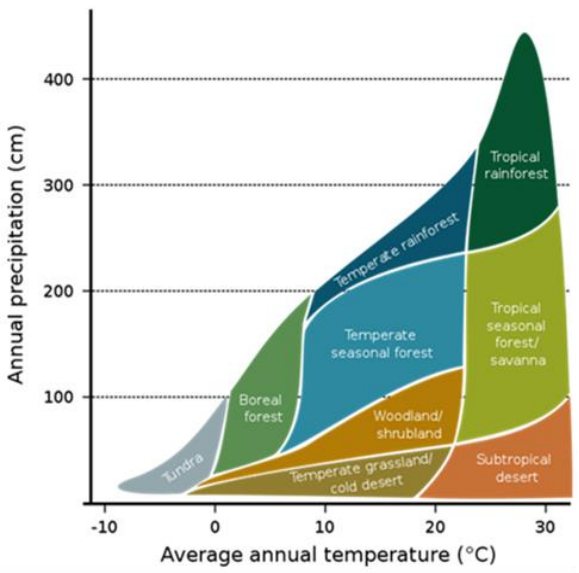


Figure 1.1 Precipitation and temperature determine the distribution of biomes on Earth. Credit: "Climate influence on terrestrial biome" by Navarras is in the Public Domain.

However, besides climate, water availability is also determined by **soil properties**; i.e. soils with low capacity to retain water typically have low fertility and productivity (Bünemann *et al.*, 2018), while soils with higher water holding capacity can buffer the effects of climate and produce more biomass even in the regions with lower levels of precipitation (Wang *et al.*, 2009). Moreover, soil properties also determine soil fertility and the amount of nutrients available for plant growth. The importance of soil fertility has been recognised by farmers for centuries; e.g. large amounts of fertile anthropogenic soil was found in the Amazonian basin dating back to the 5th century BC (Lehmann *et al.*, 2003). The role of nutrients for soil fertility and the concept of nutrient limitations for plant growth were later described and universally acknowledged by scientists (Von Liebig, 1840; Chapin, 1980) and extensively studied in agricultural contexts. In the context of grasslands as (semi)natural systems, numerous fertilization experiments demonstrated that grassland productivity is often limited by two main macronutrients (nitrogen and phosphorus) (Elser *et al.*, 2007; Craine & Jackson, 2010; Harpole *et al.*, 2011; Ågren *et al.*, 2012) while other experiments have demonstrated that multiple different nutrients can also co-limit grassland productivity (Cech *et al.*, 2008; Fay *et al.*, 2015; Lannes *et al.*, 2016). Related to this is the effect of anthropogenic atmospheric nutrient

deposition that can act as (primarily nitrogen) fertilizer thereby alleviating nutrient limitations and potentially increasing grassland productivity (Phoenix *et al.*, 2012; Stevens *et al.*, 2015).

Finally, **biotic factors** also play a role in determining plant productivity. For instance, plant diversity could enhance community productivity through niche complementarity, i.e. the higher the plant species diversity, the better the capacity of the community to use different available resources which promotes overall productivity (Hector *et al.*, 1999; Craven *et al.*, 2016). Other biotic factors, such as herbivory, parasitism and mutualism can also exert an important control on grassland biomass production (van der Heijden *et al.*, 1998; Frank *et al.*, 2002; Maron *et al.*, 2011) either by increasing plants' capacity to survive (e.g. by providing them limiting nutrients as in the case of mycorrhizal fungi) or by decreasing their fitness by causing diseases and damages to tissues (in the case of plant pathogens and herbivores).

1.1.2 Predicting grassland productivity patterns – state of the art, challenges and the way forward

Climatic factors have almost exclusively been used as predictors of global grassland productivity (e.g. Fay *et al.*, 2003; Huxman *et al.*, 2004; Sala *et al.*, 2012) given that they correlate very well with coarse productivity patterns and they are relatively easy to measure and compare across different sites. With the development of global models of N deposition, this factor has also been used to explain global grassland productivity (Stevens *et al.*, 2015). However, using only broad-scale predictors such as climate and nitrogen deposition leaves a high amount of unexplained variation which, for example, can lead to increased uncertainty in the models aiming to predict plant productivity under changing climate (Folberth *et al.*, 2016).

Even though the importance of soil fertility for plant growth is undisputed, we know surprisingly little about whether and which soil physicochemical properties and nutrients determining soil fertility can contribute to predicting the variation in grassland productivity on a global scale. Thus far, few global-scale studies included any of the soil properties as predictors of grassland productivity (e.g. Stevens et al. (2015)), however, to the best of our knowledge, no study investigated a comprehensive set of soil predictors. Given that different soil properties interact to determine soil fertility (Larcher, 2003; Kirkham, 2005), it is necessary to simultaneously assess the importance of several key soil properties (including soil texture, soil organic matter and pH) as predictors of global grassland productivity. Moreover, it remains unclear to what extent the in-situ availability of different soil nutrients (other than nitrogen, phosphorus and potassium) and particularly micronutrients, can contribute to explaining global grassland productivity patterns.

One major issue impeding the analyses of soil properties and nutrients as predictors of grassland productivity is a lack of comprehensive soil datasets. This is due to the fact that

important soil properties are not consistently measured and/or reported across studies; hence the number of predictors that can be used in meta-analyses is limited (Vicca *et al.*, 2018). Moreover, different methodologies applied to measure biomass production and soil properties across studies may be difficult to compare, further complicating the syntheses studies. One emerging solution to this issue is the establishment of coordinated distributed experiments where standardized, controlled protocols are used to perform measurements at each site which allows for effective comparison of data across different sites (Fraser *et al.*, 2013). These globally distributed experiments thus provide an excellent opportunity to elucidate the role of soil properties and nutrients as global-scale predictors of grassland productivity.

1.2 Part II: Soil microbial community composition

1.2.1 The importance of understanding microbial community composition

Researchers have long been trying to explain the patterns of aboveground diversity. Until recently, similar studies were impossible for soil microorganisms, especially at larger spatial scales. This is due to the extreme complexity of soil microbial (primarily bacterial and fungal) communities with possibly thousands of different species in just one gram of soil (Daniel, 2005). For microbial biogeography, the traditional view has held that "Everything is everywhere, but the environment selects" (Baas Becking, 1934). This long-standing paradigm refers to the remarkable potential for dispersal of microorganisms (Fuhrman, 2009) suggesting that microorganisms can be found everywhere and environmental conditions entirely determine which species can thrive and reach high abundances. This means that if any species disappears from the system due to environmental changes, its role will be easily replaced by another species (because of a high degree of functional redundancy) leaving no substantial effect on the overall functioning of the system. Therefore, although it has been acknowledged that microbes are crucial players in the processes that sustain the life on Earth, including nutrient cycling, decomposition and soil formation (Prosser *et al.*, 2007; Falkowski *et al.*, 2008), the importance of soil microbial community composition has often been marginalised.

However, recent research has disputed the theory of unlimited dispersal of soil microbes (Hanson *et al.*, 2012) as numerous continental-scale studies have shown that soil microbial community composition exhibits predictable biogeographical patterns (Fierer & Jackson, 2006; Drenovsky *et al.*, 2010; Nemergut *et al.*, 2011; Tedersoo *et al.*, 2014). These patterns are less clear for bacteria than fungi given that some common bacterial taxa are globally distributed (Delgado-Baquerizo *et al.*, 2018a), however, dispersal limitation was found to contribute to bacterial community composition (Albright & Martiny, 2018) even at local scales (Martiny *et*

al., 2011). Likewise, the theory of strong functional redundancy in microbial communities was challenged (Reed & Martiny, 2007; Allison & Martiny, 2008) and numerous studies experimentally demonstrated that changes in community composition can alter the ecosystem functioning (Strickland *et al.*, 2009; Reed & Martiny, 2013; Wagg *et al.*, 2014). Moreover, many soil microorganisms were shown to form species-specific interactions with plants (van der Heijden *et al.*, 2008; Bever *et al.*, 2015) and they thus perform unique functions in the ecosystem.

In light of these findings, it has become increasingly evident that understanding how soil microbial communities are shaped is necessary to better understand other components of the ecosystem and to improve the accuracy of global-change models. With the recent improvement of technologies that allow community profiling such as phospholipid fatty acid (PLFA) analysis and high-throughput DNA sequencing, there has been a growing amount of scientific literature examining and describing the patterns and drivers of soil microbial diversity and community composition (e.g. Fierer & Jackson, 2006; Fierer *et al.*, 2009; Martiny *et al.*, 2011; de Vries *et al.*, 2012; Tedersoo *et al.*, 2014; Thompson *et al.*, 2017; Waldrop *et al.*, 2017; Ramirez *et al.*, 2018; Delgado-Baquerizo *et al.*, 2018; Rasmussen *et al.*, 2018; Chalmandrier *et al.*, 2019). Moreover, large-scale initiatives such as the Earth Microbiome Project (Gilbert *et al.*, 2014), the Global Soil Biodiversity Assessment Initiative and TerraGenome (Vogel *et al.*, 2009) were established with the aim to construct a global catalogue of the microbial diversity on Earth and eventually elucidate relationships between diversity, community composition and ecosystem functions.

1.2.2 Drivers of soil microbial community composition – state of the art

Community assembly of species contained in an ecosystem is thought to depend on three main processes: i) environmental filters, ii) biotic interactions and iii) dispersal constraints and historical contingencies (Belyea & Lancaster 1999; Lortie *et al.*, 2004).

In the case of soil microbes, the environmental filter has been the main research focus, and on a global level, microbial diversity and community composition were found to be primarily influenced by soil abiotic factors (Fierer & Jackson, 2006; Tedersoo *et al.*, 2014) and climate (Tedersoo *et al.*, 2014; Chen *et al.*, 2015). For instance, it is commonly acknowledged that soil pH is one of the most important drivers of soil bacterial community composition (Fierer & Jackson, 2006; Männistö *et al.*, 2007; Lauber *et al.*, 2009) which could be a result of the relatively narrow pH ranges for optimal growth of individual bacterial taxa (Rousk *et al.*, 2010). Plant productivity was also found to affect microbial diversity and community composition (Zak *et al.*, 2014; Chen *et al.*, 2015; Waldrop *et al.*, 2017; Delgado-Baquerizo *et al.*, 2018a), likely because it determines the quantity and quality of labile carbon inputs (Waldrop *et al.*, 2017).

Moreover, numerous studies have shown that plant diversity and community composition as well as plant community functional traits can influence soil microbial community composition (Chung *et al.*, 2007; Orwin *et al.*, 2010; Millard & Singh, 2010; Eisenhauer *et al.*, 2011; de Vries *et al.*, 2012; Zak *et al.*, 2014; Prober *et al.*, 2015; Sayer *et al.*, 2017). Plant communities can affect soil microorganisms by increasing the diversity of environmental conditions, root exudates and litter, and by providing a diverse set of hosts for mutualistic and antagonistic microorganisms (Prober *et al.*, 2015). This effect is likely to be particularly strong for microorganisms that engage in intricate, species-specific interactions with plants, such as symbiotic N fixing bacteria (van der Heijden *et al.*, 2006), fungal plant pathogens (Bever *et al.*, 2015) and mycorrhizal fungi. Mycorrhizal fungi, for instance, form associations with up to 90% of terrestrial plants helping them to access limiting nutrients while receiving organic carbon in return (Smith & Read, 2008). Therefore, direct and indirect plant-microbe interactions may play a central role in both plant and microbial community assembly processes (Wubs *et al.*, 2019).

Finally, dispersal limitation determines the order and timing of species immigration during microbial community assembly (Fukami, 2015). Incoming species and the interactions they form can cause historical contingency (also called priority effects) that can strongly shape the development of the system (Dickie *et al.*, 2012; Vannette *et al.*, 2014). For instance, Wubs et al. (2019) demonstrated that single introductions of soil and plant material generated long-term legacies in microbial community assembly in two contrasting systems.

1.2.3 Emerging questions about soil microbial community assembly

To explain local community composition, ecologists have typically been using a hierarchical structure where dispersal limitation determines a larger pool of species that can reach a certain habitat. This pool of species is then 'filtered' by the environment and the final pool is determined by interactions between the remaining species (Figure 1.2). The current use of the environmental filtering concept originates from the research on plant community assembly and dynamics (Nobel & Slatyer, 1977; van der Valk, 1981; Woodward & Diament, 1991) which described the environment as a metaphorical 'sieve' that allows species with particular traits to establish and persist under certain conditions, excluding all others. This paradigm of hierarchical filters has recently been criticized in several studies (Anderson *et al.*, 2011; Cadotte & Tucker, 2017; Aguilar-Trigueros *et al.*, 2017); more specifically, the importance of the environment as a primary "filter" is thought to be overestimated (Kraft *et al.*, 2015) and it was proposed that historical legacies, and biotic interactions can determine the strength of the environmental filter (Cadotte & Tucker, 2017).

While the individual predictors of microbial communities are well studied, it is currently poorly understood how different mechanisms – environmental filtering, biotic interactions and geographic barriers and other historical contingencies that limit dispersal – interact to shape microbial community assembly.

Furthermore, even if the processes that shape microbial community assembly would be disentangled, the question remains about whether the mechanisms identified under certain environmental conditions would be valid under entirely different contexts, e.g. across contrasting climates and soil types. Several studies showed that that the predictors of microbial community composition are dependent on spatial scale and/or environmental conditions (Hendershot *et al.*, 2017; Alzarhani *et al.*, 2019; Chalmandrier *et al.*, 2019) confirming the notion that generality in community ecology is exceptionally rare (Lawton, 1999). This calls to question whether any general predictors can be found for soil microbial communities at different contexts over large spatial ranges.

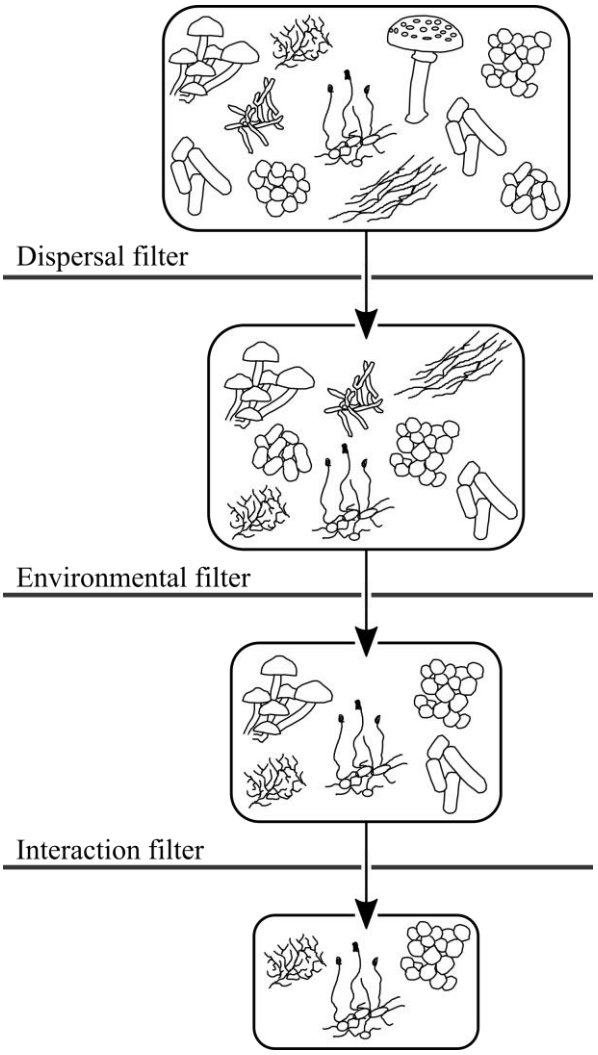


Figure 1.2 The traditional concept of community assembly processes according to which a series of hierarchical "filters" determine community assembly. Modified from Cadotte & Tucker (2017). Drawing credit: Miguel Portillo-Estrada.

1.3 Aims and objectives

1.3.1 Part I: The role of soil properties as predictors of grassland productivity

In the first part of the thesis, we used two global-scale experiments in grasslands; the Nutrient Network – NutNet (Borer *et al.*, 2014) and the Herbaceous Diversity Network – HerbDivNet (Fraser *et al.*, 2015) to unravel the role of soil properties as predictors of grassland productivity. Specifically, we focussed on the following two research questions (Figure 1.3):

Question 1 (Chapter II)

- Which soil properties determining nutrient availability can best predict global-scale variation in aboveground grassland productivity and what is their contribution compared to broad-scale drivers such as climate and N deposition?

To answer this question, we used the NutNet datasets including data on 72 worldwide distributed grassland sites where the concentrations of 12 different soil nutrients were measured in a standardized manner.

Question 2 (Chapter III)

- Which soil physicochemical properties can best predict local-scale variation in aboveground plant productivity within different grassland sites and how much they can contribute to explaining the global-scale variation in grassland productivity in addition to climate?

Here, we used the HerbDivNet experiment which consists of 21 globally distributed grassland sites many of which contain a local-scale gradient in productivity (low, medium, high productivity); making it ideal for examining and comparing local- and global-scale predictors of grassland productivity.

1.3.2 Part II: Soil microbial community assembly – mechanisms and predictors

In the second part of the thesis, we used an experiment in developing heathlands and the HerbDivNet experiment in grasslands to explore the predictors and mechanisms driving microbial community assembly (Figure 1.3):

Question 3 (Chapter IV)

- What are the most commonly used methods of soil microbial community assembly analyses?

Question 4 (Chapter V)

- What is the relative importance and interaction between ecological filters (timing of colonization/dispersal, abiotic conditions and biotic interactions) for the development of soil fungal community composition over multiple years?

For this study, we made use of a multi-year, large scale heathland restoration experiment that included soil and plant inoculation treatments (creating historical legacies) crossed with pH manipulation treatments (creating environmental filters) to examine the development of plant-fungal interactions (biotic filters).

Question 5 (Chapter VI)

- Is there generality in predictors of microbial (bacterial and fungal) community composition in grasslands across two different spatial scales (local and global)? Moreover, we investigate if plant productivity affects microbial community composition in a consistent manner across different grassland sites.

To accomplish this, we used the HerbDivNet grassland experiment with globally replicated local-scale productivity gradients which allowed us to investigate and compare the predictors of microbial community composition at different spatial and ecological contexts.

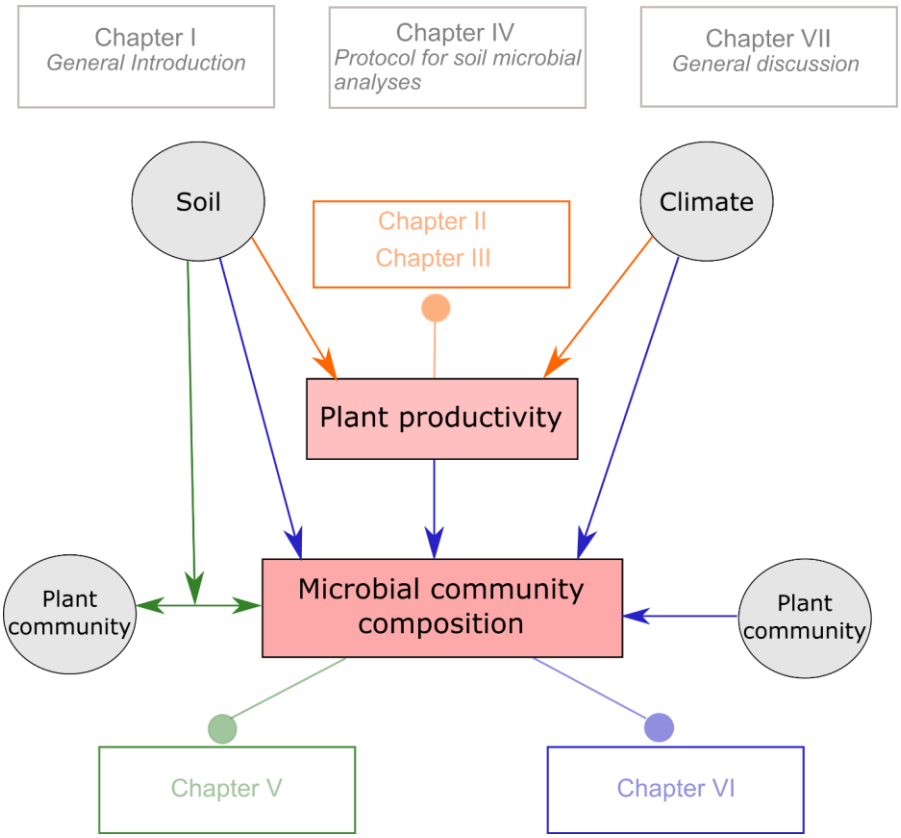


Figure 1.3 Schematic representation of the questions examined in the thesis and overview of the topics covered in each chapter. The colours of lines and boxes correspond to the topics examined in different chapters.

Soil nutrient availability as a predictor of global aboveground grassland productivity

CHAPTER II

CHAPTER II

Soil nutrient availability as a predictor of global aboveground grassland productivity

2.1 Abstract

Plant productivity is a key determinant of biodiversity and carbon sequestration in grasslands. Previous research analysed the role of climate and atmospheric N deposition as predictors of global grassland biomass production but the possible contribution of various soil properties determining nutrient availability has not been comprehensively examined, despite their unquestionable importance for plant productivity. Using data on climate, N deposition, soil organic matter (SOM), pH, cation exchange capacity (CEC) and concentrations of different soil nutrients in 72 worldwide distributed grasslands, we investigated whether and which soil properties contribute to explaining the variation in global grassland aboveground biomass (AGB). Our results demonstrate that besides soil properties associated with soil fertility (SOM and CEC), soil zinc (Zn) concentrations predicted additional amount of variation in AGB (with a possible indirect influence of SOM through Zn). Soil properties together explained 32% of the unique (non-shared) variation, while climate (precipitation) and N deposition uniquely explained 16% and 4%, respectively. Moreover, the relationship between soil micronutrient Zn and AGB was absent in the sites with relatively low nitrogen (N) or phosphorus (P) availability, as well as in the subset of grasslands previously shown to be limited/co-limited by N/NP. These results suggest that in areas where N and P are not strongly (co)limiting plant growth, soil Zn availability might have an important influence on global grassland productivity.

2.2 Introduction

Climatic factors, particularly precipitation, have long been recognized as major determinants of grassland aboveground productivity at a global scale (e.g. Sala *et al.* 1988; Huxman *et al.*, 2004). The important role of soil nutrients in determining biomass production patterns has likewise long been acknowledged (Chapin, 1980) and extensively studied in agricultural contexts. In the context of grasslands as (semi)natural systems, numerous fertilization experiments demonstrated that grassland productivity is often limited by two main macronutrients (nitrogen, phosphorus) and sometimes also potassium (Olde Venterink *et al.*, 2001a; Elser *et al.*, 2007; Craine & Jackson, 2010; Harpole *et al.*, 2011; Ågren *et al.*, 2012). In line with this, it has been shown that modelled anthropogenic N deposition patterns can predict 16% of the variation in global grassland productivity patterns (Stevens *et al.*, 2015). Besides

N, P and K, other nutrients (including calcium, magnesium, sulphur, iron, boron, copper, manganese, zinc) were found to co-limit grassland productivity, either added jointly (Cech *et al.*, 2008; Fay *et al.*, 2015; Lannes *et al.*, 2016) or alone, e.g. in the case of the micronutrient boron (Lannes *et al.*, 2020).

Even though certain nutrients are needed in much smaller concentrations than the others (hence "micro" nutrients), their influence on plant growth is as important as that of macronutrients given that they are constituents of prosthetic groups that catalyse redox processes, form enzyme-substrate complexes, enhance enzyme reactions or play a role in protein synthesis (Fageria *et al.*, 2002; Broadley *et al.*, 2011). Therefore, the deficiency in any essential nutrient limits plant development and productivity. While various micronutrients have been shown to influence crop yield and limit the productivity of agricultural plants (e.g. Fageria *et al.*, 2002; Dimkpa & Bindraban, 2016), they have rarely been investigated in (semi)natural systems. However, the study by Fay et al. (2015) has put micronutrients into the spotlight as potentially important limiting factors of plant productivity in non-agricultural grasslands worldwide.

Although the role of soil nutrients in limiting grassland productivity is undisputed, a general assessment of the importance of different soil factors determining natural variation in nutrient availability as predictors of grassland productivity is currently lacking. This is particularly true for soil micronutrient availability which is not often (consistently) measured or reported in studies quantifying grassland productivity and to the best of our knowledge, it has never been used to predict productivity patterns across large spatial ranges. Besides the concentrations of different macro- and micro-nutrients, nutrient availability is also determined by soil physicochemical properties such as pH, the amount of organic matter and soil cation exchange capacity (Vicca *et al.*, 2018; Van Sundert *et al.*, 2020). For instance, soils with low CEC have limited capacity to retain cations which are therefore easily leached resulting in nutrient imbalances and reduced nutrient availability (Lehmann & Schroth, 2005). The question thus remains which soil properties determining nutrient availability can contribute to predicting global aboveground grassland productivity in addition to broad-scale drivers such as climate and atmospheric N deposition.

Here, we made use of the comprehensive and harmonized soil dataset from NutNet – a globally distributed network of (semi)natural grasslands (Borer *et al.*, 2014) – to examine the relationship between the natural variation in soil properties and nutrient concentrations and the variation of global grassland biomass production. The data on biomass production, measured in a consistent manner in 72 sites around the globe (Figure 2.1), was collected along with the data on the concentrations of 12 different soil nutrients, soil pH, SOM and CEC, atmospheric N deposition and climatic conditions. This dataset thus contained information about a wide set of soil nutrients (including different macro- and micro-nutrients) across globally distributed grassland sites with contrasting climatic conditions. We hypothesised that, besides climate, soil physicochemical properties (pH, SOM, CEC) would show the strongest associations with biomass production due to their important effect on overall soil fertility. Moreover, given the experimental proof of widespread N limitation in terrestrial systems including grasslands (LeBauer & Treseder, 2008; Fay *et al.*, 2015) as well as co-limitation by P (Elser *et al.*, 2007; Harpole *et al.*, 2011) and other nutrients (Fay *et al.*, 2015), we expected that factors influencing N availability (C:N and atmospheric N deposition) and P availability would predict additional variation in AGB along with other macro- and/or micro-nutrients. The latter would particularly be expected in the regions where limitation by N/NP is less pronounced.

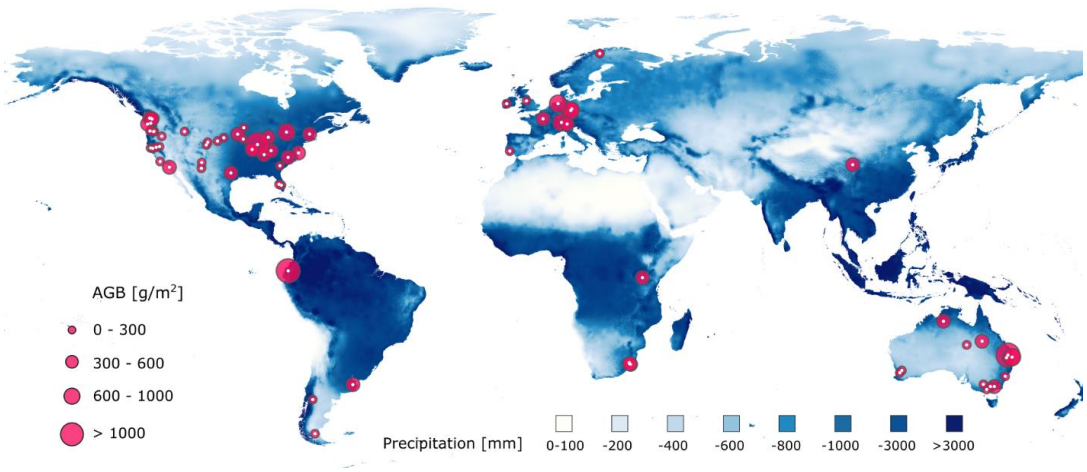


Figure 2.1 The distribution of 72 NutNet grassland sites along the precipitation gradient. White points indicate the location of different sites and different sizes of points around them correspond to the amount of aboveground biomass (AGB) per site.

2.3 Materials and methods

2.3.1 Data collection

Plant aboveground biomass data was collected from 72 Nutrient Network (NutNet) grassland experimental sites (www.nutnet.org) (Table S2.1). Sites were located on six continents and spanned a wide range of peak biomass (58 – 1602 g/m^2), mean annual precipitation (211 – 2813 mm) and mean annual temperature (-2.7 – 27.8°C) values (Figure 2.1). At each site, standing crop (live biomass and recently senescent material) was estimated destructively by clipping aboveground vegetation at the peak of the growing season from two 0.1 m^2 (10 x 100 cm) strips for a total of 0.2 m^2 within 5 x 5 m permanent plots. More details on experimental design for NutNet sites are described in Borer et al. (2014). Total live biomass was then dried at 60°C and weighted to the nearest mg. Single-time-point biomass measurements were performed between 2007 and 2017, depending on the site (Table S2.1). At each site, the data was collected from non-fertilized plots. Most sites contained 30 plots (i.e. pre-treatment plots), while 12 sites (control plots within fertilization experiment) contained less than 10 plots (a minimum of three). We calculated the average of standing biomass from all the plots within a site to obtain a proxy of aboveground grassland biomass – AGB [g/m^2] production per site. While peak standing crop is not a perfect measure of grassland productivity, it has been shown that this method can be a fairly good indicator for the general ranking of grassland productivity (Scurlock *et al.*, 2002) and that it can produce similar estimates compared to those obtained by more complex methods (Lauenroth *et al.*, 2006). Given that 29 sites were exposed to certain disturbances (mowing, burning, grazing) that could affect biomass productivity estimates, we created disturbance scores for each site; mowing, burning and low-intensity grazing each had a score of one, medium intensity grazing a score of two and high-intensity grazing a score of three. The other 43 sites, which did not have any form of management within a period of more than a year before the biomass was collected, received a disturbance score of zero (Table S2.1).

Soil sampling was conducted in the same 5 x 5 m plots by taking three soil cores (2.5 cm diameter) at a depth of 10 cm. The soil was subsequently pooled in one sample per plot, air-dried and analysed for different nutrients (total N and total C, extractable soil P, K, Ca, Mg, Na, S, Zn, Fe, B, Cu, Mn), pH, SOM and CEC (except for the latter two at a few sites, all measurements were performed in the same years of biomass sampling). Total soil C and N [pct] were determined using dry combustion gas chromatography on an Elemental Analyzer (Costech ECS 4010 CHNSO Analyzer, Valencia, CA USA). pH was determined by a pH meter in 1:1 soil: water suspension (A&L Labs, Memphis, TN USA). The concentrations of extractable nutrients [ppm] were analysed using Mehlich-3 analysis (A&L Labs, Memphis, TN USA) which is considered to be suitable for the determination of both macro- and micronutrients in a wide range of soil types and pH conditions (Mehlich, 1984; Jones, 1990).

The measured concentrations were in all cases above the minimum detection level for different micronutrients. Cation exchange capacity [meq/100g] was calculated based on the concentrations of Ca, Mg and K using the method described by Ross & Ketterings (1995). The percentage of soil organic matter was determined using the loss on ignition method, by performing soil combustion at 400 °C. The values of soil parameters were averaged per site.

Based on the site locations, we obtained climatic data using global databases. Mean annual precipitation (MAP) and temperature (MAT) estimates for the period between 1979 and 2013 were derived using the 'Climatologies at high resolution for the earth's land surface areas' database (Karger *et al.*, 2017); hereafter referred to as 'CHELSA' for brevity. We also obtained long-term weather-station climate data from 41 sites and compared the values with those obtained by CHELSA. While CHELSA estimates were very similar to the weather station measurements in the majority of sites, we corrected the CHELSA precipitation values for nine sites for which the measured values differed by more than 15% (Figure S2.1). We further calculated the length of the growing season as the number of months with a mean monthly temperature higher than 5^{o}C. This threshold is considered to be appropriate especially for mid-latitudes (Frich *et al.*, 2002), where the majority of our sites are located, but it was used here as a rough indicator of growing-season length for all the sites. Based on this, mean annual growing season precipitation (MAPgs) was calculated and included in the analyses in addition to MAP because it might better represent the amount of water available to plants during the period of their activity than annual means. The aridity and potential evapotranspiration (PET) data was obtained using CGIAR-CSI Global-Aridity and PET Database (http://www.cgiar-csi.org; Zomer et al. 2008). Data on total inorganic nitrogen deposition [kg/ha/yr] was derived from (Ackerman *et al.*, 2018). We used the average values over the period of years available in the database (1984-1986, 1994-1996, 2004-2006, and 2014-2016) in order to account for long-term patterns of N fertilization via atmospheric deposition.

2.3.2 Statistical analyses

Disentangling the predictors of AGB

To disentangle the role of different climatic and soil properties as predictors of grassland AGB, a structural equation model (SEM) was built based on prior knowledge about the drivers of grassland productivity. The factors that were hypothesised to be overarching global drivers of grassland AGB were climate, N deposition, soil physicochemical properties determining soil fertility (SOM, CEC, pH) and disturbance intensity (Sala *et al.*, 1988; Fay *et al.*, 2003; Huxman *et al.*, 2004; Stevens *et al.*, 2015; Grace *et al.*, 2016; Bünemann *et al.*, 2018). Furthermore, the availability of main macronutrients (N and P) was expected to explain additional variation in addition to the hypothesised main drivers (or to reflect their indirect effect) due to their key role in (co)-limiting grassland productivity worldwide (Elser *et al.*, 2007; Craine & Jackson,

2010; Harpole *et al.*, 2011; Ågren *et al.*, 2012); followed by other nutrients which have also been demonstrated to co-limit grassland productivity (Olde Venterink *et al.*, 2001a; Fay *et al.*, 2015; Lannes *et al.*, 2020). Apart from the direct effects, precipitation was additionally expected to have an indirect influence through atmospheric N deposition as it affects N deposition rates (Prado-Fiedler, 1990), particularly wet N deposition (Kryza *et al.*, 2011; Wałaszek *et al.*, 2013), while climate, in general, was expected to have additional indirect effects via its influence on SOM and pH (Zhao *et al.*, 2019). All these factors were also expected to affect CEC and available soil nutrient concentrations. For example, CEC and availability of different soil nutrients are strongly determined by soil pH and SOM (Havlin, 2004; Bünemann *et al.*, 2018). Nutrient concentrations can also be influenced by atmospheric N deposition (which is more pronounced in the region of strong anthropogenic influences and can also indirectly indicate increased anthropogenic deposition of other nutrients).

We built a SEM representing the influence of different variables in four steps based on the expectations described above. The 'core' SEM was constructed using the most important overarching drivers: climate (MAPgs, MAP, MAT, aridity, PET), N deposition, SOM, CEC, pH and disturbance intensity. We tested an indirect influence of climatic variables through N deposition (precipitation), SOM and CEC. Moreover, we examined an indirect influence of SOM and pH through CEC. The variables with significant paths (either direct or indirect; $P < 0.05$) were retained in the model. In the following steps, groups of nutrients were sequentially added (in all cases, indirect paths from retained climatic factors, N deposition, SOM and pH were tested); main macronutrients were added first (total N, C:N and P) followed by other macronutrients (K, Ca, Mg, Na, S) and finally micronutrients (Zn, Fe, Mn, Cu, B). Only the nutrients with significant paths were retained in each step. The variables with significant paths in the preceding models were always retained even if with additional paths added in subsequent steps their direct effect was no longer significant. The rationale behind this approach is that the variables that are expected to have the most important role on AGB either directly or indirectly through other factors (e.g. climate through soil properties) were given the advantage in the model construction so that their potential direct and indirect influence could be fully explored.

Structural equation models were constructed using the *lavaan* package (Rosseel, 2012). The fit was assessed using standard indices, where model chi-square (χ^2) $P > 0.05$, comparative fit index (CFI) > 0.95, Tucker-Lewis index (TLI) > 0.95, root mean squared error of approximation (RMSEA) < 0.08, and standardised root mean square residual (SRMR) < 0.08 were considered as indicators of a good fit (Hooper *et al.*, 2008). In each step, the model with a good fit and the highest R^2 was selected and reported.

In additional analyses, we constructed a multiple regression model using the factors with a significant direct path on AGB in the final SEM. We examined if different two-way interactions between the selected variables in the best model would emerge as significant predictors of AGB by separately adding each interaction term to the multiple regression model and testing its

significance (using $P < 0.05$ as a significance threshold). The model performance was evaluated via repeated (100 times) k-fold (k = 10) cross-validation using the *caret* package. Finally, we performed variance partitioning between three distinct groups of variables in the best model (climatic variables, N deposition, soil variables) using the *varPart* function in the *modEvA* package to determine the percentage of unique vs shared variance explained by these groups.

The variables were $\log_e$-transformed prior to analyses in case of a skewed distribution to improve normality and linearity. All analyses were performed in R (version 3.3.2) (R Core Team, 2015).

Examining the influence of N and P availability levels and N/NP (co)limitation on the relationship between other selected nutrients and AGB

We hypothesised that the influence of soil nutrients (other than N and P) selected as important predictors of AGB in the prior step would depend on the levels of N and/or P availability. To test the first hypothesis, we assigned each site to two groups according to their C:N levels (low and high) and N deposition levels (low and high) and combined them to obtain a variable with four categories (low C:N - low N deposition, low C:N - high N deposition, high C:N - low N deposition, high C:N - high N deposition). The threshold between 'low' and 'high' levels of N deposition and C:N was based on 50% quantiles (cut-offs of 3.64 kg/h/y and 13.2, respectively). The median value for C:N in our study was comparable to the average C:N value found in worldwide-distributed grasslands (Cleveland & Liptzin, 2007) supporting its use to contrast relatively low and high C:N. Nonetheless, to test the sensitivity of the threshold chosen, we also performed the analyses based on the mean values (N deposition = 5.1 kg/ha/y and C:N = 14). This provided very similar results (Table S2.3). The group with high C:N and low N deposition is here considered as the 'low N availability level'. This assumption is based on the general finding that C:N is a relatively robust indicator of spatial variation in N availability, where increasing C:N indicates decreasing N availability (Andrianarisoa *et al.*, 2009; Wang *et al.*, 2014; Alberti *et al.*, 2015; Vicca *et al.*, 2018), while atmospheric N deposition can substantially increase N availability but it can take very long for this effect to be translated in a decrease of soil C:N (Vicca et al., 2018). Similarly as for N, to test the effect of P availability, the dataset was divided into two parts using the median value of P (29 ppm) on low and high P availability and the relationship between selected nutrients and AGB was assessed for these two datasets.

Given that soil C:N and N deposition may not be accurate indicators of soil N availability for different sites, to further examine if the relationship between selected soil nutrients and AGB is absent in grasslands with low N availability, we also explored this relationship for NutNet sites that had previously been demonstrated to be either N limited / NP co-limited or without N (co-)limitation according to the experiment by Fay et al. (2015). There were 38 sites from our dataset (out of 72 sites in total) for which N (co-)limitation had been assessed in the study by Fay et al. (2015), but these comprised only a few of sites in our 'low N availability' group

and it was therefore not possible to confirm that this group generally contained sites that are N limited.

2.4 Results

Disentangling the predictors of AGB

The 'core' SEM model included a significant direct influence of MAPgs, N deposition, SOM, CEC as well as indirect effects of MAPgs, SOM and pH (Figure S2.5). This model explained 47% of the variation in AGB. After adding macronutrients to the core model, the best model included C:N as an additional variable and this model explained 53% of the variation (Figure S2.5). When micronutrients were added in the final step, only Zn was retained in the best final model, with indirect paths from SOM, pH and N deposition (Figure 2.2). Adding Zn to the model, however, rendered the direct effect of SOM not statistically significant. None of the other soil nutrients were chosen, likely because their influence was already contained in the direct influence of other variables in the core model. The final model explained 61% of the variation in AGB with the highest direct factor loadings for Zn and MAPgs.

A multiple regression model including the variables with a significant direct effect on AGB (MAPgs, N deposition, CEC, C:N and Zn; the individual relationship between these variables and AGB are shown in Figure S2.2) explained 56% of the variation in AGB with no significant two-way interaction effects ($P > 0.05$). Repeated K-fold cross-validation demonstrated that this model can predict 54% of the variation in the validation set. Variance partitioning revealed that the highest proportion of unique (non-shared) variation in the model was explained by soil properties (32%,) followed by climate (16%) and N deposition (4%) (Figure 2.3).

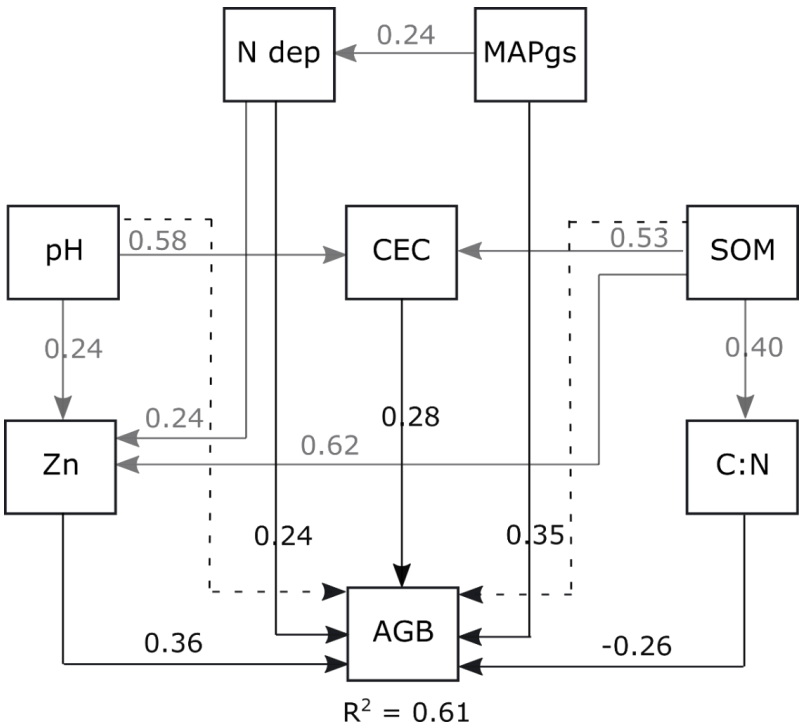


Figure 2.2 Final structural equation model depicting the direct (black lines) and indirect (grey lines) influence of different predictors of AGB. Full lines indicate significant paths and dashed non-significant paths. Factor loadings are indicated only for significant paths. SRMR = 0.06, RMSEA = 0.00, CFI = 1, TLI = 1.002, P ($\chi 2$) = 0.45, df = 11, R2 (AGB) = 0.61. The model has a good fit based on each of the goodness-of-fit criteria.

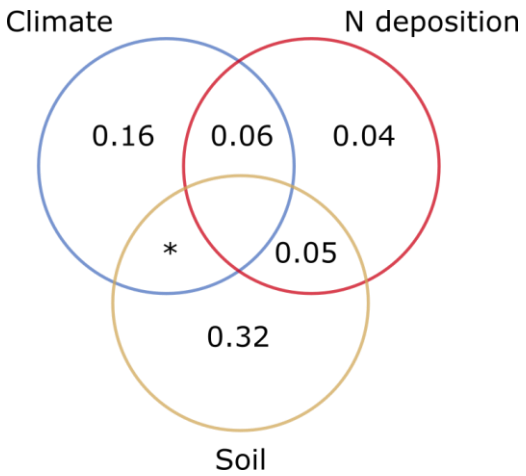


Figure 2.3 Variance partitioning between climate (MAPgs), N deposition and soil properties (CEC, C:N and Zn). *Values lower than 0 are not shown.

The influence of N and P availability levels / N(P) limitation on the relationship between Zn and AGB

According to Liebig's law of the minimum (von Liebig, 1840), the potential important influence of Zn on AGB detected in the SEM may depend on the availability of the scarcest nutrients in relation to the need, i.e. N and P. To create different 'N availability' levels, we used the 50% quantiles as a threshold between low and high levels of C:N and N deposition, splitting the dataset into four classes where the group with high C:N ratio – low N deposition was considered as low 'N availability' group. Simple linear regressions performed for each of these four groups showed that in the 'low N availability group', there was no significant relationship between Zn and AGB, while in the other three groups, there was a significantly positive Zn-AGB with comparable slopes and R^2s ranging from 0.34 to 0.43 (Figure 2.4a, Table S2.2). Correspondingly, when 'low N availability' sites were excluded from the full dataset, Zn explained more variation in AGB (R^2 = 0.37) than in the full dataset. Similarly, Zn-AGB relationship was significant at high P availability group (R^2 = 0.36) and absent at low P availability group (Figure 2.4b, Table S2.2).

To confirm that the Zn-AGB relationship is absent when soil N/P availability is low, we further explored the Zn-AGB relationship in the subset of sites that had previously been demonstrated to be N limited or NP co-limited compared to those that had no N limitation / NP co-limitation in the fertilization study by Fay et al. (2015). In line with the previous analyses, we found that the sites that were N limited / NP co-limited showed no relationship between Zn and AGB while this relationship was significantly positive for sites with no signs of N limitation / NP co-limitation (Table S2.4, Figure 2.5).

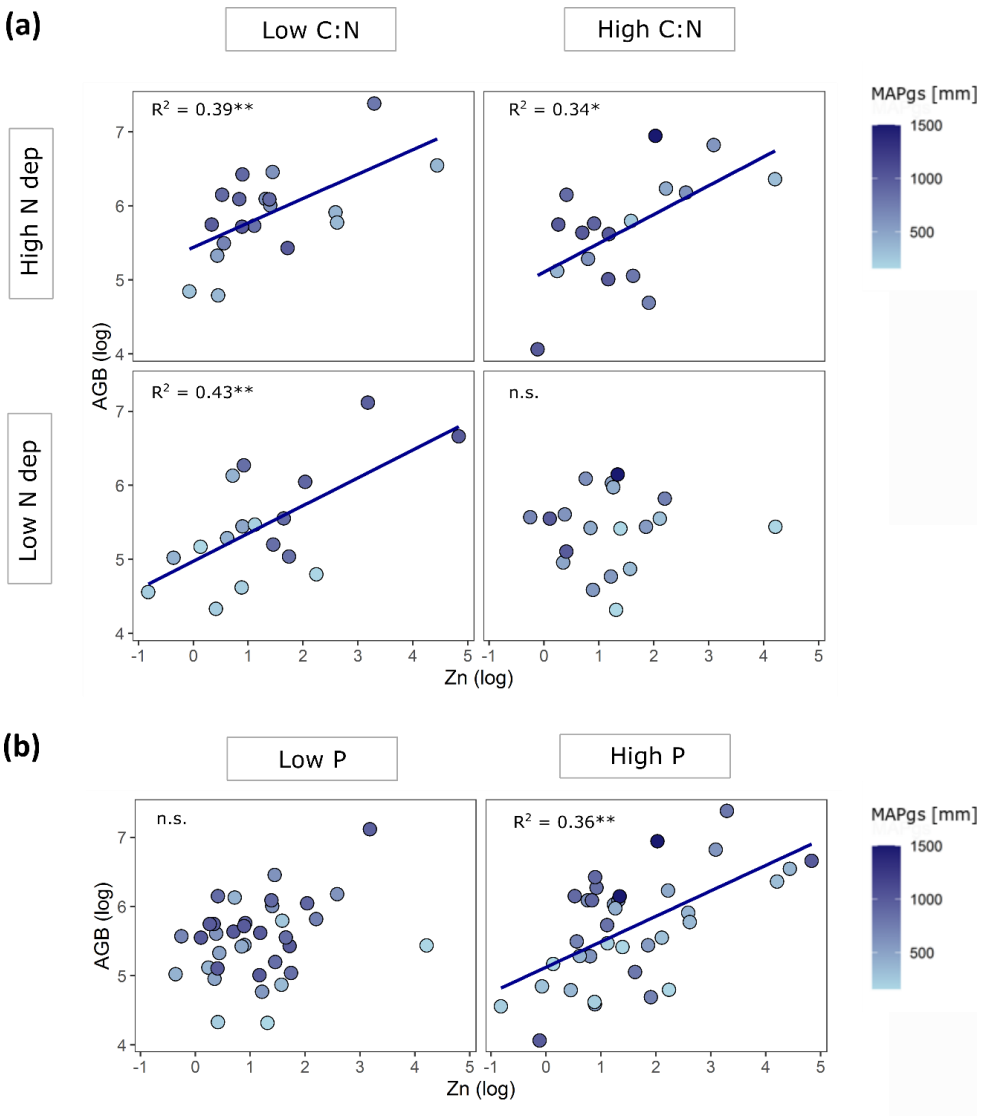


Figure 2.4 a) The relationship between log Zn and log AGB under different levels of C:N and N deposition; from top-left to bottom-right: low C:N - high N deposition, high C:N - high N deposition, low C:N - low N deposition, high C:N - low N deposition. The median values of C:N and N deposition were taken as thresholds based on which the dataset was split into 4 equal groups. Different colours of the points represent different levels of growing season precipitation (which ranges from 160 mm to > 1500 mm per year); **b**) The relationship between log Zn and log AGB at two different P availability levels (low and high). The median value of available P was taken as thresholds based on which the dataset was split into 2 equal groups.

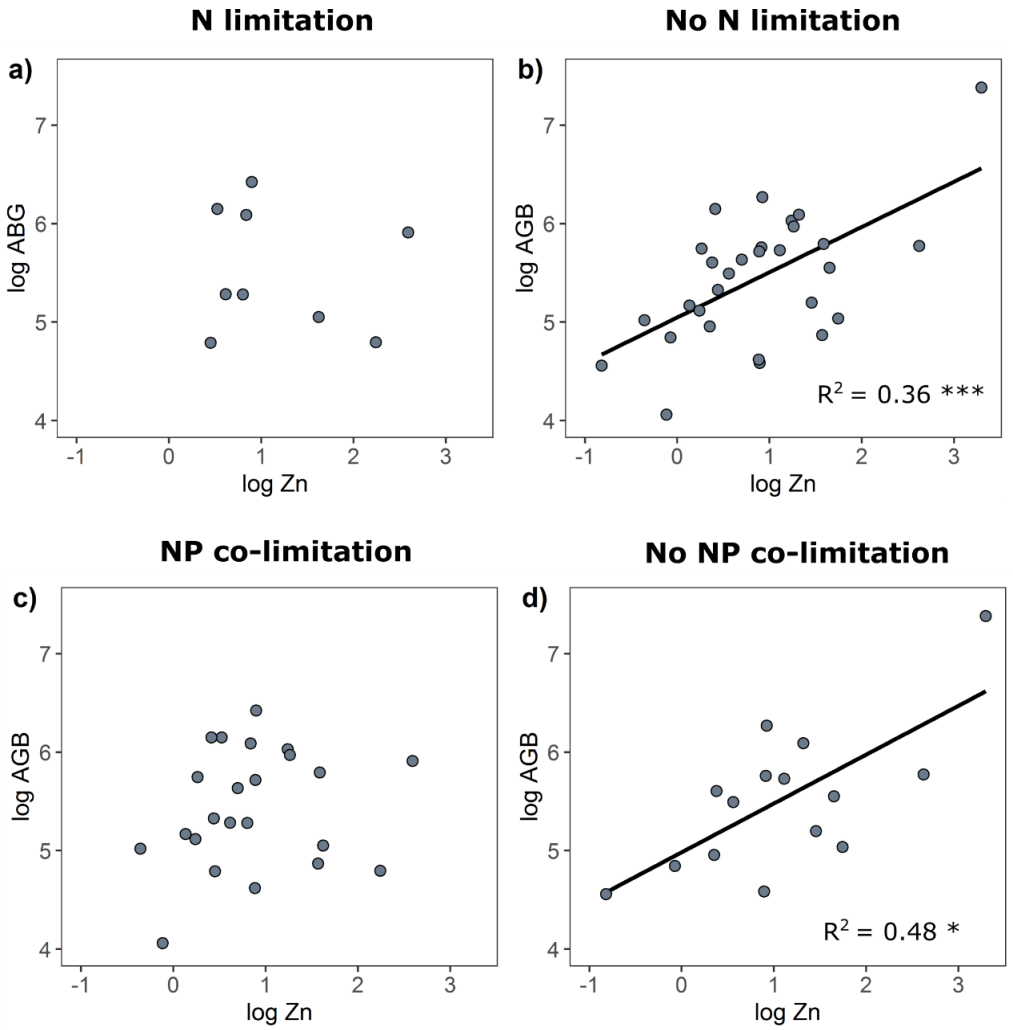


Figure 2.5 The relationship between Zn and AGB in the subset of NutNet sites (n=38) for which the effect of nutrient additions was assessed in the study by Fay et al. (2015). The relationship in the soils that were shown to be **a**) N limited (n = 9) or **b)** without N limitation (n = 29); **c**) NP co-limited (n = 23); **d**) without NP co-limitation (n = 15). Here, it is demonstrated that in the grasslands limited by N or co-limited by N and P, there was no relationship between AGB and Zn while in those that showed no signs of N limitation there was a significant, positive relationship between AGB and Zn.

2.5 Discussion

Our findings demonstrate that the factors that determine soil nutrient availability i.e. CEC, C:N and particularly Zn (all of them likely strongly driven by SOM) are important predictors of global grassland productivity, together explaining 32% of the unique (non-shared) variation in addition to the effect of the broad-scale predictors, climate and N deposition.

Soil nutrient availability as a predictor of global grassland AGB

Soil organic matter and cation exchange capacity both determine the amount of nutrients that can be retained in soil and they are thus among the most important factors driving soil fertility and overall nutrient availability (Havlin, 2004; Bünemann *et al.*, 2018). In this study, both CEC and SOM were strongly correlated with different soil nutrients but, in general, they were better predictors of grassland AGB than the concentrations of primary and secondary macronutrients. This suggests that these fertility indicators already largely incorporate the effect of soil macronutrients on AGB. Nonetheless, indicators of soil N availability (N deposition and C:N) were also selected in the best model in addition to the fertility indicators, where particularly N deposition contributed substantially to explaining the variation in AGB, as previously shown by Stevens et al. (2015). Surprisingly though, soil micronutrient Zn was the single best predictor of variation in AGB, and according to our SEM, much of the effect of SOM on biomass went through its effect on Zn. Indeed, SOM has been shown to play a critical role in the availability and transformation of Zn in soil (Obrador *et al.*, 2003; Cakmak, 2008; Chen *et al.*, 2017).

Zn is a micronutrient essential for plant growth and, even though it is only needed in relatively small concentrations Zn deficiency is known to be widespread in agricultural systems (Sillanpää, 1982, 1990) where it was often found to strongly influence plant growth and limit plant yield (Rashid & Ryan, 2004; Alloway, 2009; Shukla *et al.*, 2014). For instance, it has been shown that almost 50% of 190 agricultural soils investigated in the study by Sillanpää (1990) were deficient in plant-available Zn, which was more frequent than any other micronutrient. Zn serves as a structural component of a large number of proteins (Andreini *et al.*, 2006) and is involved in important metabolic functions including photosynthesis and defence against drought and disease (Cakmak, 2000). When the amount of Zn in the soil is inadequate, plants can suffer due to dysfunction of different enzymes and changes in physiological processes resulting in inhibited plant growth and development and hence reduced productivity (Cakmak, 2000). Micronutrient deficiencies rarely occur due to low total concentrations of nutrients in soil (primary deficiencies) but they rather occur as a result of soil factors reducing the availability to plants of otherwise ample supplies of micronutrients (secondary deficiencies) (Sillanpää, 1982). For instance, alkaline soils, such as those rich in $CaCO_3$, are often Zn deficient due to low Zn solubility at high pH and/or because Zn gets

adsorbed to carbonates and become unavailable to plants (Chen & Barak, 1982; Fageria *et al.*, 2002). Moreover, drylands have been found to have particularly low Zn availability due to their high soil pH that limits the formation and preservation of clay minerals and soil organic matter (Moreno-Jiménez *et al.*, 2019). On the other hand, Zn can be toxic to plants in higher concentrations, but Zn toxicity is expected to be much less common in natural soils compared to Zn deficiency (Alloway, 2008) and there were no indications of possible Zn toxicity in the dataset used in this study.

The potential importance of micronutrients for the productivity of non-agricultural grasslands on a global scale was previously hinted at in the synthesis study on fertilization experiments by Fay et al. (2015) where multiple nutrients (including micronutrients) co-limited productivity in many grasslands worldwide. Moreover, other experimental studies (Cech *et al.*, 2008; Lannes *et al.*, 2016) have also demonstrated the potential effect of micronutrient additions in non-agricultural grasslands but few studies examined additions of individual micronutrients. For example, a recent study in Brazilian Cerrado grasslands showed that a single micronutrient (boron) limits plant productivity in these grasslands (Lannes *et al.*, 2020). As for other natural systems, Zn deficiency has been linked to fruit production in European forests (Fernández-Martínez *et al.*, 2014). Therefore, it seems plausible that micronutrient limitations of plant productivity, widely observed in agricultural systems, could likewise occur in natural systems including (semi)natural grasslands. Moreover, given that many grasslands are located in relatively arid areas and that aridity is expected to increase due to climate change, the availability of some micronutrients could decrease further (Moreno-Jiménez *et al.*, 2019), potentially leading to stronger micronutrient deficiencies in future.

The influence of N (co-)limitation on Zn-AGB relationship

We find that the relationship between Zn and AGB prominent in the overall dataset, was entirely absent in the regions with high C:N and low N deposition (potentially indicating low N availability) as well as in the grassland with relatively low P availability. The same was found in the subset of NutNet grasslands previously shown to be N limited/ NP co-limited according to the fertilization study by Fay et al. (2015). The lack of the relationship between Zn and AGB under low N/P availability may be due to decreased uptake of Zn at low N (Cakmak *et al.*, 2010) and/or because, in accordance with Liebig's law of the minimum, grassland AGB is primarily constrained by the lack of the scarcest nutrients in relation to the need. It has previously been demonstrated that the effect of some micronutrients on productivity depends on soil N availability; the grain yield from plants grown on soils deficient in N and Zn only responded to Zn additions when N fertilizer was added (Loneragan & Webb, 1993). Similarly, micronutrient additions (Zn and B) in relatively infertile agricultural regions of India had the strongest effect on crop yield when added together with N and P (Sahrawat *et al.*, 2010). Moreover, numerous studies have shown that N is important for Zn uptake and translocation (Shi *et al.*, 2010; Cakmak *et al.*, 2010; Erenoglu *et al.*, 2011; Gupta *et al.*, 2016)

which is why higher concentrations of Zn were observed in plants treated with N compared to control (Kutman *et al.*, 2010; Erenoglu *et al.*, 2011). These results suggest that, besides NP co-limitation, also N(P)-Zn co-limitations could be common in grasslands and that in addition to anthropogenic N/P deposition, also metallic micronutrient (Zn) deposition (which is likewise tightly related to human activities) may be important to consider when predicting the fate of grassland systems under environmental changes.

2.6 Conclusion

In this study, we demonstrate that soil properties determining soil fertility – soil organic matter and cation exchange capacity – can significantly improve the predictions of grassland productivity patterns given that they can integrate the effect of overall nutrient availability. Nonetheless, soil Zn was shown to explain additional variation in productivity indicating that Zn deficiency might be common in grasslands. While the effects of the widespread deficiency of micronutrients on plant productivity are well-documented for agricultural plants, the individual role of micronutrients, such as Zn has been rarely examined in natural ecosystems. The study by Fay et al. (2015) was the first to reveal the importance of micronutrients as limiting factors across non-agricultural grasslands worldwide and our findings demonstrate there is also a clear link between in-situ soil Zn concentrations and grassland productivity. We argue that it would be beneficial to measure and report the concentrations of micronutrients, particularly Zn, (both in plants and soil) in studies investigating grassland productivity, including nutrient fertilization studies. This would help to further examine the extent of deficiencies of these nutrients, their link with grassland productivity, as well as their role in grassland responses to environmental changes, such as increased anthropogenic nutrient depositions.

No observational study can fully disentangle the effect of different correlated soil properties and nutrients (e.g. soil organic matter and Zn in this study). Our results, however, motivate future experimental studies in grasslands to focus more on micronutrient additions specifically, alone, together and in combination with organic matter (that can help increase the availability of added nutrients) and N(P) additions, to assess the potential interaction between N(P) and Zn limitation and to further unravel the role that nutrients play in determining grassland productivity.

Soil physicochemical properties as key predictors of local and global variability in grassland aboveground biomass production

CHAPTER III

CHAPTER III

Soil physicochemical properties as key predictors of local and global variability in grassland aboveground biomass production

3.1 Abstract

Plant biomass production has long been known to depend on the soil they grow on and the climate they live in. While the effects of climate have been widely studied, which of the multitude of individual soil properties determining soil fertility are the best predictors of global grassland productivity patterns has been rarely assessed, nor has it comprehensively examined how much of the total variation in productivity they can explain. In this study, we used globally replicated local-scale gradients in plant productivity to examine which soil physicochemical properties contribute to explaining variation in aboveground biomass (AGB) production in (semi)natural grasslands on the local and the global scale. We found that the amount of soil organic matter, percentage of sand, bulk density and the amount of calcium in the soil were the best soil predictors of AGB production and this was consistent for both spatial scales. On the global scale, the variation of AGB was better explained by soil properties than by climate. Our results emphasize that these soil properties, which are generally associated with soil's capacity to retain and provide nutrients and water, are critical to understand current and predict future grassland productivity patterns.

3.2 Introduction

Plant productivity is one of the main grassland ecosystem services and an important component of the global carbon budget (White *et al.*, 2000). The drivers of grassland productivity patterns have thus been a long-standing topic of ecological research demonstrating the key role of climate, biotic factors and soil physicochemical properties (van der Heijden *et al.*, 1998; Hector *et al.*, 1999; Fay *et al.*, 2003; Huxman *et al.*, 2004; Elser *et al.*, 2007; Sala *et al.*, 2012).

Soil physicochemical properties such as texture and structure, amount of organic matter, cation exchange capacity, pH and concentrations of different nutrients jointly determine nutrient availability (Vicca *et al.*, 2018; Van Sundert *et al.*, 2019) and water holding capacity (Kirkham, 2005), both of which are of crucial importance for plant growth (Chapin, 1980; Kirkham, 2005; Vicca *et al.*, 2012; Fay *et al.*, 2015; He & Wang, 2019). For example, sandy soils with a low

amount of organic matter have a low capacity to retain water (Kirkham, 2005). They also have low cation exchange capacity, and cations are therefore easily leached from sandy soils (Lehmann & Schroth, 2005) typically resulting in suboptimal nutrient availability. Soil pH also determines plant nutrient availability as it influences the chemical composition of the soil; low pH favours leaching of base cations and the mobilization of iron (Fe) and aluminium (Al) further impedes their uptake. Low pH can also reduce phosphorus (P) availability by enhancing complex formation with Al and Fe compounds. In contrast, in alkaline soils, high concentrations of calcium ions (Ca^{2+}) in soil solution limit plant uptake of ions like Fe^{2+} and magnesium (Mg^{2+}) and promote the precipitation of calcium phosphates, thereby reducing P availability (Larcher, 2003).

Soils can also affect the relationship between climate and productivity by buffering or aggravating climatic impacts (Sala *et al.*, 1988; Fernández-Illescas *et al.*, 2001; Wang *et al.*, 2009). Folberth *et al.* (2016) argued that the effect of climate changes on plant yield can strongly depend on soil texture where, for example, a moderate decrease in precipitation is expected to cause a more severe impact on yield in soils with lower water holding capacity, while the soils with fine texture might buffer the decrease in precipitation by storing water for longer periods. They stress that accounting for soil variability is essential to improve models assessing the impacts of climate change on biomass yield.

Surprisingly though, few in-depth studies have attempted to assess the influence of soil properties and nutrients and determine the best soil-derived predictors of global grassland productivity. For instance, Stevens *et al.*, (2015) included nitrogen, phosphorus and soil pH in their global study of predictors of grassland productivity where they found little support that global AGB was related to these soil properties. On the other hand, on a regional scale, Yang *et al.* (2009) and Sun *et al.* (2013) found that soil texture, moisture, nitrogen and organic carbon were important predictors of AGB productivity on the Tibetan Plateau while Olde Venterink et al. (2001) showed that AGB was related to the factors representing N and P availabilities and soil bulk density in wet meadows of the Netherlands and Belgium. Given that the main soil drivers of productivity – nutrient availability and water holding capacity – depend on the interplay of several soil properties and nutrients, ideally, they all need to be studied in concert (Vicca *et al.*, 2018). The lack of harmonized, comparable databases containing plant biomass production data as well as comprehensive soil data has been a great challenge hampering the large-scale analyses of the importance of soil physicochemical properties as predictors of biomass productivity (Vicca *et al.*, 2018).

Here, we collected data on climate, aboveground biomass and 17 soil physicochemical properties measured in a standardized manner from worldwide distributed grassland sites within the Herbaceous Diversity Network (HerbDivNet) (Fraser *et al.*, 2015). Most HerbDivNet sites contain a 'productivity' gradient, where the existing variation in plant AGB cannot be explained by climate (which is relatively stable within the sites). Hence, this setup

provides a unique opportunity to determine the influence of soil properties for both within-site (i.e. local-scale) and between-site (i.e. global-scale) variation in AGB. While the within-site variation enabled a detailed assessment of the relationship between soil properties and AGB for different sites exposed to contrasting climates, the between-site variation allowed us to disentangle the relative importance of soil properties and climate globally. Specifically, we aimed to investigate: i) which soil properties are the best predictors of local-scale variation in grassland AGB; ii) which climatic and soil factors are the best predictors of global-scale variation in grassland AGB and what is the relative contribution of climate versus soil predictors and iii) whether the local-scale relationships between AGB and soil properties persist on the global scale when climate is taken into account.

3.3 Materials and methods

3.3.1 Sampling sites and data collection

Data was collected from 21 Herbaceous Diversity Network (HerbDivNet) sites located in 12 countries distributed over six continents (Figure 3.1). All HerbDivNet sites are grassland areas dominated by vegetation representing the regional species composition. Each of the 21 sites contained between two and six plots of 8 x 8 m; nine sites contained six plots, three sites contained four plots, one site three plots and eight sites contained two plots (Table S3.1). In total the dataset thus contained 85 plots. The sites with six plots were chosen to represent an estimated gradient in productivity (low, medium and high) with two plots at each productivity level, while the sites with four plots normally had two levels of productivity and other sites did not contain a clear productivity gradient (Figure S3.1). More details on experimental design can be found in Fraser et al. (2015). All the plots within one site were subject to the same or very similar climatic conditions (Table S3.1). Mean annual temperature (MAT) across sites ranged from 1.5 °C to 20.1 °C and mean precipitation ranged from 294 mm to 1237 mm (Figure 3.1). Peak annual AGB values spanned a wide range from 13 g/m^2 to 1187 g/m^2 (Figure S3.1).

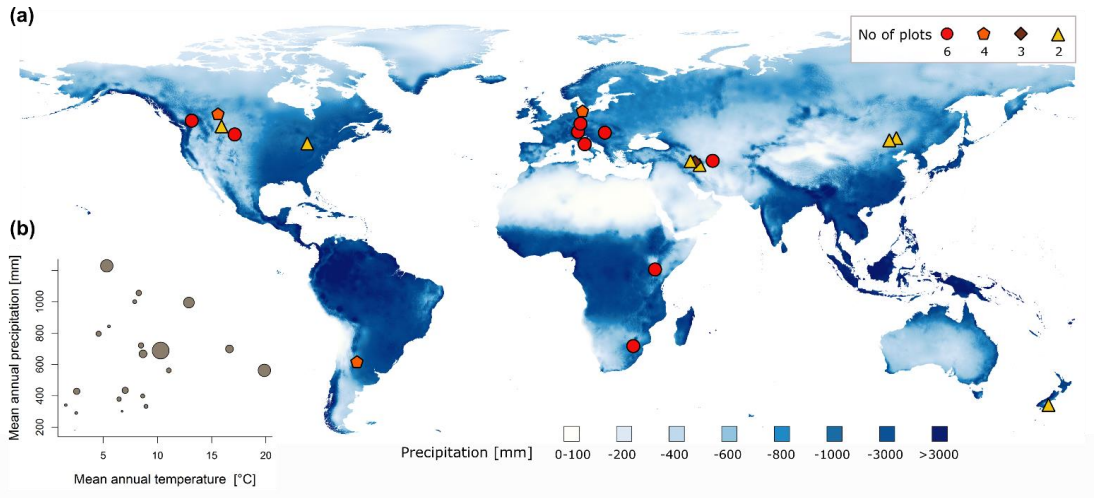


Figure 3.1 a) The location of 21 HerbDivNet sites across a precipitation gradient (colours of points indicate the number of plots per site). **b**) The distribution of sites under different combinations of mean annual temperature (MAT) and mean annual precipitation (MAP) values. The size of points corresponds to mean AGB values per site.

Biomass sampling

Total aboveground biomass was harvested from each m^2 of each 64 m^2 plot at the peak of the growing season. Sampling was performed in the period between 2012 and 2017 in a single sampling event per site (Table S3.1). Litter was first excluded from the total biomass and live biomass was dried and weighed. Based on this, average peak AGB production [g/m^2] was calculated for each plot. The sites were not fertilized, but most of them were subject to some form of low-intensity management (mowing, grazing, burning) and sampling was therefore performed at least three months after the last disturbance event (the plots were fenced before biomass collection).

Soil sampling

Soil samples were taken in a single sampling event at the peak of the growing season in the period between 2017 and 2018, depending on the site (Table S3.1). For each plot within a site, five samples were taken using soil corers from four corners and the centre of the plot at two soil depths (0-10 cm and 10-20 cm) Samples were pooled into one composite sample per soil depth (a total of 170 samples), air-dried and sieved at 2 mm after which they were sent to the University of Antwerp for further analyses. Besides this, additional samples were taken at two soil depths with corers of a known volume to determine the average bulk density (w/v) and gravel content (v/v) in the soil.

Climatic data

Mean annual precipitation (MAP) and temperature (MAT) for the period between 1979 and 2013 were derived from the CHELSA database (Karger *et al.*, 2017) based on the geographical position (latitude and longitude) of each plot. We also calculated the mean growing season precipitation (MAPgs) by summing the monthly precipitation of the months with a mean monthly temperature higher than 5ºC as in Chapter II. With this threshold, the growing season precipitation might be overestimated for two African sites that have short periods with drought (with a temperature higher than 5ºC) during which plants are dormant, but given that precipitation in these months is very low compared to the other months, the overestimation is negligible. We derived data on aridity and potential evapotranspiration (PET) using CGIAR-CSI Global-Aridity and PET Databases.

3.3.2 Soil analyses

Besides bulk density (BD) and gravel content, the following soil parameters were analysed: soil organic matter (SOM), total nitrogen (N), total carbon (C), total phosphorus (P), available P, exchangeable bases (including calcium - Ca, potassium - K, magnesium - Mg), base saturation (BS), cation exchange capacity (CEC), pH and soil texture (% sand, % silt, % clay). For each plot, it was estimated whether the soil was 'shallow' (< 20 cm) or deeper than 20 cm. These particular soil properties were chosen to be measured because they are known to be among the most important factors determining soil fertility and plant productivity (Vicca *et al.*, 2018; Van Sundert *et al.*, 2020). Micronutrients were not measured in this study as it was initially assumed that their contribution would be negligible (but see the results of Chapter II; however, the analyses were later hampered by COVID restrictions).

SOM [%] was calculated as the loss of dry matter at 550°C expressed as a percentage of dry matter (Heiri *et al.*, 2001). Total soil N and total C [%] were determined on ground soil, dried 48h at 70°C, using the Flash 2000 CN analyser (ThermoFisher Scientific, Waltham, MA, USA). Each sample was analysed in triplicate and averaged. Total P [ppm] was determined using acid digestion with H_2SO_4, salicylic acid, H_2O_2 and selenium, following the method of Novozamsky et al. (1983). Available P [ppm] was analysed following both the Olsen (Olsen *et al.*, 1954) and the Bray (Bray and Kurtz, 1945) methods using a Continuous Flow Analyser (CFA) SAN++ (Skalar, Breda, The Netherlands). CEC [meq/100g] and BS [%] were estimated based on the exchangeable H^+ and total exchangeable bases – TEB [meq/100g]. For this, cations in ammonium acetate extract were measured following Reeuwijk (2002) using an Inductively Coupled Plasma Spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA) and acidity was determined following the method by Brown (1943). pH of air-dried soil was measured using a pH meter (Hanna HI 3222; Hanna Instruments, Woonsocket, RI, USA) in 1:2.5; soil : 1M KCl suspension (Blakemore *et al.*, 1987). Soil texture was analysed by

determining the percentage of primary particles (sand: 2000-53 µm, silt: 53-2.0 µm, and clay: < 2.0 µm), following the method of Gee & Bauder (1986). Bulk density [kg/m^3] was determined by dividing the dry (60°C) weight of soil (sieved at 2 mm) by in-situ soil volume. Given that gravel weight was excluded from soil weight, BD measured in this way was an indicator of the amount of fine-soil per total soil volume (this measure thus represents fine-soil bulk density). Gravel volume was determined by submerging the gravel in a known volume of water in a measuring cylinder and subtracting that volume from the total volume read after adding the gravel. This value was then divided by in-situ soil volume to obtain the percentage of gravel volume.

The soil properties measured here are relatively stable over time and they are recommended to be used when comparing the soil status across different sites (Vicca *et al.*, 2018).

3.3.3 Statistical analyses

Local-scale (within-site) variation in AGB

First, 17 soil variables measured from the soil depth 0-10 cm (d1) were considered as potential predictors of within-site variation in AGB: SOM, total N, total P, available P-Olsen, available P-Bray, Mg^{2+}, Ca^{2+}, K^{+}, BS, BD, CEC, pH, C:N, % sand, % silt, % clay and % gravel. We also tested for potential non-linear effects of SOM and pH. Prior to the analyses, variables (besides pH and BS) were $\log_e$-transformed to reduce positive skewness and scaled and centred (mean = 0, sd =1) to allow the comparison of effect-sizes.

Forward selection of the best linear mixed effect model (constructed using the *lmer* function from the package *lme4*), where 'site' was included as a random effect, was performed by starting from a single predictor with the lowest AICc (second-order AIC criterion). Variables were added sequentially until the decrease in AICc was smaller than 2. At each step, normality of the residuals of the model was tested using the Shapiro-Wilk test and the model was accepted only if the residuals did not significantly deviate from a normal distribution ($W > 0.9$, $P > 0.05$). R^2 of selected models was calculated using the *r.squaredGLMM* function from the package *MuMln*. We used the *lme* function from the package *nlme* to examine if different variables included in the best mixed-effect models had a significant effect on AGB ($P < 0.05$).

The model selection described above (conducted using the variables measured at soil depth d1) was repeated using the variables measured at the soil depth 10-20 cm (d2) to compare the predictive power of d1 versus d2 soil variables. All additional analyses were conducted using the surface (or subsurface) soil dataset that explained the greatest amount of variation in AGB.

Given that shallow soils can constrain plant productivity, we performed sensitivity analyses to test if the findings were substantially affected by the differences in soil depth (i.e. very shallow

compared to deeper soils). To this end, another model selection was performed excluding the plots with very shallow soils (not deeper than 20 cm).

Global-scale (between-site) variation in AGB

To examine the most important global-scale predictors, we first subset the dataset 1000 times, where each subset contained exactly one randomly selected plot from each site (hence, n = 21 for each subset). This way, we excluded the effect of within-site variability in AGB from the analysis.

We analysed these subsets for climate variables and soil variables separately. To determine which climatic variables are important predictors of global variation in AGB we performed model selection for each of the subsets using the *glmulti* function based on the AICc (second-order Akaike Information Criterion). Glmulti is a package for automated model selection and multi-model inference that fits all possible models based on the input variables (Calcagno & de Mazancourt 2010). Considering the small sample size per subset, to avoid overfitting, we constrained the number of variables that can be included in the model to a maximum of four variables. After obtaining the best models for 1000 subsets, for each of the input variables, we calculated the percentage of best models in which a variable was selected. We removed those variables that appeared in a small number of the best models (less than 20%) in order to obtain the most influential variables. We calculated the mean R^2 of the model containing the retained variables. This way, we determined how much variation in global AGB could be explained by climatic variables alone.

In the same manner, we selected the most important soil predictors. To reduce the number of input variables, we did not include four soil variables that were highly correlated other variables selected in the local-scale analysis; i.e. total N, CEC, silt/clay were not included given that they were highly correlated with SOM, Ca, sand, respectively (Figure S3.2) and they would thus not add much to the interpretation of the model. This was also done to avoid potential discrepancies between the local- and global-scale models that are merely due to strong correlations among some of the soil measurements in our dataset. Moreover, as an indicator of P availability, only P-Olsen was included given that it was highly correlated with P-Bray but is showed a slightly better correlation with AGB (Figure S3.2). The variables that were included in the model selection were: SOM, % sand, BS, Ca, K, Mg, C:N, pH, total P, P-Olsen, BD and % gravel.

Finally, we combined the important climatic and soil variables and performed backward model selection by removing variables that increased model AICc until we obtained the model with the lowest AICc. From this model, we used mean linear regression coefficients as indicators of mean effect sizes for each variable. We also tested if including an interaction term between climatic factors and soil factors in the model would result in a better model (lower AICc).

To examine the direct and indirect effects (based on prior knowledge) of soil and climatic variables selected in the previous step, structural equation models (SEM) were created using the *lavaan* package in R (Rosseel, 2012). We tested both the direct influence of climatic factors on AGB and the indirect paths, through soil variables (Zhao *et al.*, 2019). Regarding soil variables, direct links from indicators of soil texture (e.g. % sand, silt, clay, gravel) were tested as well as indirect links through other soil variables which can all be influenced by soil texture. For organic matter, both direct and indirect links through bulk density and soil nutrients were considered because organic matter content influences soil bulk density (Heuscher *et al.*, 2005) and the amount of nutrients in the soil (Kononova, 1966). The model with a good fit, the highest R^2 and all significant paths was reported. The fit was assessed using standard indices, where model chi-square (χ^2) $P > 0.05$, comparative fit index (CFI) > 0.95, Tucker-Lewis index (TLI) > 0.95, root mean squared error of approximation (RMSEA) < 0.08, and standardised root mean square residual (SRMR) < 0.08 were considered as indicators of a good fit (Hooper *et al.*, 2008). An alternative SEM was created to test the potential causal relationship from AGB to SOM since aboveground plant inputs can determine the total amount of organic matter in the soil. This model was then compared with the original model (with a path from SOM to AGB) in terms of the strength of the path coefficients and the overall goodness-of-fit of the models.

All statistical analyses were performed using the R software, version 3.6.1 (R Core Team 2019).

3.4 Results

Local-scale (within-site) variation in AGB

Among 17 soil variables measured at depth d1, the best mixed-effect model explaining the variation in AGB included: SOM, Ca, BD, % sand and total P (marginal $R^2 = 0.52$, conditional $R^2 = 0.74$, AICc = 186) (Figure 3.2). Therefore, most of the variation in the model was explained by fixed factors (52%), and the remaining 22% was explained by differences between the sites (i.e., the random intercept accounting for between-site variation). Other models with AICc within 2 units difference included total N instead of total P, % silt instead of % sand and CEC instead of Ca. Given that silt/sand, CEC/Ca and N/P are strongly correlated (Pearson r >= 0.8, Figure S3.2), we continued with the initial best model. In this model, SOM had the highest effect size (0.66), followed by Ca (-0.55), bulk density (0.36), % sand (-0.32) and P (0.30) (Figure 3.3). To investigate if soils with a high amount of Ca, that are potentially rich in $CaCO_3$, are the reason for the unexpected negative coefficient of Ca in the model, we excluded the sites with high values of extractable Ca (Ca > 10 meq/100g according to Hazelton & Murphy (2019)). This analysis showed that the effect of Ca in the model was not significant for soils with Ca > 10 meq/100g (P > 0.05), while the effects of SOM and BD remained significant (Table S3.3). The plots with a high amount of soil Ca were dominant in our dataset (n = 50), indicating a potentially high prevalence of calcareous soils in these grassland sites.

When using the soil predictors measured at the soil depth d2, the best model included SOM, Ca and % sand (Table S3.3). This model was comparable to the one with variables measured at soil depth d1, but fixed factors in this model explained less variation in total (marginal $R^2 = 0.45$, conditional $R^2 = 0.79$). Hence, we continued with the d1 measurements.

The model selection where the plots with shallow soils were excluded from the dataset showed that the most important predictors were again SOM, Ca and % sand (marginal $R^2 = 0.51$, conditional $R^2 = 0.87$), while BD and P were not selected in the best model (Table S3.3).

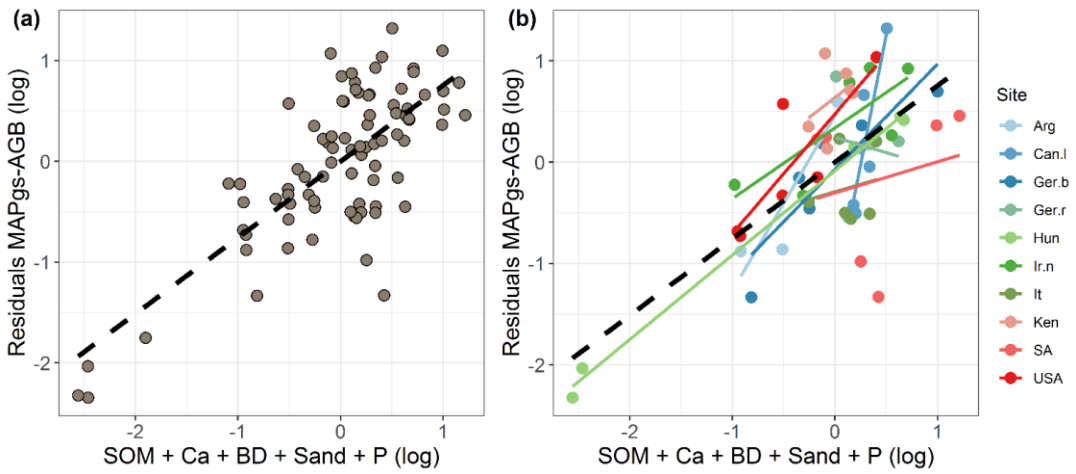


Figure 3.2 a) Relationship between a composite variable (sum of coefficient-weighted variables selected in the best mixed-effect model) and residuals of MAPgs-AGB relationship (aboveground biomass corrected for the effect of growing season precipitation). The graph demonstrates that this composite variable has a good global relationship with AGB across all the plots (n = 85). **b**) The same relationship focusing on the within-site (local-scale) patterns. Here, only the plots belonging to the sites that contained a substantial within-site variation of AGB (at least 50% difference between the plot with the lowest and the plot with the highest AGB value; n = 61) are shown; sites are depicted in different colours and connected with full lines of the same colour. The dashed line represents the association of the residuals of MAPgs-AGB with the composite variable across the sites.

Global-scale (between-site) variation in AGB

On the global scale, five climatic variables (MAP, MAPgs, PET, aridity and MAT) together explained 45% of the variation in AGB (sd. 11%). Out of these five variables, PET, aridity and MAT were selected in less than 20% of the best models and they were therefore excluded from further analyses. MAP and MAPgs each occurred in more than 20% of the best models (Figure S3.3) and together they explained on average 38% (sd. 10%) of global AGB variation.
The model selection analysis for soil predictors of AGB (i.e. without accounting for the effect of climatic variables) resulted in seven out of 12 soil variables being selected in more than 20% of the best models. These were: SOM, Ca, BD, % sand, available P, pH and BS (Figure S3.3).

When climatic and soil variables were combined, the best model following backward selection included MAPgs, SOM, Ca, BD and % sand (mean R^2 = 0.77, sd. 0.08, mean AICc = 53). Variance partitioning demonstrated that out of 77% of the variation explained by this model,

7% of the unique variation could be attributed to MAPgs, 45% to soil properties and 25% was shared between the two (Figure S3.4). SOM had the highest mean effect size, followed by Ca, % sand, BD, and finally MAPgs (Figure 3.3). Adding an interaction term between MAPgs and different soil properties did not improve the model (all the models with an interaction term had a higher AICc than the initial model).

Overall, these results demonstrate that almost the same set of soil variables were the best predictors of AGB both on the local scale and the global scale (Figure 2.3 & Figure 3.3).

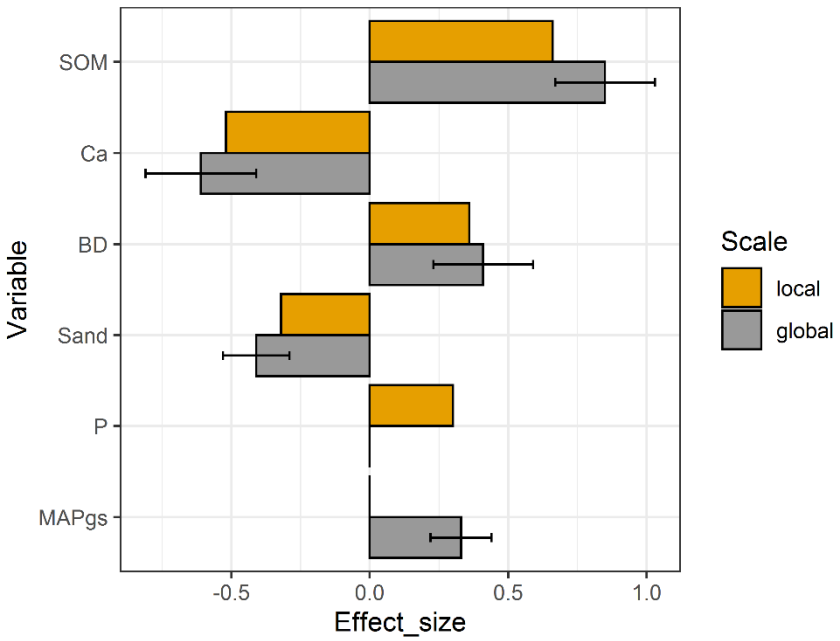


Figure 3.3 Effect sizes / mean effect sizes of different variables included in the best models explaining local- (within-site) vs global-scale (between-site) variation in AGB. Error bars for the global-scale models indicate the standard deviation of the mean effect sizes of each variable from 1000 best models.

The best structural equation model (with an appropriate fit and all significant paths) showed that, while SOM had the highest direct factor loading on AGB, precipitation and % sand had an indirect influence through SOM and Ca, and SOM had an indirect influence through Ca and bulk density (Figure 3.4). Precipitation had a negative loading on Ca because sites with higher Ca levels were located in relatively arid areas. An alternative model with a path from AGB to SOM (Figure S3.5), had a very poor fit and although the path was significant, the loading was low compared to the effect of SOM on AGB (0.29 vs 0.77, respectively). Moreover, even though the alternative model contained the same paths to AGB as the original model (except

the one from SOM), the amount of variation in AGB explained by this model dropped substantially compared to the model with the path from SOM (from 64% to 44%).

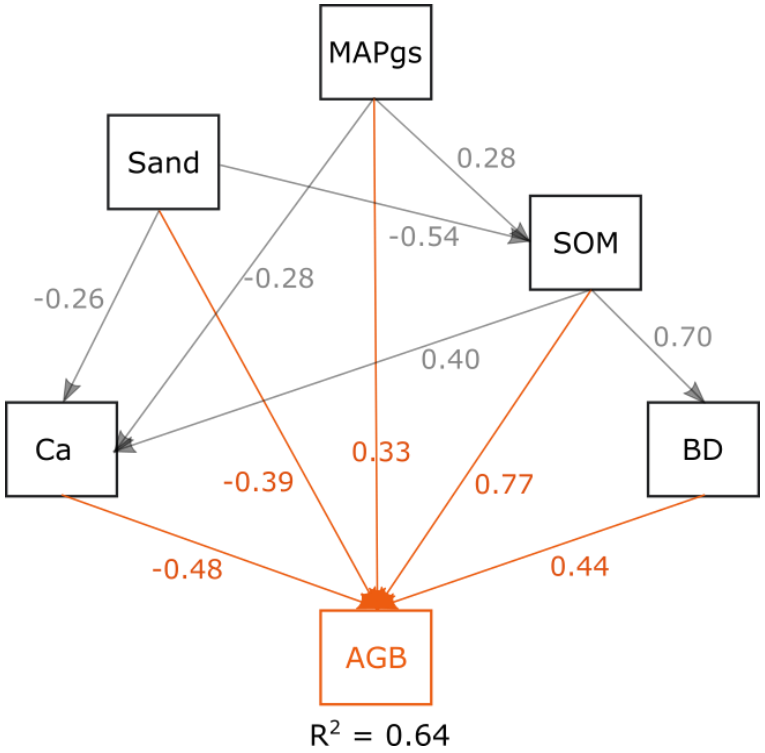


Figure 3.4 SEM depicting the direct (orange arrows) and indirect (grey arrows) influence of the most important climatic and soil variables on global-scale variation in grassland AGB (values on the arrows represent factor loadings). All paths were significant. SRMR = 0.02, RMSEA = 0.00, CFI = 1, TLI = 1.04, P (χ2) = 0.76, df = 3, R2 (AGB) = 0.64. The model has a good fit based on each of the goodness-of-fit criteria.

3.5 Discussion

Climatic factors, such as mean annual precipitation, have typically been used as key determinants of aboveground biomass patterns in grasslands worldwide (Sala *et al.*, 1988; Jobbágy *et al.*, 2002) and predictions regarding the influence of future global changes are primarily focused on the changes of future precipitation patterns on grassland productivity (Ma *et al.*, 2010; Hsu *et al.*, 2012). In this study, we examined which soil physicochemical properties measured at two soil depths (0-10 cm and 10-20 cm) can contribute to explaining additional variation in AGB production on a local and global scale. The amount of soil organic matter, percentage of sand, fine-soil bulk density and the amount of calcium in soil at 0-10 cm were the best predictors of AGB production and this was consistent for both spatial scales. On the global scale, these soil factors together had better predictive power and explained a higher amount of unique variation than climatic factors. This suggests that the variation in soil properties can outweigh the importance of climate for grassland productivity across different sites, which is line with the findings by Folberth et al. (2016) for modelled biomass yield in agricultural systems.

The amount of soil organic matter was the strongest and most consistent predictor of both local- and global- scale variation in AGB. SOM has long been considered as the key indicator of soil fertility and soil health (Kononova 1966; Bünemann *et al.*, 2018; Van Sundert *et al.*, 2019) due to its impact on soil chemical, physical, and biological properties; it increases water holding capacity of the soil, affects nutrient availability by providing and retaining nutrients and provides a substrate for soil organisms that drive nutrient cycling (Reeves 1997; Oldfield *et al.*, 2018). In their meta-analysis of the predictors of aboveground biomass in alpine grasslands in Tibet, Sun et al. (2013) found that soil organic carbon (related to the amount of SOM) was strongly associated with productivity, while Oldfield et al. (2018) experimentally demonstrated the important direct role of SOM on plant productivity.

The formation of SOM depends on the interactions between climate, inherent soil physicochemical properties (e.g. texture), and the nature of inputs (Bot & Benites, 2005). The inputs can partially come from aboveground plant biomass itself, but it has been argued that much of SOM derives from below-ground inputs (Schmidt *et al.*, 2011; Cotrufo *et al.*, 2013). Our SEMs also indicate that SOM likely had a much stronger influence on plant AGB than the other way around. However, it is probable that the observed relationship between SOM and plant AGB reflects, to some extent, the two-directional relationship between plant growth and soil organic matter content.

Besides the direct link with AGB, we found that SOM had an indirect link through fine-soil bulk density and the amount of Ca in the soil. Ca was negatively associated with AGB in our models even though in itself it is one of the essential soil macronutrients that regulates and promotes plant growth (Hepler, 2005). However, the negative effect was not apparent for the

soils containing low and medium amounts of Ca (based on Hazelton & Murphy (2019)). Thus, the amount of Ca was likely indicative of calcareous soils that have high representation in many grasslands worldwide, particularly those located in arid and semi-arid areas (FAO, 2020). Many of the sites included in the current study had Ca-rich soils and high soil pH but did not sustain high productivity. This is possibly due to low water-holding capacity/high infiltration rate, poor structure, low availability of some nutrients (such as P) and/or decreased plant uptake of soil micronutrients, which are all typical for calcareous soils (FAO, 2020). For instance, soils such as calcareous rendzinas can be rich in clay and humus with a high amount of Ca but can still be relatively infertile due to nutrient imbalances and shallowness (Miechówka & Drewnik, 2018).

Bulk density is another important indicator of soil fertility (Bünemann *et al.*, 2018) as it affects soil aeration and water and nutrient movement through soil (Stirzaker *et al.*, 1996). In our study, AGB was positively associated with fine-soil BD because lower fine-soil BD was found in relatively shallow soils with higher amounts of soil gravel. This was confirmed in the analyses excluding the plots with shallow soils, where fine-soil BD was omitted from the best model while the other original variables were retained. Low BD thus served primarily as an indicator of shallow and gravelly soils with limited amount of water and nutrients available for plant growth. Moreover, BD has shown to be positively related to soil N mineralisation (Olde Venterink *et al.*, 2001b) indicating that the increase of AGB with increasing BD in our study may also be partly due to increased N availability.

Soil texture (% sand) had a strong overall effect on AGB with both indirect effects through its influence on SOM and Ca as well as a direct effect. Along with precipitation patterns, soil texture is the most important determinant of soil water availability (Sala *et al.*, 1988), where the decreasing sizes of soil particles generally indicate increased water holding capacity. Sala et al. (1988) showed that the influence of soil texture (as a proxy for water holding capacity) on biomass production in grasslands of central USA depended on the level of precipitation, but here we found no interaction between texture or other determinants of soil water availability (such as SOM and BD) and precipitation, suggesting that precipitation does not affect the relationship between water holding capacity and AGB production in our dataset. This is also apparent from the local-scale patterns where, in general, there was a consistent relationship between indicators of soil water holding capacity and AGB across contrasting climates.

Finally, although this study included numerous different soil properties, there was of 48% of the unexplained variation in AGB on the local scale and 22% on the global scale. This unexplained variation might be attributed to the influence of biotic and/or other important factors that were not assessed in this study or it could be a result of differing past management practices which may have caused additional variation in biomass productivity estimates.

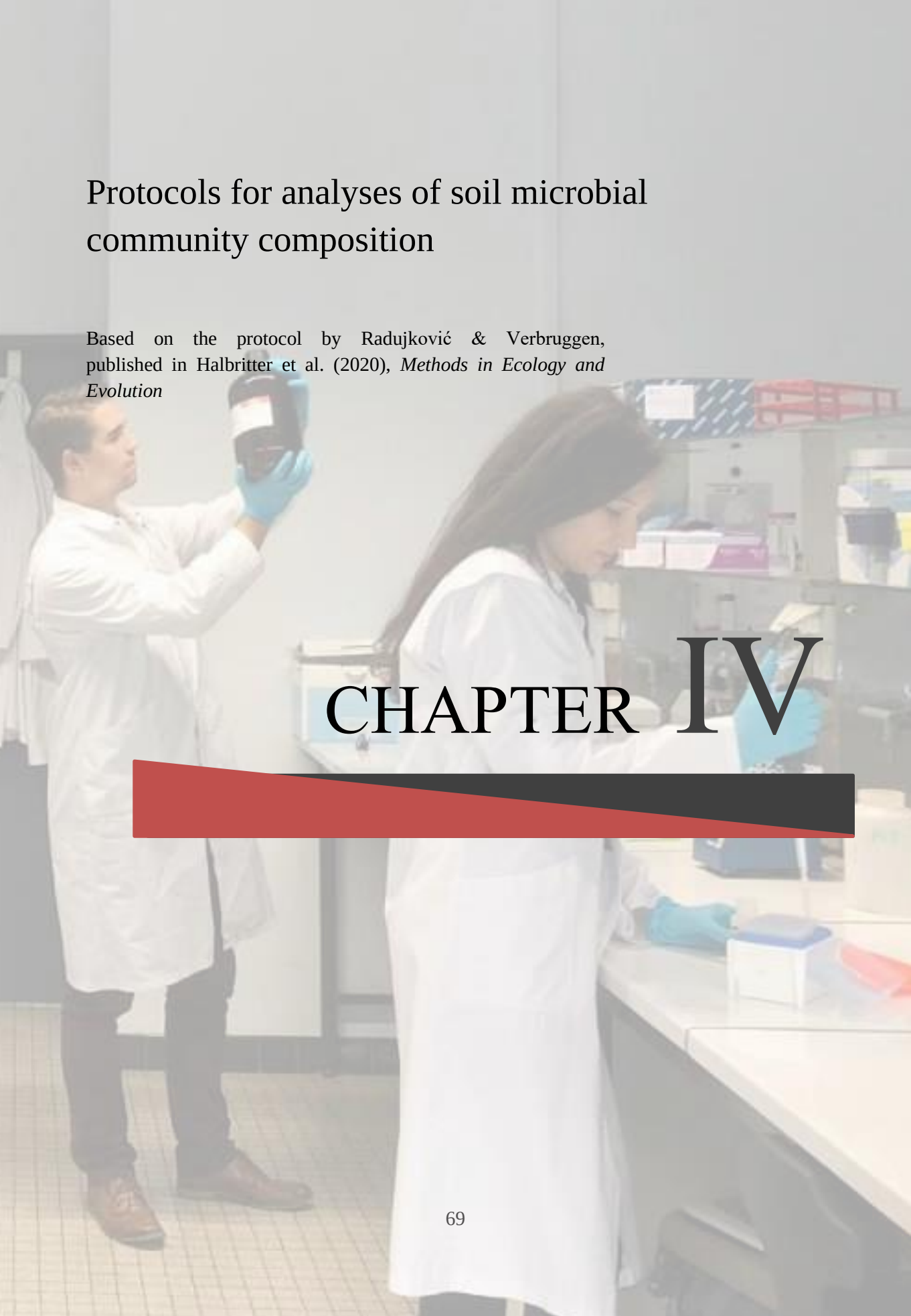
3.6 Conclusion

This study demonstrated that soil properties are critical predictors of variation in both global-scale and local-scale aboveground grassland productivity. Soil organic matter, soil texture and indicators of soil shallowness and calcareous soils, all of which are important determinants of the amount of nutrients and water that can be retained in the soil, explained up to 70% of the variation in aboveground biomass production. On the global scale, these soil factors were stronger predictors of aboveground biomass than commonly used climatic factors.

We argue that modelling the effects of environmental changes on primary productivity, as one of the keystone ecosystem processes in grasslands, can hugely benefit from incorporating the main indicators of soil fertility and water holding capacity reported here. Furthermore, we stress the importance of measuring these soil properties in grasslands as well as other ecosystems such that their role can be fully explored in future synthesis efforts.

Protocols for analyses of soil microbial community composition

Based on the protocol by Radujković & Verbruggen, published in Halbritter et al. (2020), *Methods in Ecology and Evolution*



CHAPTER IV

CHAPTER IV

Protocol for analyses of soil microbial community composition

Based on the protocol by Radujković & Verbruggen, published in Halbritter et al. (2020), *Methods in Ecology and Evolution*

4.1 Why to measure?

The microbial community composition represents the number and relative abundance of microbial taxa in a given system. This measure provides insight into the diversity and variability of the relative abundances of microbial taxa and thus aspects of their community dynamics. Moreover, changes in overall soil microbial community composition may point to corresponding changes in the various processes in which these communities are involved (Zogg *et al.,* 1997; Balser & Firestone, 2005; Strickland *et al.,* 2009). Microbes in soil are essential for the decomposition of organic matter (Allison & Martiny, 2008), they can play a key role in long-term carbon storage (Clemmensen *et al.*, 2013), and they are important drivers of biogeochemical cycling processes, including carbon and nitrogen cycling (Prosser *et al.*, 2007; Falkowski *et al.*, 2008). Specific microbial functional groups (such as nitrifying bacteria, mycorrhizal fungi, plant parasites) may affect ecosystem functioning by altering nutrient availability or plant productivity (van der Heijden *et al.*, 2008). Mycorrhizal fungi, for example, are associated with up to 90% of terrestrial plants (Smith & Read, 2008) and they have important effects on plant productivity (Wilson *et al.*, 2016; Yang *et al.*, 2016) and carbon dynamics in soil. It has been demonstrated that changes in different environmental factors (e.g. precipitation, CO_2, temperature, nutrient concertation) can cause shifts in microbial community composition (Zogg *et al.*, 1997; Nemergut *et al.*, 2008; Castro *et al.*, 2010). These changes may, directly or indirectly, affect important ecosystem processes (e.g. carbon cycling), thereby mediating the feedback responses to global change (Davidson & Janssens, 2006; Pold & DeAngelis, 2013).

4.2 What and how to measure?

4.2.1 Gold standard

With the development of high-throughput DNA sequencing techniques, the composition of soil microbial communities can be studied in more detail at a lower cost than using traditional culture-dependent approaches (Shokralla *et al.*, 2012). High-throughput molecular identification of microbial communities requires the isolation of nucleic acids from environmental samples, followed by DNA amplification using primers (small manufactured sections of DNA) that bind specifically to phylogenetically conserved regions of genes, which flank so-called barcode markers (Winsley *et al.*, 2012). The accuracy of these analyses is strongly dependent on the choice of primers (Klindworth *et al.*, 2013). Genes encoding components of the nuclear ribosomal units (small subunit, SSU; large subunit, LSU; internal transcribed spacer, ITS) are by far the most commonly used genetic markers for taxonomic identification of microorganisms *(Lindahl et al.,* 2013). Current high-throughput sequencing techniques allow the simultaneous sequencing of millions of reads (Bartram *et al.*, 2011), which are then typically clustered into operational taxonomic units (OTUs; typically at 97% sequence similarity) and assigned to taxonomic/functional groups using various bioinformatical tools and reference databases. These methods are semi-quantitative given that they provide information of relative abundances of taxa rather than absolute abundances, however, methods have been developed where absolute abundances of bacteria can be estimated by incorporation of internal standards into DNA pools (Smets *et al.*, 2016; Harrison *et al.*, 2020). The changes in microbial community composition exposed to certain climate treatments (e.g. warming, drought) compared to control communities, can be statistically assessed based on the differences in the number and (relative) abundance of OTUs between these communities and/or changes in the (relative) abundance of taxonomic/functional groups.

4.2.2 Bronze standard

Phospholipid fatty acid analysis (PLFA) is another culture-independent method that is commonly used to assess the changes in microbial community composition. It has been demonstrated that PLFA analyses and genetic sequencing can detect similar overall patterns in bacterial community composition (Orwin *et al.*, 2010). However, compared to genetic sequencing, PLFA has a very limited taxonomic resolution, especially for groups other than bacteria, but it can provide quantitative information about microbial biomass (Brewer *et al.*, 2015). It can thus be used in cases when quantitative shifts in both biomass and broad functional groups (fungi, gram-positive v. gram-negative bacteria) are to be delineated. For a detailed

protocol and possible applications of PLFA see Frostegård et al. (1993) and Frostegård et al. (2011), respectively.

4.2.3 Soil sampling and storage

Soil samples are collected using soil corers, usually at depths of 0–5 cm and/or 5–10 cm (e.g. Rinnan *et al.,* 2007; Kuffner *et al.,* 2012; Hayden *et al.,* 2012). The corers must be cleaned between the samples in order to avoid cross-contamination. When collecting samples for fungal analysis, it should be borne in mind that fungi can have very long mycelia and thus it is recommended to keep a minimum distance of 3 m between different samples when independence is required for statistical analysis (Lindahl *et al.,* 2013). Typically, a few soil samples are taken per plot and pooled into a composite sample (e.g. for a good representation of a plot, samples can be taken in four corners and the centre). Depending on the study question, samples can be taken one time only (e.g. in the peak of the growing season, if the aim is to examine the effect of treatments at the peak of vegetation growth) or multiple times in the same plot (e.g. if the aim is to examine inter- or intra-annual changes in community composition). The samples can be stored in sterile plastic ziplock bags. After sampling, the soil is sieved (2 mm mesh size is a standard in soil science), taking care to prevent contamination. The samples should be kept in a cold place and processed as soon as possible to avoid the degradation of DNA and microbial growth (Rochelle *et al.,* 1994). Over longer periods, samples can be optimally stored by freezing at -20 °C or -80 °C (Song *et al.,* 2016). Alternatively, they can be freeze-dried (Lindahl *et al.,* 2013) or stored in pure ethanol (Hale *et al.,* 2015) or commercially available preservation solutions.

4.2.4 DNA extraction

Most extraction methods are based on direct cell lysis which generally provides high yields of DNA with relatively short processing times (Robe *et al.,* 2003). Commercially available soil DNA extraction kits provide detailed protocols for extraction procedures. Because of the typical low sample size for extraction (0.25–0.5 g dry weight), care should be taken to thoroughly homogenise material for subsampling, or isolate DNA from multiple technical replicates. Ideally, extraction should yield high and uniform amounts of DNA and minimal concentrations of amplification inhibitors (Lindahl *et al.,* 2013). DNA yield can be assessed and concentrations can be adjusted through dilution. The same DNA extraction protocol should be used for all samples (Tedersoo *et al.,* 2010) ensuring that potential extraction-related biases are equally distributed across all samples.

The procedures described next are sometimes outsourced to a commercial laboratory (even including taxonomic annotation of obtained sequences) or can be performed in-house when facilities are available.

4.2.5 DNA amplification – PCR

Following extraction, DNA is amplified using primers that target a barcode marker region which is conserved within a particular microbial group (prokaryotes, eukaryotes, fungi, arbuscular mycorrhizal fungi), but includes variable regions that allow the distinction at the phylogenetic level of interest (Lindahl *et al.*, 2013). The primers also include artificial barcode sequences that allow identification of different samples after sequencing, or these barcodes are added in a second step. Amplification of the marker is accomplished by the successful binding of the two primers to the flanking sections and generating copies of it through a "polymerase chain reaction" (PCR). In order to assess the variation resulting from stochastic processes during laboratory work, replicate PCR reactions can be performed using independently obtained DNA extractions from the same sample (Kauserud *et al.*, 2012). PCR conditions (see e.g. Bartram *et al.*, 2011; Klindworth *et al.*, 2013; Zhang *et al.*, 2016) need to be optimised to the marker region and lab conditions (e.g. enzymes and thermal cycler), where the annealing temperature, in particular, deserves attention. Optimal annealing temperatures range between 45 and 68 °C depending on primer sequence and are, as a rule of thumb, set at 5 °C below the calculated temperature of the lowest primer melting point (T_m) (Roux, 2009).

To assess the success of a PCR, the products are visualised on an agarose gel where the presence and length of a product can be determined. If the annealing temperature is too low (primers do not anneal specifically to the target region) there will be more bands visible on the gel (more than expected based on natural length variation of the marker); if it is too high (primers do not anneal to target region at all) there will be no bands on the gel. The optimal annealing temperature for a particular primer pair can be determined by gradually increasing the annealing temperature (gradient PCR). PCRs can also fail due to different inhibitors present in the starting template. A 5–100-fold dilution of the template may dilute out the inhibitor (Roux, 2009). Other possible solutions in case of PCR failure include re-extraction, re-amplification, ethanol precipitation, changing the number of PCR cycles, or adding stabilising proteins such as bovine serum albumin (BSA) (see Roux, 2009 for more details on optimisation of PCR process).

4.2.6 Primer choice

There are multiple valid reasons to choose one primer-pair over another for a particular group of microbes. Main reasons are i) the sequencing technology used: some instruments (e.g.

Illumina MiSeq) work optimally with DNA sequences between 250–500 base pairs (bp) in length, while others (e.g. PacBio) can sequence whole DNA strands with thousands of bp; ii) sequence variability: ideally there should be a so-called "barcode-gap" (Schoch *et al.*, 2012) making it easy to delineate within-species vs between-species variability, however, this varies between taxonomic groups and markers and so the choice will often be a trade-off where higher quality data for one group will come at a cost of another group; and iii) historical reasons will cause a marker for a group of interest to have a much better representation in databases (e.g. 16S/18S rRNA for many microbial groups), which means that even when in principle other regions would be more suitable, having a well-filled database to compare against will improve the quality of the eventual data.

Bacteria. The 16S rRNA gene (encodes SSU in prokaryotes) has been by far the most commonly used genetic marker for analyses of bacterial communities (Klindworth *et al.*, 2013) for a number of reasons: it is present in all bacteria; it contains both highly conserved regions and hypervariable regions; and it is sufficiently long (1,500 bp) for bioinformatic purposes (Janda & Abbott, 2007). 515F-806R primer (Caporaso *et al.*, 2011) pair is often used to amplify bacterial V4 region of 16s rRNA gene (Chapter VI), while the combination of Bakt_341F and Bakt_805R primers (Herlemann *et al.*, 2011) can be used to amplify both V3 and V4 variable regions. The latter primer set was evaluated by Klindworth et al. (2013) as one of the most efficient in amplifying a wide range of bacterial phyla.

Fungi. Molecular analyses of fungal communities mainly rely on amplification of the ITS region (spanning the ITS1, 5.8S, and ITS2 regions), which was selected as the universal genetic barcode for fungi (Schoch *et al.*, 2012). However, whether the ITS1, ITS2, or a combination of these two regions is better suited for characterisation of fungal communities is still under debate (Blaalid *et al.*, 2013). The ITS1 region is frequently amplified using the combination of ITS1f and ITS2 primers (Op De Beeck *et al.*, 2014; Smith & Peay, 2014) (Chapter V and Chapter VI). fITS7, gITS7, and fITS9 primers target binding sites in the 5.8S region and together with the ITS4 primer, they can be used to amplify the ITS2 region (Ihrmark *et al.*, 2012). The combination of ITS1f and ITS4 primers span both ITS regions together with 5.8S region (Smith & Peay, 2014).

Arbuscular mycorrhizal fungi (AMF) – Glomeromycota. For AMF analysis, the most commonly used marker region is SSU (18S rDNA in eukaryotes), followed by LSU(28S rDNA) and ITS rDNA region (Öpik *et al.*, 2014). The SSU rDNA region alone is not suitable for identification of species (Öpik *et al.*, 2014), but in the cases when species resolution is not the primary goal, primers that target SSU region, for example, AML1 and ALM2, designed by Lee et al. (2008), can provide useful information regarding the overall AMF community composition. Primer set SSUmAf–LSUmAr (1800 bp) and SSUmCf–LSUmBr (1500 bp) developed by Krüger et al. (2009) spans a fragment covering the partial SSU, the entire ITS, and the partial LSU rDNA region. This combination of primers enables the detection of

additional AMF, but the sequences are too long for some high-throughput sequencing and alternative sequencing methods must be used (Schlaeppi *et al.,* 2016).

Protists. A comprehensive overview of different SSU primers designed to target protists is provided by Adl et al. (2014). The combination of primers TAReuk454FWD1 and TAReukREV3 (Stoeck *et al.,* 2010) that targets the V4 region of SSU, can be used for the detection of a wide range of eukaryotic lineages (Mahé *et al.,* 2017). A recently developed combination of primers (ITS3 primer mixes, ITS4ngs) described in Tedersoo et al. (2015) that target ITS2 region can be used to characterise certain protist groups: Cercozoa, Ciliophora, and Chlorophyta, as well as soil animals (Acari, Nematoda, Collembola, Rotifera, Annelida) which are thought to be the most abundant and species-rich eukaryotic taxa in soil (Tedersoo *et al.,* 2015). Given the paraphyletic nature of protists (spanning the entire eukaryotic phylogenetic tree), no primers specifically targeting this group as a whole can be designed. For this reason, samples containing a high concentration of plant, fungal, or animal DNA, such as when one aims to elucidate protists that are part of their "microbiomes", are at risk of primarily generating non-target sequences.

4.2.7 Library preparation and sequencing

Following purification from PCR artefacts (primers and primer-dimers), different samples with specific barcodes are equimolarly pooled into a single library ready for sequencing. The Illumina MiSeq platform (Illumina Inc; San Diego, CA, USA) is currently the most commonly used platform for high-throughput sequencing of environmental microbial samples (see Chapter V and Chapter VI). This platform enables sequencing of 200–550 bp-long paired-end reads (forward and reverse) which is, in most cases, enough to cover the entire marker region for different microbial groups. Longer reads can be sequenced using single-molecule real-time (SMRT) methodology (PacBio; Manlo Park, CA, USA).

4.2.8 Quality control and bioinformatics analyses

UPARSE (Edgar, 2013), QIIME (Caporaso *et al.,* 2010), and mothur (Schloss *et al.,* 2009) are some of the most commonly used bioinformatics pipelines that allow quality filtering and construction of OTUs from next-generation sequencing reads. The main result of these analyses is an OTU table (Figure 4.1). The downstream analyses (e.g. standardisation of read number through downsampling (Weiss *et al.,* 2017), calculation of alpha and beta diversity) can be performed using QIIME and mothur, but also in statistical programs such as R (e.g. using the ‘vegan’ or ‘phyloseq’ packages). Typically, OTU tables are used to create distance matrices, which include pairwise distances between the microbial communities of different samples

(Figure 4.1). It should be noted that for bacterial sequences, it is common to create phylogenetic trees and use phylogenetically informed distance metrics (i.e. UniFrac).

Taxonomic identification is performed by aligning sequences to the reference sequences (using BLAST or other methods implemented in UPARSE/QIIME/mothur) deposited in publicly available databases. An overview of different databases is given by Santamaria *et al.* (2012). For instance, the Greengenes database contains a collection of bacterial 16s rDNA sequences (DeSantis *et al.,* 2006), UNITE is a comprehensive reference database for fungal ITS sequences (Abarenkov *et al.,* 2010), and PR^2 (Protist Ribosomal Reference) database is suitable for annotation of protist SSU sequences (Guillou *et al.*, 2013). Other databases, such as Silva (Quast *et al.,* 2013) and Ribosomal Database Project (Cole *et al.,* 2014) contain collections of SSU and LSU sequences for various groups of prokaryotic and eukaryotic microorganisms. Following the taxonomic assignment, fungal OTUs can also be assigned to different functional categories (i.e. saprotrophic fungi, white rot decomposers, yeasts, plant pathogens, mycoparasites, animal parasites, arbuscular mycorrhizal fungi – AMF, ectomycorrhizal fungi – EcM) by matching their genus/family level with the known lifestyles (e.g. as in Tedersoo *et al.,* 2014) using specialised tools such as FUNGuild (Nguyen *et al.,* 2016).

4.2.9 Statistical analyses

Statistical analyses on distance matrices or OTU tables can be performed using various multivariate types of analyses such as PERMANOVA, ANOSIM, and ordination methods (e.g. PCoA, (G)NMDS, CCA, an example is depicted in Figure 4.1 and Figure 5.1). Moreover, network analyses can be performed to investigate potential interactions between species (see Chapter V). Although correlation networks do not necessarily represent the real biological interactions, they can provide valuable insights in species co-occurrence patterns and elucidate the mechanisms driving their community assembly and functioning (Barberán *et al.,* 2012). For instance, examining the architecture of ecological networks (e.g. the number and strength of connections, network connectedness, network modularity) and identifying the taxa that are key players in these networks can be used to predict community stability (Thebault & Fontaine, 2010) examine functional redundancy (Banerjee *et al.*, 2016) and reveal the taxa driving microbiome functioning (Banerjee *et al.,* 2018). Numerous network analysis methods have been developed and used in different studies: from simple correlation-based methods (e.g. in Encinas-Viso et al. (2016) and de Vries et al. (2018)) to more complex methods such as hierarchical modelling of species communities (Ovaskainen *et al.*, 2017) and extended local similarity analysis (Xia *et al.*, 2011).

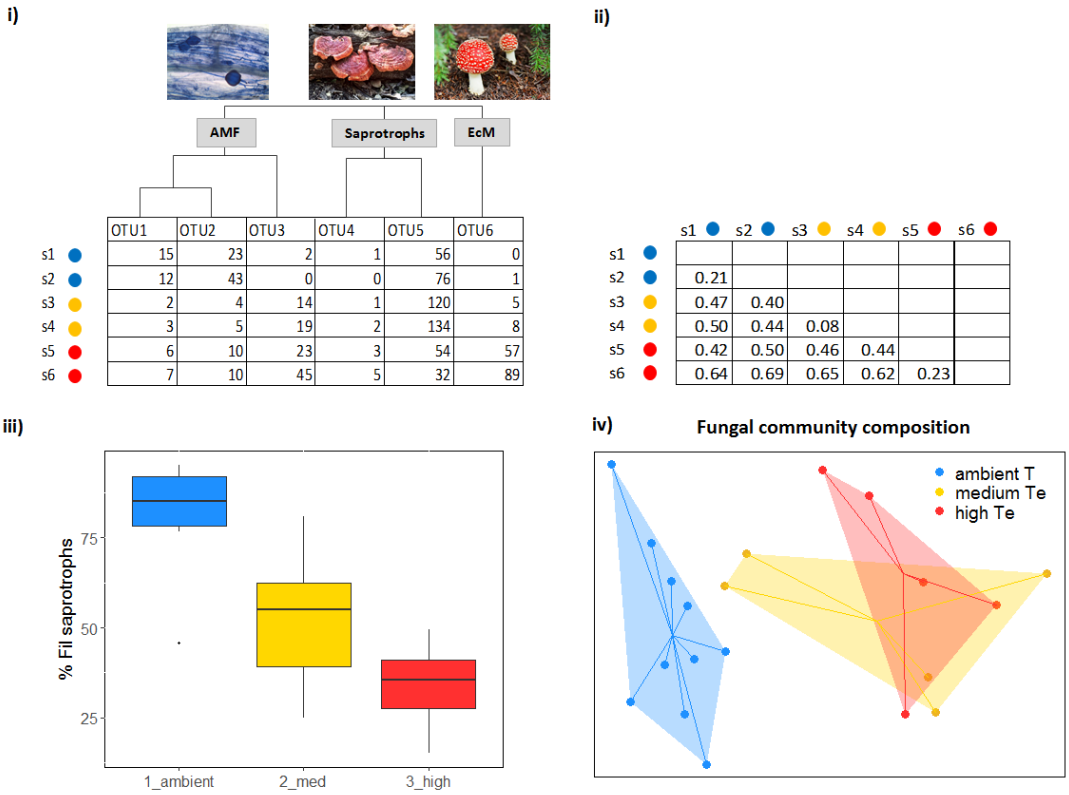


Figure 4.1 Main output of microbial genetic sequencing. Simplified representation of an OTU table containing the number of fungal OTUs in soil samples (s1–6) exposed to 3 different treatments (depicted in different colours). Different OTUs are assigned to taxonomic and/or functional groups by comparing them against a database. ii) Based on the OTU table, it is possible to quantify the dissimilarities between the samples and summarise them in a distance matrix. Lower panels show the differences between soil fungal communities exposed to different intensities of natural warming, based on a subset of actual data from the ForHot natural experimental site (microbial data: Radujković et al. (2018); ForHot experiment: Sigurdsson et al. (2016). iii) The relative abundance (% of the total amount of sequences in a sample) of filamentous saprotrophic fungi exposed to different intensities of warming and iv) the multidimensional ordination of samples based on Bray-Curtis distances. Points and the corresponding polygons are coloured according to temperature elevations (Te): blue – ambient temperatures; orange – medium temperature elevation; (+3 °C to +5 °C); red – high temperature elevation (+7 °C to +11 °C).

4.3 Special cases, emerging issues and challenges

The methods of molecular analysis of microbial communities are evolving very rapidly with the development of new technologies. Previously commonly used 454 pyrosequencing is now almost entirely replaced by Illumina sequencing by synthesis. SMRT technology, such as PacBio, is now being increasingly used since it can provide longer reads (albeit with high error rates). The choice of sequencing platform is therefore currently a trade-off between the quality of the produced reads and the maximum length of the reads (Kennedy *et al.*, 2018), but these or other platforms will likely become cost-efficient at low error rates in the near future.

Recently, there has been a lot of discussion regarding the common practices for bioinformatics analysis of sequencing data. The conventional approach is to perform the clustering of OTUs, usually based on 97% similarity. However, this approach has been challenged and it has been proposed that instead of OTU clustering, amplicon sequence variants (ASVs) should be used. It is argued that, compared to OTUs, ASVs represent a biological reality independent of the data analysis, they have a better taxonomic resolution and they can be validly compared across different studies (Callahan *et al.*, 2017). However, ASVs are highly sensitive to the quality of the data and this approach could be problematic for downstream analysis due to significantly increased diversity. While OTU clustering still remains the most common approach in ecological studies, it would be useful to also report the sequence variants in order to enable the effective comparison between different studies.

Mechanisms driving the development of soil fungal community composition in restored heathlands

Based on the paper by Radujković et al. (2020), *New Phytologist*.

CHAPTER V

CHAPTER V

Mechanisms driving the development of soil fungal community composition in restored heathlands

Based on the paper by Radujković et al. (2020), *New Phytologist.*

5.1 Abstract

Dispersal limitation, biotic interactions and environmental filters interact to drive plant and fungal community assembly, but their combined effects are rarely investigated. This study examines how different heathland plant and fungal colonization scenarios realized via three biotic treatments - addition of mature heathland derived sod, addition of hay and no additions - affect soil fungal community development over six years along a manipulated pH gradient in a large-scale experiment starting from an agricultural, topsoil removed state. Our results show that both biotic and abiotic (pH) treatments had a persistent influence on the development of fungal communities, but that sod additions diminished the effect of abiotic treatments through time. Analysis of correlation networks between soil fungi and plants suggests that the reduced effect of pH in the sod treatment, where both soil and plant propagules were added, might be due to plant-fungal interactions since the sod additions caused stronger, more specific, and more consistent connections compared to no addition treatment. Based on these results, we suggest that the initial availability of heathland fungal and plant taxa, that reinforce each other, can significantly steer further fungal community development to an alternative configuration, overriding otherwise prominent effect of abiotic (pH) conditions.

5.2 Introduction

The incidence and abundance of local above- and below-ground species in an ecosystem are dependent on three main processes or 'filters': i) dispersal constraints ii) environmental (habitat) filters, and iii) biotic interactions (Belyea & Lancaster 1999; Lortie *et al.* 2004). Contrary to the traditional view that biotic interactions only operate after environmental filtering has taken place (Belyea & Lancaster, 1999; Raevel *et al.*, 2013), it is increasingly recognized that biotic interactions can significantly mediate species' responses to the environment and therefore determine the strength and extent of this filter (Wisz *et al.*, 2013; Cadotte & Tucker, 2017; Aguilar-Trigueros *et al.*, 2017). The same is true for dispersal, where the timing of arrival may dictate which biotic interactions prevail, with a cascading effect on

future community assembly through priority effects (Fukami *et al.*, 2005; Fukami, 2015). Understanding and predicting the development of communities thus requires knowledge of how these three processes act in concert (Wisz *et al.*, 2013).

Numerous studies have demonstrated that plant-soil interactions (particularly those between soil fungi and plants) are key biotic interactions that can shape above- and below-ground communities (Kardol *et al.*, 2006; Smith & Read, 2008; Wagg *et al.*, 2014; van der Putten, 2017). For instance, they have been shown to be major drivers of plant community composition patterns in restored tallgrass prairies (Bauer *et al.*, 2015) and pristine tropical forests (Mangan *et al.*, 2010). Moreover, manipulation through soil inoculation promoted the development of heathland and grassland systems, possibly through positive feedbacks among plants and their associated soil biota (Wubs *et al.*, 2016, 2019, van der Bij *et al.*, 2018). Studies investigating plant-soil interactions have particularly emphasised the importance of mycorrhizal fungi as mediators between below- and aboveground communities (Bauer *et al.*, 2015) showing, e.g. that the presence and identity of mycorrhizal fungi determined whether late or early successional plant species came to dominate in a prairie restoration experiment (Koziol & Bever, 2017). Characterization of plant-soil interactions and the mechanisms by which they steer community assembly has been very challenging, particularly in field conditions, considering the myriad of interactions between plant and soil organisms (Toju *et al.*, 2018). Nevertheless, incorporating real-life complexity is crucial to accurately characterize the influence of the environment on plant-soil interactions (Lekberg *et al.*, 2018).

The complexity of plant-soil interaction can be captured by network approaches since they incorporate the whole community rather than limited number of preselected taxa (Ramirez *et al.*, 2018; *Toju et al.*, 2018). Several recent studies have utilized the network approach to examine putative biotic interactions (Banerjee *et al.*, 2016; Encinas-Viso *et al.*, 2016; Tylianakis *et al.*, 2018; de Vries *et al.*, 2018), showing for instance that the architecture of ecological networks is related to community stability (Thebault & Fontaine, 2010) and that hubs of highly connected soil microbes mediate interactions between plants and microbes (Agler *et al.*, 2016). Characterizing plant-soil network structure (e.g. the number and strength of connections) and identifying the taxa that are key players in these networks can thus help us understand how plant-soil interactions influence community development.

The present study examines the importance of plant-soil interactions for soil fungal community development in a large-scale heathland restoration experiment. Heathlands are species-poor systems thriving on nutrient-poor, acidic soils, with high dominance of ericaceous plants and associated ericoid mycorrhizal (ERM) fungi (Gimingham, 1989; Webb, 2008). Therefore, they represent a relatively tractable model system to explore typically complex plant-fungal interactions. In our study system, the upper soil layer from an ex-arable field was removed and different plots were subjected to three biotic addition treatments crossed with three pH manipulation treatments. Biotic treatments represent different dispersal scenarios (different timing of colonization): an initial presence of both soil and plant propagules derived from a heathland system, an initial presence of primarily plant propagules only, or "natural" colonization through gradual dispersal in the control. The abiotic – pH – treatments created a gradient with the potential to act as an environmental filter within each of the biotic treatments. pH is known to strongly influence the success of heathland restoration (Marrs *et al.*, 1998) since it affects the germination of heathland plants and the development of their interactions with ERM fungi (Díaz *et al.*, 2008). By censusing the plant and soil fungal community composition through time, we followed the development of plant-fungal correlation networks under different treatments.

This experimental setup, therefore, allowed us to investigate the combined effect of three different mechanisms (timing of colonization, abiotic conditions, biotic interactions) on the development of soil fungal communities over multiple years. We hypothesized that (1) initial biotic manipulations had a lasting effect on fungal community development, as evidenced by significant differences in community composition at the end of the experiment; (2) that the effect of different biotic treatments and abiotic conditions were contingent on each other, as evidenced by interactions between biotic and abiotic treatments and variation in within-group dispersions between biotic treatments. Furthermore, we explored (3) whether and in what way the interactions between fungi or between plants and fungi may have contributed to fungal community development through co-occurrence and network analyses. Together, these approaches shed light on the relative importance and interaction between the ecological filters operating in heathland fungal community assembly.

5.3 Materials and methods

5.3.1 Study sites and sampling

Study sites were located at Dwingelderveld National Park (lat: 52.7810, long: 6.3709, alt: 10 m) in the Netherlands. The study area had previously been used for intensive agriculture. In 2011, the top-soil layer (30-40 cm) was removed to eliminate the excess of nutrients and other legacies (e.g. seed bank) of agricultural land as an attempt to restore a typically nutrient-poor heathland ecosystem. Subsequently, 27 large plots (15m x 15m) were established with nine different treatments, three biotic treatments crossed with three abiotic treatments, each in three replicates in a randomized block design. The biotic treatments included biotic control = no additions, addition of hay material or addition of sod material, from well-developed heathlands. The abiotic treatments consisted of: no additions = abiotic control, addition of dolomite $CaMg(CO_3)_2$ = liming, or addition of elemental S = acidification. The donor heathland sites for sod and hay material was a dry mature heathland dominated by *Calluna vulgaris* L, located 100 – 200 m from the experimental site. For all treatments, the material was added in late autumn 2011 (first abiotic then biotic additions), except for hay material which was not available in late autumn and was added in early autumn 2012. For the hay / sod treatment, 1 m^2 of fresh heathland hay / sod material (the vegetation and soil down to 5-6 cm depth) was added per 2 m^2 and 15 m^2 of experimental site, respectively. For the liming / acidification treatment 2 t of dolomite / 1.5 t of elemental sulphur were added per hectare of experimental site, respectively. None of these treatments significantly altered the amount of organic matter in the soil, and except for the abiotic treatments, none altered the soil chemistry (Van der Bij *et al.*, 2018), including pH (Figure S5.1, Table S5.7). Initially, liming increased soil pH by approximately 0.3-0.5 units and acidification decreased it by 0.3 units (averaged across biotic treatments). Six years after the additions, soil pH under different abiotic treatments still differed significantly (mean pH_{2017}: control = 4.7, liming = 5.2, acidification = 4.5) (Figure S5.1).

Every year from 2012 to 2017, plant cover in the centre 10 x10 m of each plot was estimated according to the Tansley scale, and three soil samples were taken at a depth of 0-5 cm from each of the 27 plots and pooled into one composite sample per plot for microbial analysis and measurements of soil pH. In addition, three soil samples were taken in three different well-developed (reference) heathland plots in the same area in 2017 and pooled in one sample per reference. Samples taken in the first five years were immediately air-dried, homogenized and kept under cool, dark and dry storage conditions before the DNA was isolated in 2017, while the samples from 2017 were immediately frozen, shortly after which DNA was isolated. Further tests indicated that storage conditions and storage time did not affect perceived variation in fungal community composition. See Supplementary material (Annex S5.1, Figure S5.2) for more details on additional tests and analyses concerning sample preservation.

5.3.2 Sample preparation and sequencing

DNA was isolated from 0.25-0.35 g of soil using the DNeasy PowerSoil Kit according to the manufacturer's protocol (Qiagen, Venlo, the Netherlands). The ITS1 region was amplified using fungal primers ITS1f (Gardes & Bruns, 1993) and ITS2 (White *et al.*, 1990), modified according to (Smith & Peay, 2014). In the first PCR, primers were amended with Illumina Nextera labels (Illumina Inc; San Diego, CA, USA). Each 25 µl reaction mixture contained 2 µl of the sample, 0.5 µM of each forward and reverse primer, 1X PCR buffer, 200 µM dNTPs and 1 U Phusion High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA, USA). PCR conditions were as follows: initial denaturation at 98°C for 60 s, followed by 35 cycles of: denaturation at 98°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 30 s; and an additional extension of 72°C for 10 min. A second PCR was performed using dual barcoded primers with Illumina adapters (2.5 µl of 50 x diluted PCR products template and 0.1 µM of each primer). The conditions were: 98°C for 60 s, 12 cycles: at 98°C for 10 s, 63°C for 30 s, 72°C for 30 s; and 72°C for 5 min. PCR products were run on an agarose gel to confirm successful PCR amplification and successful amplicons were normalized and purified from primers and primer-dimers using the SequalPrep Normalization Plate Kit (ThermoFisher Scientific). Samples were then pooled into a single library, and subjected to a gel extraction using QIAquick Gel Extraction Kit (Qiagen, Venlo, the Netherlands). The library was quantified with qPCR (KAPA Library Quantification Kits, Kapa Biosystems, Wilmington, MA, USA) and sequenced on the Illumina MiSeq platform (Illumina Inc; San Diego, CA, USA) with 300 cycles for forward and reverse reads. Several negative controls and technical replicates were also sequenced in order to test the reproducibility of sample preparation and the sequencing procedure (Figure S5.3). The raw sequences were deposited in SRA-NCBI database under the accession number: PRJNA566105.

5.3.3 Quality filtering and bioinformatics analyses

Fungal sequences were analysed using the USEARCH (v8.1.1861) and VSEARCH (Rognes *et al.*, 2016) software following the UPARSE pipeline (Edgar, 2013). After trimming to 250 bp the paired-end reads were merged and primers were removed. This trim length was chosen because it was the optimal length for merging pairs by removing the low-quality bases at the end. Merged sequences were quality filtered using expected number of errors (E) as a measure of read quality, as implemented in UPARSE. We imposed a relatively stringent criterion of $E_{max} = 0.5$, keeping the reads that have a maximum 50% chance to contain one erroneous base (Edgar & Flyvbjerg, 2015), leaving 3.01 M sequences. Following singleton removal, the sequences were clustered into OTUs (operational taxonomic units) based on 97% similarity using the UPARSE-OTU algorithm (Edgar, 2013) which automatically detects and filters out chimeras with high efficiency. All original reads were mapped to the OTUs with an identity

threshold of 0.97, yielding an OTU table with a total of 2,192 OTUs and 3.5 M reads. Using all original reads does not compromise quality of OTUs but allows sequences erroneously labelled as low-quality to be counted. Further steps were performed using R software (R Core Team, 2015). The number of reads per sample was rarefied to 1,275. This rarefaction depth was chosen because it included almost all samples (except for four which were omitted), and although it does not represent the entire diversity, rarefaction curves showed that the number of taxa was levelling off for most samples at this depth. We also calculated Chao coverage (*entropart* package (Marcon & Herault, 2015)) as an indication of the amount of unsampled taxa, which was the same for different biotic treatments (Figure S5.4, Table S5.1). Representative OTUs were aligned to the fungal sequences in the UNITE database (Kõljalg *et al.*, 2005) (release date 10.10.2017), using the NCBI's BLAST algorithm with default settings. OTUs were retained and assigned to particular taxa if they had a minimum alignment length of 75 bp. and a maximum E-value of 10^{-36} (as in Waring *et al.*, (2016)).

5.3.4 Statistical analyses

The differences in fungal community composition were examined with PERMANOVA analysis (Anderson, 2001) using *adonis* function in *vegan* (Oksanen & *et al.*, 2018), based on Bray-Curtis (BC) distances and visualized using Nonmetric multidimensional scaling (NMDS, *metaMDS* in *vegan*). First, PERMANOVA analysis was performed 1) on the entire dataset using year as a continuous variable and plot as strata to assess the effect of time and 2) using biotic and abiotic treatments and their interaction as explanatory variables and year as strata. In addition, a separate PERMANOVA analysis was performed for the last year of the experiment to assess whether the effect of different biotic and abiotic treatments was present at the end of the experiment. Data was $\log_e$-transformed prior to analyses to reduce the impact of abundant taxa (Anderson *et al.*, 2006) which are typically overestimated due to the exponential nature of PCR, but the results were similar using different types of transformations (Table S5.2). To assess general trends in fungal OTU richness, the effect of time and different biotic and abiotic treatments (as well as their interactions) on fungal OTU richness was tested using the *lmer* function from *lme4* package with plot as a random effect.

Multivariate dispersion (distances from group-centroids) within different biotic treatments for each year was calculated using the *betadisper* function in the *vegan* package and by calculating the mean distance between each pair of samples within a treatment (using the actual BC distances between samples). Based on the results from *betadisper*, a posthoc test was performed to examine if dispersion was significantly different between different biotic treatments and P values were corrected for multiple testing (Benjamini & Hochberg, 1995). The rationale for this analysis is to explore whether there is fungal community convergence within biotic treatments (i.e. if the dispersion within treatment decreases), which we take as evidence that the relative influence of abiotics or random variation decreases. We also calculated the BC

distances contrasting biotic treatments (sod vs control, hay vs control and sod vs hay) to visualize change through time.

We used dissimilarity overlap curve (DOC) analysis (Bashan *et al.*, 2016) to test whether the interactions between fungal taxa were important drivers of fungal community composition in different biotic treatments across all the years. Bashan *et al.* (2016) demonstrated that communities with high overlap also become increasingly similar in abundance-patterns (so reduced dissimilarity) when their constituent taxa interact predictably. Following Bashan *et al.* (2016) and Verbruggen *et al.* (2018), a significant negative relationship between community overlap and dissimilarity of the 50% of data points with highest overlap was here taken as support that interactions between fungal taxa substantially influence fungal community composition. Null models were constructed to additionally confirm that no relationship was found in randomized data (see Bashan *et al.* (2016) for more details on the analysis).

DOC analysis was performed in MatLab v.9.0 (The MathWorks, Inc., Natick, Massachusetts, United States). All other analyses were performed in R (version 3.3.2) (R Core Team, 2015).

5.3.5 Network analysis

While numerous network analyses have been developed (see Chapter IV), due to the specific nature of our data, we followed a procedure that first calculates a general relationship between taxa based on the full dataset, and then estimates the extent to which this relationship is realized in each sample. By first calculating the relationship between taxa in the full dataset we circumvent the problem of few replicates for each treatment-time combination and the issue of high within-group variance of fungal abundances and low within-group variance of plant cover data which would otherwise be very difficult to correlate. This is done by assigning higher weights to 1) better fit and 2) higher relative abundance / percentage cover compared to all other occurrences of the two queried taxa. This procedure is detailed below.

First, 65 dominant fungal OTUs (containing a minimum of 500 reads across samples) and 25 dominant plant species (occurring in more than 8% of plots) were selected and the Pearson correlations between taxa were calculated. Rare taxa were removed to reduce the effect of zero occurrences, but more than 60% of total plant cover/fungal sequences for each treatment/year were included (Table S5.3). Correlations with Pearson r higher than 0.2 were further considered for the construction of correlation networks. We imposed this threshold as an initial filter against spurious correlations but set it low enough to account for inherent error due to low precision of actual plant cover estimates and noise due to random variation. A sensitivity analysis with different thresholds and different cut-offs of the number of OTUs and plant species showed that these alternative choices did not substantially influence overall network structure (Figure S5.5). Next, a simple linear regression between each pair of fungal OTUs and plant species was performed to estimate the study-wide slopes and intercepts using ordinary

least squares (OLS) regression. In order to estimate the realization of these relationships in different samples, the values for slopes and intercepts were then used to calculate the explained variation (*EV*) of the abundance of one taxon based on the abundance of the other for each sample in each year. More specifically, *EV* for a given pair of taxa (cases with double zeros were excluded) was calculated by subtracting the residual variation – *RV* (the difference between the actual abundance of a taxon (*y*) and the abundance predicted by the abundance of the other taxon (*x*) when using the slope (*a)* and intercept (*b)* as calculated above) from the total variation – *TV* (the difference in abundance of a taxon (*y*) and the mean abundance of that taxon ($\overline{y'}$) across all the data) (Equation 1). This value was then multiplied by an index calculated as the square root of the product between the abundance of each taxon in a pair per plot per year, as a fraction of their maximum abundance in the dataset (x' and y') to obtain EV' (Equation 2). EV' was used as an indicator of connection strength. This means that the higher the abundances of both taxa relative to their maximum abundance, the score gets a higher weight. The reasoning behind this is that under lower abundances, which are less variable, the scores would be inherently higher than the scores at higher abundances (due to the positive correlation between mean and variance). Finally, this calculation was performed for each year and obtained values were averaged: i) per biotic treatment and ii) per each combination of biotic and abiotic treatments. Negligibly low coefficients (< 0.001) and those lower than zero were set to zero.

$$EV = \frac{\mathrm{TV} - \mathrm{RV}}{\mathrm{TV}} = \frac{|y-\overline{y'}| - |y-(ax+b)|}{|y-\overline{y'}|} \quad (1)$$

$$EV' = EV \times \sqrt{\frac{x}{\max(x')} \times \frac{y}{max(y')}} \quad (2)$$

To further investigate the development of typical heathland community networks, all taxa were divided into two groups: i) heathland plants (i.e. *Calluna vulgaris* L., *Erica tetralix* L., *Rumex acetosella*, L., *Betula pendula* Roth, *Molinia caerulea* L., *Carex pilulifera L.* and *Juncus sp.*; often found in mature heathland vegetation), and heathland-related fungi belonging to the order Archaeorhizomycetales, Helotiales and the genus *Clavaria*, based on that they were found in high abundance in reference heathlands in the current study and/or that they are known to be abundant in heathlands (Englander & Hull 1980; Rosling *et al.*, 2011) or to contain ERM fungal taxa (Zijlstra *et al.*, 2005); ii) non-heathland taxa including all other plant species and fungal taxa. The list of all plant species included in the network analysis is shown in the Supporting Information (Table S5.4).

The change in the total strength of heathland vs non-heathland links between plants and fungi over time (from 2013 to 2017) was plotted for biotic and abiotic treatments. The first year

(2012) was not included since the hay treatment had only been established earlier that year. The links between fungi and plants in the early (2013) and the late phase of the experiment (2017) were visualized and overall network properties (number of connections, strength and modularity) were calculated. The strengths of links for individual taxa were normalized to a 0-1 range by dividing them with the highest overall strength value in the dataset. Weighted modularity was calculated based on the Walktrap algorithm (Pons & Latapy, 2005) which assesses the extent to which the network is divided into modules or clusters. It can range from -1 to 1, where positive values indicate that the number of edges within groups exceeds the number expected based on a randomly connected network, whereas higher values indicate stronger clustering (i.e. dense connections within and sparse connections between the clusters).

All calculations and network visualizations were performed in R using base functions and the *igraph* package.

5.4 Results

Fungal community composition

Over the six years of the ecosystem development, there was a clear directional change in fungal community composition (Figure 5.1) where time explained 12% of the variation ($F_{1,153}$ = 21.67, P = 0.001). When controlling for the effect of time, both biotic and abiotic treatments significantly influenced the fungal community composition (R^2 = 0.06, $F_{2,146}$ = 4.92 P = 0.001 and R^2 = 0.05, $F_{2,146}$ = 4.41, P = 0.001, respectively) and there was a significant interaction between them (R^2 = 0.04, $F_{4,146}$ = 1.81 P = 0.001). The direction of fungal community change was orthogonal to the reference heathlands community composition, indicating that overall community development across treatments was not directed towards the local reference communities (Figure 5.1).

In the reference heathlands, the most dominant orders were Archaeorhizomycetales and Helotiales comprising 57% and 15% of total reads, respectively. The relative abundance of these fungi consistently increased in experimental plots over time in all treatments (Figure S5.6). This increase was fastest and reached the highest levels in the sod treatment where the sum of the relative abundances of Archaeorhizomycetales and Helotiales in 2017 was comparable to that in the reference heathlands (mean = 69%, sd = 16 vs mean = 72%, sd = 6, respectively).

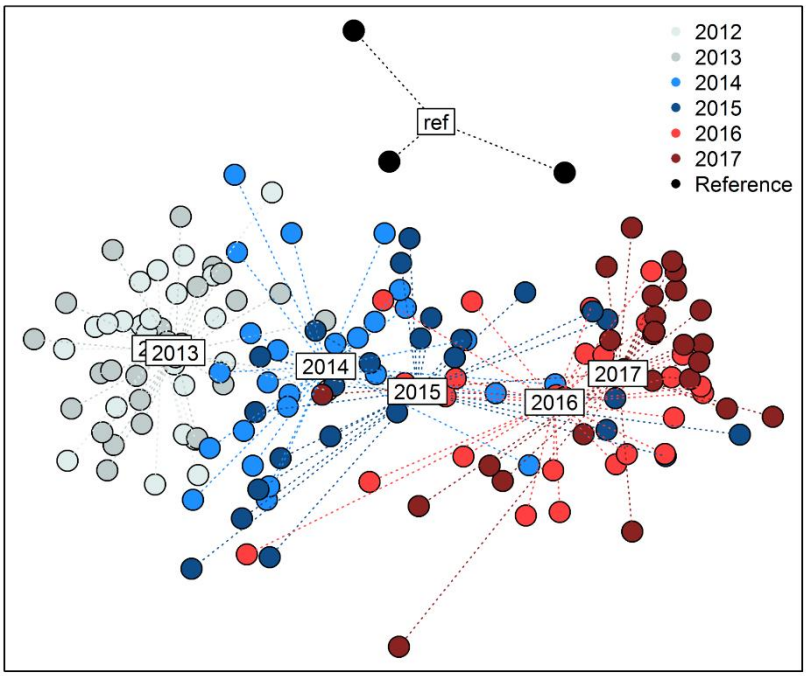


Figure 5.1 NMDS ordination showing the change in fungal community composition over the course of six years (from 2012 to 2017) compared to the reference heathland communities (ref). Different colours represent different years and dotted lines connect the samples from the same year with their group centriod. First two dimensions are shown (stress: 0.15). The ordination with the third dimension is presented in Figure S5.9.

In the last year of the experiment, both biotic and abiotic treatments still had a significant influence on fungal community composition ($P < 0.001$), with a slightly higher effect size of the former than the latter ($R^2 = 0.15$ and $R^2 = 0.13$, respectively), and a significant interaction between them ($R^2 = 0.17$, $P < 0.05$) (Figure 5.2a). Within biotic treatments, both hay and sod treatments differed from the control ($R^2 = 0.11$, $P = 0.01$ and $R^2 = 0.14$, $P = 0.003$; respectively), to a similar extent as in previous years (see Figure 5.3 for temporal development of between-treatment differences). In the case of abiotic treatments, fungal community composition significantly differed between the liming and the acidification treatment in 2017 ($R^2 = 0.12$, $P = 0.006$). The interaction between biotic and abiotic treatments is related to a larger response of fungal communities to abiotic treatments in the biotic control (grey symbols in Figure 5.2a) than in the sod treatment; there was a steadily decreasing dispersion (dissimilarity between samples across abiotic treatment levels) of fungal communities under sod treatment over time (Figure 5.2b), that was significantly lower than that of the control communities in 2017 ($P_{adj.} < 0.05$).

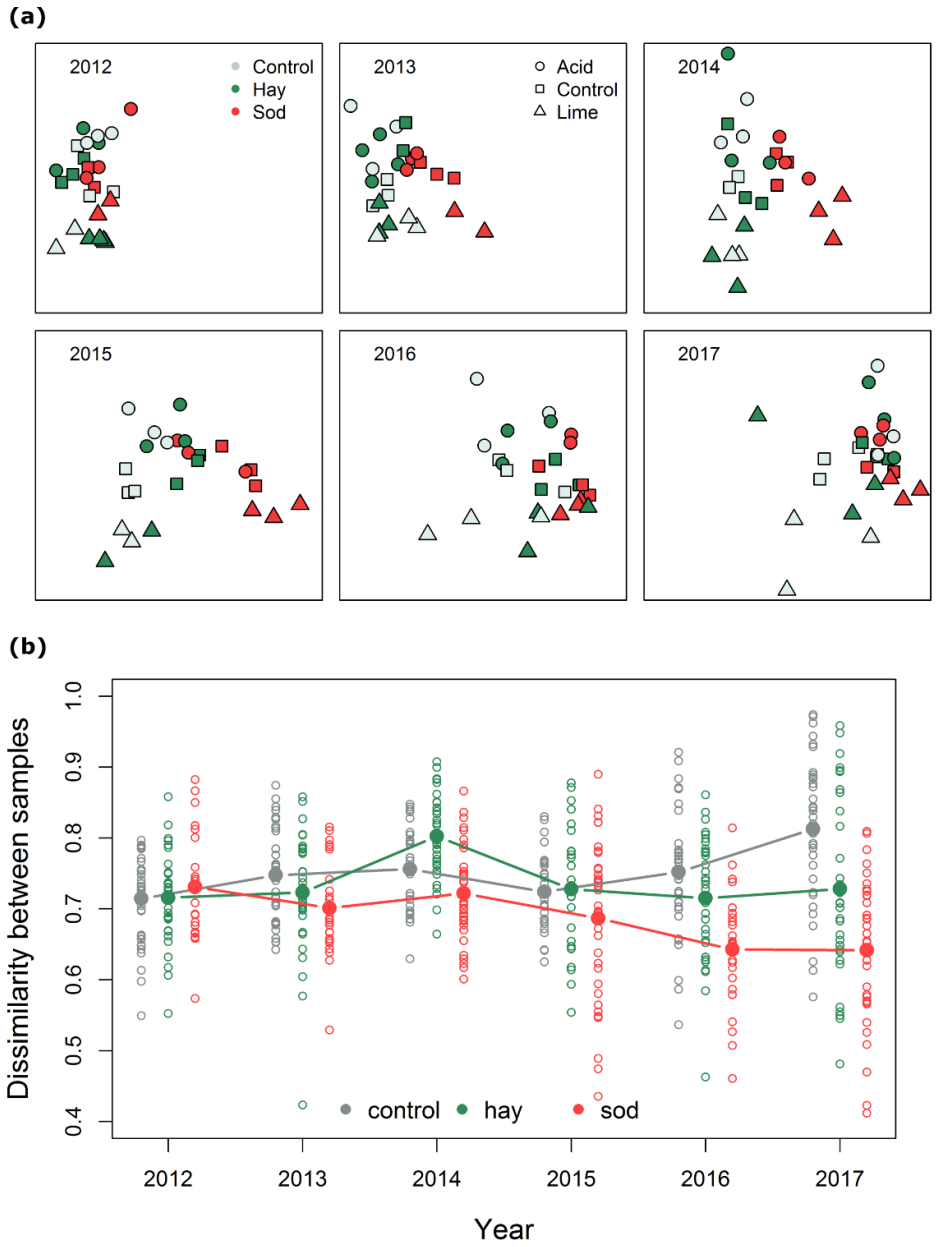


Figure 5.2 a) NMDS ordination of fungal community composition throughout six years (2012 – 2017) where each year is presented separately to emphasize biotic and abiotic treatments. The first two dimensions are shown (stress: 0.15). The ordination with the third dimension is presented in Figure S5.10. Different colours represent biotic (control, hay, sod) and shapes abiotic treatments (control, acidification, liming). **b**) Bray-Curtis distance (dissimilarity) between each fungal community in the biotic treatment to any other sample from that treatment (i.e. dispersion within biotic treatments but across abiotic treatments) over the same six years as in (**a**). Values are slightly shifted to increase visibility.

Fungal OTU richness was also significantly affected by time ($F = 15.9$, $P < 0.001$), biotic treatments ($F = 6.4$, $P < 0.01$), interactions between biotic and abiotic treatment ($F = 3.3$, $P < 0.05$) and interaction between biotic treatment and time ($F = 5.9$, $P < 0.001$). OTU richness tended to decrease over time in all treatments (with high variation between replicate plots), and this decrease was the most prominent in the sod treatment, in that it had the highest mean richness in 2012 and the lowest in 2017 of all biotic treatments. The other significant effects (interaction between biotic and abiotic treatments, and biotic main effect) are more complex and not straightforward to discern (Table S5.5).

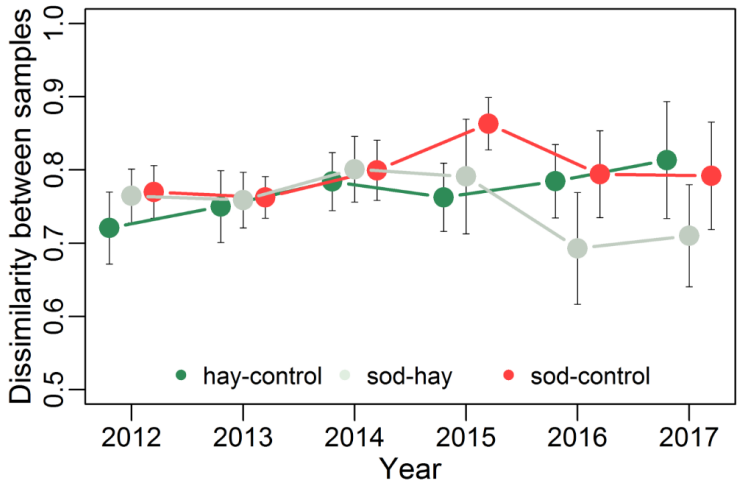


Figure 5.3 Mean Bray-Curtis dissimilarity between fungal communities exposed to different biotic treatments through time. Different colours represent different combinations of biotic treatments (sod vs hay = grey, hay vs control = green, sod vs control = red). 75 percentiles are shown as error bars. If values decrease with time there is a tendency for fungal communities in treatment-pairs to become more similar, and vice-versa.

DOC analysis

We used DOC (dissimilarity overlap curve) analyses to test whether biotic interactions between fungal taxa were important factors in shaping their community composition for each biotic treatment. The results indicate that biotic interactions had a significant influence in shaping fungal community composition in the sod and the hay treatment, evidenced by a negative relationship between community overlap and dissimilarity at high overlap region (sod: slope = -0.24, P_{real} = 0.005, P_{null} = 0.3, hay: slope = -0.18, P_{real} = 0.02, P_{null} = 0.8). For the control treatment, there was no significant relationship between community overlap and dissimilarity (slope = -0.02, P_{real} = 0.47, P_{null} = 0.7) (Figure S5.7).

Plant-fungal correlation networks

In 2013 (one year after all treatments were in place), the structure of plant-fungal correlation networks was very similar in the control and the hay treatment, consisting of relatively strong links between non-heathland taxa. In the sod treatment, however, the overall network strength was very low, with a relatively high number of links (Figure 5.4a). During the course of the experiment, the strength of links between heathland taxa increased while the strength of links between non-heathland taxa decreased, particularly in the hay and the sod treatment (Figure 5.4b).

The increase in strength of heathland taxa links occurred in the early stages of development for the sod treatment and was consistent across each abiotic treatment (Figure 5.4b). Furthermore, while the overall strength of connections increased by approximately 200%, the number of connections decreased by half (from 77 to 36). The core (most strongly connected) plant species was *C. vulgaris* with 12 links and a normalized strength of 1 (the highest strength for any taxon in any treatment). Modularity, which represents the extent of division of a network into modules or groups, decreased from 0.5 to 0.2 from 2013 to 2017. These results demonstrate that the taxa in the sod treatment became more interconnected over time, and the connections became stronger and more specific (i.e. occur almost exclusively between heathland taxa).

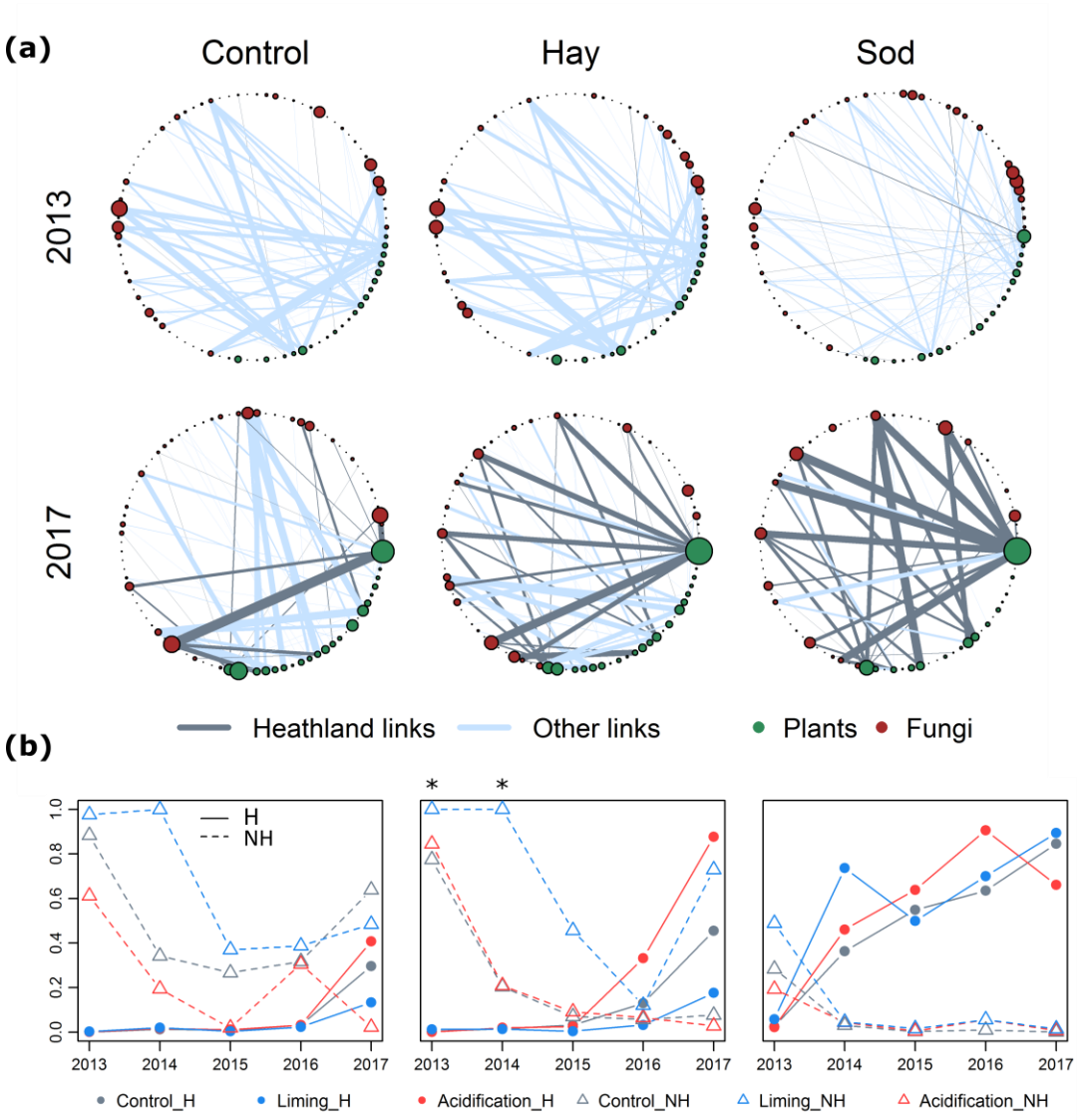


Figure 5.4 a) Positive plant-fungal interaction networks for 2013 and 2017 for three biotic treatments (control hay, sod). Green and red circles represent plant and fungal taxa, respectively. The size of the circles is proportional to the percentage cover for plant species and relative abundance for fungal OTUs. Lines represent the edges (connections) between the taxa and their width is proportional to the strength of connections. Darker lines represent links between the heathland taxa and lighter represent links between other taxa (note: this includes the links between the pairs where one or both taxa were classified as non-heathland and those that could not be classified). **b**) Change in the strength of links between heathland – H (full lines) and non-heathland – NH (dashed lines) taxa in time for control, hay and sod treatment. Different line colours represent abiotic treatments (grey – abiotic control, blue – liming, red – acidification). * = values higher than the maximum presented here are set to one for visibility.

Overall network structure in the hay treatment in 2017 was similar to the one in the sod treatment, consisting primarily of strong links between heathland taxa (Figure 5.4a) with *C. vulgaris* as a central species (12 links, strength 0.7). During previous years, the increase in heathland taxa in the hay treatment was 2-3 years delayed compared to the sod treatment and was altogether diminished in the liming treatment, where the strength of links between non-heathland taxa was still relatively high (Figure 5.4b).

In the biotic control treatment, the increase in the strength of links between heathland taxa started only in 2016 and was weaker than in the two other treatments, particularly under liming conditions. Therefore, the network structure in 2017 (Figure 5.4a) was still substantially different from the network structure in the sod and the hay treatments, with positive links both within heathland and non-heathland taxa (therefore higher modularity of the network = 0.5). Moreover, there were multiple core plant species; one from the heathland group – *C. vulgaris* with 7 connections (strength 0.4) and the other from the non-heathland group – *P. lanceolata* with 5 connections (strength 0.3).

Finally, given that most plant and fungal taxa in the network analysis occurred in all biotic treatments in 2013 at least once (Table S5.6), we expect there was no absolute dispersal limitation hindering the development of communities in the control treatment. Moreover, heathland taxa (plant and fungal) were present with similar frequencies in the control and the hay treatment at the beginning of the experiment (Figure S5.8).

5.5 Discussion

In the current study, we used a large-scale heathland restoration experiment to estimate the combined effects of different drivers of fungal community assembly. We found that 1) the initial presence of heathland soil communities and plant seeds had a persistent influence on fungal community composition and plant-fungal correlations networks after six years; 2) the early presence of the soil communities diminished the effect of abiotic (pH) conditions on both of these community aspects compared to the treatments without sod additions.

Timing of colonization alters the development of fungal communities – the role of biotic interactions

It has previously been shown that soil inoculation can significantly affect heathland community composition (Wubs *et al.*, 2016; van der Bij *et al.*, 2018), indicating that plant-soil biotic interactions are important in this ecosystem type. Here, we present three further lines of evidence to demonstrate the dynamic and nature of biotic interactions in the development of fungal community composition over a six-year time-scale. Firstly, there was a persistent difference in fungal community composition between biotic addition treatments and the control. This was true despite that biotic additions did not alter the initial soil abiotic conditions,

and fungi could easily colonize the non-inoculated plots from the adjacent inoculated plots. Similar findings were reported by Wubs *et al.*, (2019), where single introductions of soil biota and plant seeds led to long-term legacies on the trajectory of community assembly. Secondly, the DOC analysis indicates consistent biotic interactions among fungal taxa under sod additions and to a lesser extent hay additions, but this signal was absent in control communities. Thirdly, at the end of the experiment, the structure of plant-fungal correlation networks in the sod and in the hay treatment was clearly different from that in the control. In the first two treatments, the networks contained strong connections between "typical" heathland plant and fungal taxa whereas the control treatments exhibited relatively loose connections for either heathland and non-heathland taxa. Morriën *et al.* (2017) have previously shown that during the course of primary succession soil networks can become more tightly connected. Here, we show that after six years of development such connectivity is highly dependent on the initial biotic community, as only the networks formed under biotic additions become more strongly connected and more specific.

The importance of the initial presence of not only plant but also soil fungal partners is further corroborated by the slower development of links between heathland plants and fungi in the hay treatment compared to the sod treatment. Such dependence of plant community composition on soil biota is in line with many previous reports in greenhouse (van der Heijden *et al.*, 1998; Koziol & Bever, 2017) and field (Wubs *et al.*, 2019) settings. Specifically for heathlands, van der Bij *et al.* (2017) found that typical heathland vegetation developed much faster and typical heathland plants reached a much higher cover when a heathland soil community was already present. Our results suggest that when heathland seeds are present from the beginning, but a matching soil fungal community is absent or present at low abundance, it is more difficult for heathland plants and their associated fungal communities to develop. Apparently, additional heathland-related fungi first have to disperse into the plots and become established, causing heathland plant-fungal links to develop later as compared to the sod treatment. However, once their abundance reaches a certain threshold, further development of the heathland system is relatively fast and ultimately resembles the sod treatment. This means that, in terms of heathland restoration, hay additions can in longer-term provide similarly successful results as sod additions.

In the control treatment, both plant seeds and soil microbes were introduced gradually through dispersal. These plots were situated next to the inoculated plots and close to a larger area of abundant heathland vegetation, which poses a significant source of heathland taxa available to colonize them. It has been shown that the vicinity of source sites is an important factor promoting heathland community development (Torrez *et al.*, 2016; van der Bij *et al.*, 2017). Surprisingly though, despite the fact that control plots collectively contained the majority of plant and fungal taxa observed in other treatments, including heathland taxa, the increase in the strength of links between heathland plants and fungi was notably delayed or absent compared to the sod-inoculated plots. A small-scale mismatch between heathland plants and fungi in time

and space is likely the reason that links between them are not often formed, leaving opportunities for non-heathland plants and fungi to establish. This could result in the local development of competing plant-microbe systems, as evidenced by higher network modularity in the control treatment; one consisting of heathland and the other of non-heathland plant and fungal taxa, with relatively weak positive links within these modules. Whether these links between plants and fungi are strong enough to fuel positive feedback will likely determine the long-term trajectory of the non-inoculated plots, and whether the heathland system can successfully be restored or an alternative one will eventually prevail. The stochastic processes operating in this heathland system are likely to contribute to the 50% of variance not accounted for by different biotic and abiotic treatments or time.

Together, these observations suggest that initial simultaneous presence of a relatively large pool of heathland fungi and plant seeds in the sod treatment promotes the early formation of strong positive plant-fungal feedbacks between heathland taxa, thus reinforcing their further development. These early feedbacks can create priority effects (Kardol *et al.*, 2007) and hamper the successful development of non-heathland fungi, leading to lower overall OTU richness observed in the sod treatment. Mechanisms behind these feedbacks could be both symbiosis, such as between plants and mycorrhizal fungi (Kerley & Read, 1998) but also competition for limiting nutrients or direct antagonism between plants or fungi, as has been shown to elicit priority effects in nectar-yeasts (Vannette *et al.*, 2014; Fukami, 2015). That plant-fungal soil interactions have indeed a high potency in creating priority effects has previously been demonstrated by Peay (2018), where the timing of ectomycorrhizal inoculation had a strong effect on the development of pine seedlings and on their success against competitors associated with AMF.

Which fungi would be responsible for the differences between treatments and control? Members of two dominant fungal orders, Archaeorhizomycetales and Helotiales strongly increased under biotic additions, particularly in the sod treatment, where they reached an abundance similar to that in the reference heathlands. Therefore, even though soil communities in the experimental site did not move towards those in the reference in terms of OTU identities, they became similar in terms of dominant fungal groups, which might play similar roles in the ecosystem. It is well known that Helotiales contain taxa that are associated with heathland plants (Zijlstra *et al.*, 2005; Leopold, 2016). Archaeorhizomycetales are relatively poorly investigated fungi that are typically found in roots and rhizosphere (Rosling *et al.*, 2011, 2013) and might depend on root-derived carbon (Schadt *et al.*, 2003). Given that these fungi are very abundant in the reference heathlands, they potentially form important associations with heathland plants as symbionts or decomposers. Further research is needed to reveal more about the nature of connections of these fungi with heathland plants and their possible importance in heathland restoration.

Convergence of communities under sod additions – biotic interactions override the effect of pH

The factorial experiment with a crossed abiotic and biotic additions allows us to test whether this abiotic filter has precluded biotic interactions to play out, as a hierarchical model of community assembly would suggest (Belyea & Lancaster, 1999). Under this model, we should expect communities to increasingly sort according to the environmental gradient as species disperse in, where the biotic addition treatments are given a head start. In contrast, the multivariate dispersion analyses show that fungal communities in the sod treatment converge over time, regardless of abiotic differences. Furthermore, the plant-fungal correlation networks in this treatment were also not influenced by the differences in abiotic conditions. These results indicate that environmental and biotic filters interact with each other and do not influence heathland communities in a solely hierarchical way. In the absence of initial "target" soil communities, abiotic pressures were apparently more influential, and liming in particular favoured stronger positive links between non-heathland plants and fungi, which are typically generalist that are less successful on acidic soils. In contrast, the links between heathland taxa were promoted under acidification because heathland plants thrive under acidic conditions (Lawson *et al.*, 2004; Díaz *et al.*, 2008, 2011) and likely heathland fungi too, as known to be the case for Helotiales (Rousk *et al.*, 2010).

This, however, raises the question of why the development of connections between heathland taxa in the sod treatment was not affected by sub-optimal (increased pH) conditions. It is possible that plant-associated heathland fungi can strengthen the heathland plant performance (and vice-versa) even under sub-optimal conditions through positive feedbacks, and hinder the establishment of other, otherwise competitively superior species that are developing in the control plots. Research on facilitation has highlighted that positive interactions between species - particularly mutualistic ones - can expand their tolerance to the abiotic environment (Callaway & Walker, 1997; Bruno *et al.*, 2003; Poisot *et al.*, 2011; Kazenel *et al.*, 2015; Peay, 2016; Gerz *et al.*, 2018). For instance, it has been shown that ectomycorrhizal fungal symbionts can help seedlings establish and persist under suboptimal conditions (Simard, 2009). Our results strongly suggest that, in heathland systems, biotic links can override "environmental filters" supporting the proposal of Cadotte & Tucker (2017) and Aguilar-Trigueros *et al.* (2017) that these are much less rigid than previously thought.

5.6 Conclusion

The findings presented here suggest that the timing of colonization has an important effect on the development of fungal community composition in heathland systems through shaping plant-fungal interaction networks. We propose that the early stage presence of heathland soil communities and the interactions they form can reinforce the development of a heathland system and alleviate the abiotic filter imposed in the absence of these interactions. If the system is exposed to slow dispersal, other incoming plant and fungal species establish their own, alternative interactions possibly leading to a strongly altered community trajectory that is more sensitive to the abiotic context. These results have clear implications for our capacity to steer community development, for instance in the context of heathland restoration, through manipulation of keystone plants and fungi.

Consistent predictors of bacterial and fungal community composition in grassland soils worldwide

CHAPTER VI

CHAPTER VI

Consistent predictors of bacterial and fungal community composition in grassland soils worldwide

6.1 Abstract

Potential drivers of soil microbial community assembly have been extensively studied, but it is currently unclear whether a consistent set of predictors operates in different soils, or whether context-dependency prevails. In this study, we used a network of globally distributed local-scale gradients in grassland plant productivity to examine i) if the same abiotic or biotic factors can predict both global- and local-scale patterns in bacterial and fungal community composition and ii) if community composition differs consistently with local plant productivity (low vs high) across different sites. We found that microbial community composition can be predicted by similar factors on the global and the local scale; with bacteria predominantly associated with soil properties (such as base saturation and pH) and fungi predominantly associated with plant community composition. Moreover, there was a microbial community signal that distinguished high and low productivity levels that was shared across worldwide-distributed grasslands. However, the relationship among dominant bacterial and fungal groups, and between these groups and plant functional groups and soil properties, differed substantially between productivity levels; there were many more significant relationships at low compared to the high productivity level. These findings suggest that while the predictors of overall microbial community composition in grasslands are consistent across two important contexts (spatial scales and productivity levels), the abundance of bacteria and fungi and the relative abundance of their dominant groups are likely influenced by different factors that vary with local grassland productivity and the amount of resources.

6.2 Introduction

Variation in the strength and direction of ecological relationships under different conditions (i.e. context-dependency) is common in nature (e.g. Maestre *et al.*, 2005; Chamberlain *et al.*, 2014; Tedersoo *et al.*, 2015; Song *et al.*, 2020). Biotic and abiotic predictors of microbial community composition at particular spatial scales and under specific environmental contexts have been thoroughly studied (e.g. Fierer & Jackson, 2006; de Vries *et al.*, 2012; Tedersoo *et al.*, 2014; Delgado-Baquerizo *et al.*, 2018), but it is currently not well understood whether there is generality in how these drivers operate across different contexts. Context-dependency in the processes that structure microbial communities arises for several reasons, including altered

plant-soil interactions, different historical legacies (Fukami, 2015), stochastic events in community assembly processes (Beck *et al.*, 2015), dispersal limitation (Peay *et al.*, 2010) all possibly leading to different main drivers to dominate microbial community assembly in different regions (Hendershot *et al.*, 2017).

Several studies have indeed shown that the drivers of microbial community composition might not be consistent at different spatial scales (Shi *et al.*, 2018; Martiny *et al.*, 2011; Chalmandrier *et al.*, 2019; Guo *et al.*, 2020). For instance, Alzarhani *et al.* (2019) investigated community composition of root-associated fungi in salt marshes in two geographical regions and identified different abiotic and biotic predictors depending on the site and spatial scale, indicating that the drivers of these communities are strongly context-dependent. Similarly, Chalmandrier et al. (2019) and Shi et al. (2018) found that at the landscape (regional) scale, soil microbial community composition was predicted by either environmental factors or plant community composition, but these relationships were weak or absent at the local scale.

To what extent, then, there is any commonality in processes shaping microbial community composition across different contexts has been challenging to examine because most studies up to now have either been restricted in spatial range or their sampling design did not allow to distinguish context-dependency from either noise or generality. For instance, while global-scale studies have found that microbial community composition can be strongly related to soil abiotic properties, such as pH (Fierer & Jackson, 2006; Delgado-Baquerizo *et al.*, 2018a) and plant community composition (Prober *et al.*, 2015), it remains unclear to what extent the observed correlations are due to the fact that microbial and plant communities, as well as soil properties, covary strongly with climate and geographical distance (Steidinger *et al.*, 2019). On the other hand, regional- and local-scale studies have a better potential to assess the effect of soil properties and plant communities along an environmental gradient, but their findings may not generalize across different places (Alzarhani *et al.*, 2019).

Here, we used a global network of grassland sites that contain local-scale plant productivity gradients (Fraser *et al.*, 2015) to examine whether soil bacterial and fungal community composition in grasslands can be predicted by the same factors under different contexts, or whether to the contrary context-dependency prevails. Given that grassland productivity is intrinsically related to biodiversity, soil fertility and plant-soil interactions (Craven *et al.*, 2016; Delgado-Baquerizo *et al.*, 2017; Guerrero-Ramírez *et al.*, 2019), and therefore to the overall functioning of the system, different local-scale productivity levels provide entirely different underlying environmental contexts for the development of soil microbial communities. For instance, plant competition for light is expected to increase with productivity (Grace et al 2016) favouring acquisitive fast-growing plant species (DeMalach *et al.*, 2016). Such development is mirrored in soil, where the high input of easily decomposable plant litter may select for more acquisitive biota such as many gram-negative and other bacteria (Marschner *et al.*, 2011), to the detriment of fungi and microbes engaged in nutritional symbioses with plants (de Vries *et*

al., 2007, 2012; Johnson *et al.*, 2008). The question remains if these processes structuring microbial assembly would play out in a similar manner across different sites and different (e.g. climatic) conditions.

To address this question, we first analysed the importance of climate, geographical distances, nitrogen deposition, plant biomass, plant community composition and 15 soil properties in explaining global bacterial and fungal community composition patterns. We then tested if locally varying factors such as soil properties, plant biomass and plant community composition can consistently predict bacterial/fungal community composition across different sites and climatic conditions. Further, we examined whether two different grassland productivity levels (low and high) had globally consistent effects on microbial communities across different sites. Given that consistent relationships across different contexts might be more apparent at high taxonomic levels or functional groups than in the overall community (Alzarhani *et al.*, 2019), we also explored if the dominant bacterial taxa and fungal functional groups, as well as bacterial and fungal abundances, form different correlations with soil properties and plant biomass/functional groups at different productivity levels.

If context-dependency is strong, we expect to find different, site-specific, local-scale drivers of microbial communities. In that case, the predictors identified on the global scale would not predict well the local-scale variability across different sites. For instance, we anticipate that at the global scale, climate and geographical distances would be dominant predictors while on the local scale, microbial communities would be determined by different soil properties and/or plant community composition. Furthermore, we would expect no common signal in the way community composition is shaped between two productivity levels because the effect of plant productivity on microbial community composition would vary across globally distributed grassland sites (e.g. depending on the climatic conditions, biogeography, or soil type).

6.3 Materials and methods

6.3.1 Sampling sites and data collection

Data was collected from 18 Herbaceous Diversity Network (HerbDivNet) grassland sites (Fraser *et al.*, 2015) located in 12 countries distributed over six continents (Figure 6.1). Each of the 18 sites contained between two and six plots of 8 x 8 m: 11 sites contained six plots, one site contained four plots, one site three plots and five sites contained two plots (Table S6.1); 83 plots in total. Most sites were chosen to represent an estimated gradient in productivity (low, medium and high) with six plots; two per each productivity level. However, some sites contained fewer plots and did not show a prominent productivity gradient. A clear gradient in

biomass productivity was accomplished in 11 sites; including ten with six plots and one with four plots (Figure 6.1). Other general site characteristics are described in Chapter III.

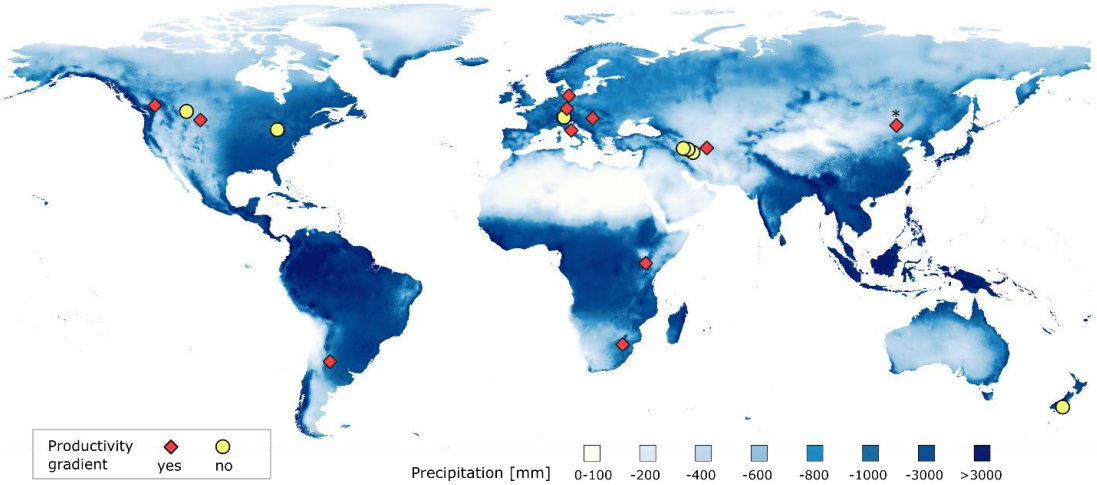


Figure 6.1 The location of 18 HerbDivNet sites across a precipitation gradient. Red diamonds indicate 11 sites that contained a clear productivity gradient and yellow circles indicate other sites (containing from 2 to 6 plots but with no clear productivity gradient). All the sites and plots were used in the analyses of global and local-scale predictors of microbial community composition while 11 sites with the productivity gradient were used in the analyses of microbial community composition at high and low productivity levels. *This site consists of 3 relatively close sites (each with 2 plots) that were kept separate in Chapter III to take into account small, but relevant, differences in precipitation between them.

6.3.2 Soil and plant sampling

Soil was sampled in a single sampling event at the peak of the growing season in the period between 2017 and 2018, depending on the site (Table S6.1). For each plot within a site, five samples were taken using soil corers from four corners and the centre of the plot at 0-10 cm depth. Subsamples for microbial analyses were taken and stored in pure ethanol and the rest of the sample was pooled into one composite sample (a total of 83 samples), air-dried and sieved at 2 mm. All samples were further analysed at the University of Antwerp. Samples for microbial analyses stored in ethanol were kept cool until the DNA extraction (see below). Although soil samples for DNA analysis should ideally be frozen shortly after sampling, due to practical constraints they were preserved in ethanol which was shown to yield similar DNA recovery as cold conservation (Harry *et al.*, 2000).

Total aboveground biomass was harvested and plant species from each m^2 of each 64 m^2 plot were identified at the peak of the growing season. Sampling was performed in the period between 2012 and 2017 in a single sampling event per site (Table S6.1). Litter was first excluded from the total biomass and live biomass was dried and weighed. Based on this, average peak AGB production [g/m^2] was calculated for each plot as in Chapter III.

The data on the presence of different plant species at each m^2 of the plot was used to derive the 'frequency' of different species per plot (with the highest possible value of 64 – for species present at each m^2 of the 64 m^2 plot) which was used as a measure of relative abundance. Further analyses of plant community composition distances were based on species aggregated to genera (as in Prober et al. (2015)) rather than to the species level given the plant species turnover across different plots and sites would often be 100% and thus produce continuous data at highly similar communities only, reducing information content.

6.3.3 Climatic and N deposition data

Mean annual precipitation (MAP) and temperature (MAT) were derived from the CHELSA database (as in Chapter III) based on the geographical position (latitude and longitude) of each plot which was also used to derive geographical distances [km] between the plots. Data on aridity and potential evapotranspiration (PET) were obtained using CGIAR-CSI Global-Aridity and PET Databases. Data on total inorganic nitrogen deposition [kg/ha/yr] were derived from Ackerman et al. (2018).

6.3.4 Analyses of soil physicochemical properties

The following soil parameters were analysed for each of the 83 plots: soil organic matter (SOM), total nitrogen (N), total carbon (C), total phosphorus (P), available P (Olsen), base saturation (BS), cation exchange capacity (CEC), pH, soil texture (% sand, % clay, % silt), bulk density (BD) and extractable Ca, Mg and K. The methods of analyses of soil properties are described in Chapter III.

6.3.5 Analyses of microbial communities

Sample preparation and sequencing

DNA was isolated from 415 soil samples (5 samples for each of 83 plots) using 0.25-0.35 g of soil with the DNeasy PowerSoil Kit according to the manufacturer's protocol (Qiagen, Venlo, the Netherlands). The bacterial 16s V4 region was amplified using the 515F-806R primer pair

and the fungal ITS1 region was amplified using general fungal primers ITS1f and ITS2 (see Chapter IV). The preparation of PCR mixture was identical as in Chapter V. PCR conditions were as follows: initial denaturation at 98 °C for 60 s, followed by 30 (35 for fungi) cycles of: denaturation at 98 °C for 30 s, annealing at 55 °C for 30 s, extension at 72 °C for 30 s; and an additional extension of 72 °C for 10 min. The success of amplification was tested on 1.5% agarose gel. For the samples that did not amplify successfully, amplification was attempted again with a modified mixture that contained 2 µl of the sample and 1 µM of forward and reverse primer. Successful PCR products were diluted 50 x and a second PCR was performed using dual barcoded primers with Illumina adapters (2.5 µl of diluted PCR products and 0.1 µM of each primer). The conditions were: 98 °C for 60 s, 12 cycles: at 98 °C for 10 s, 63 °C for 30 s, 72 °C for 30 s; and 72 °C for 5 min. PCR products were run on an agarose gel and successful amplicons were purified and normalized using the SequalPrep Normalization Plate Kit (ThermoFisher Scientific) and pooled into a single library. The library was purified through gel extraction using QIAquick Gel Extraction Kit (Qiagen, Venlo, the Netherlands) and quantified using qPCR. The sequencing was performed using the Illumina MiSeq platform (Illumina Inc; San Diego, CA, USA) with 300 cycles for forward and reverse reads.

Quality filtering and bioinformatics analyses

The sequences were analysed using the USEARCH (v8.1.1861) and VSEARCH (Rognes *et al.*, 2016) software following the UPARSE pipeline (Edgar, 2013). After trimming to 280 bp and 250 bp for bacteria and fungi respectively, the paired-end reads were merged and primers were removed. This trim length was chosen because it was the optimal length for merging paired reads by removing the low-quality bases at the end. Merged sequences were quality filtered using the expected number of errors (E) as a measure of read quality, with a threshold of $E_{max} = 0.5$. This yielded 10.8 M and 4.02 M of good-quality reads, for bacteria and fungi, respectively. Following singleton removal, the sequences were clustered into OTUs (operational taxonomic units) based on 97% similarity using the UPARSE-OTU algorithm (Edgar, 2013) which automatically detects and filters out *de novo* chimaeras with high efficiency. Filtered reads were then mapped to the OTUs with an identity threshold of 0.97, yielding an OTU table for bacteria and fungi. The raw sequences from this study will be deposited in SRA-NCBI database.

Representative OTUs were aligned to bacterial sequences in the SILVA database (Quast *et al.*, 2013) (release date 13.12.17) and fungal sequences in the UNITE database (Kõljalg *et al.*, 2005) (release date 10.10.2017), using the *sintax* command in USEARCH with 0.8 cut-off. Non-bacterial and non-fungal sequences were removed from OTU tables resulting in a total of 19,248 and 13,967 OTUs for bacteria and fungi, respectively.

Further steps were performed using R software. The number of reads per sample was rarefied using the *rrarefy* function in *vegan* (Oksanen & *et al.*, 2015) to 6,046 for bacteria and 1,231 reads for fungi as rarefaction curves showed that the number of taxa was levelling off for most

samples at these depths (Figure S6.1). Most bacterial samples had a higher amount of sequences that the chosen rarefaction depth but 60 samples had fewer sequences than this threshold. Out of these, 13 samples had too few sequences or they were outliers and they were therefore discarded and 47 (with the minimum amount of sequences of 3571) were normalized to contain 6046 sequences, leaving 402 samples for bacteria. For each of these samples, it was verified that they do not notably deviate from the other samples in their group (i.e. those sampled from the same plot). For fungi, all but 13 samples had more sequences than the chosen rarefaction depth, and the letter were therefore omitted leaving 402 samples in total. Both for bacteria and fungi, no less than 3 samples per plot were retained in any plot.

To annotate *sintax*-assigned fungal sequences to known genera in the UNITE database, we used NCBI's BLAST algorithm with default settings. OTUs were then assigned to particular taxa if they had a maximum E-value of 10^{-36} (as in Chapter V) and from this, the lowest E-value hit with a known genus was selected. If there were none, the genus level was left unassigned. OTUs were subsequently assigned to functional groups if the genus was successfully matched with one of the genera with known lifestyles in Tedersoo et al. (2014) and Liu et al. (2016a).

6.3.6 Analysis of microbial abundance

DNA extracts of the five replicate samples per plot were first pooled into one sample, leaving 83 samples in total. The abundance of bacterial and fungal genes per sample was quantified using qPCR targeting 16s V4 region (with the 515F–806R primer pair) for bacteria and 18s region for fungi (primer set FR1 / FF390 (Chemidlin Prévost-Bouré *et al.*, 2011)). Each 20 µl reaction mixture contained 4 µl of the sample, 0.5 µM of each forward and reverse primer, 1 x ROX high and 10 µl of KAPA SYBR FAST qPCR master mix (Kapa Biosystems, Wilmington, MA, USA). qPCR conditions were as follows: initial denaturation at 95°C for 3 min, followed by 40 cycles of: denaturation at 95 °C for 3 s, annealing at 57 °C (52 °C for fungi) for 20 s, extension at 72 °C for 12 s; finishing with 35 s at 50 °C. Prior gel-electrophoresis with these primers and reaction conditions showed the reactions were highly specific. Melting curve analysis of all amplicons was conducted to confirm that fluorescence signals originated from specific amplicons and not from primer-dimers or other artefacts. Standard curves were generated using duplicates of 10-fold dilutions of amplicons derived using the same primers, isolated from the gel using QIAquick Gel Extraction Kit (Qiagen, Venlo, the Netherlands) and quantified using Qubit fluorometer (Invitrogen GmbH, Karlsruhe, Germany). Fungal and bacterial gene copy numbers were derived from a regression equation based on the standard curves (with minimal $R^2 > 0.99$) by relating the quantification cycle (Cq) value of each sample to the Cq values of standards with the known number of copies. All reactions were performed in duplicate and the number of bacterial and fungal copies was then averaged (a deviation of Cq between replicates < 1 was used as a passing criterium) and expressed per g of dry weight soil.

6.3.7 Statistical analyses

Global- and local-scale predictors of microbial community composition

First, we averaged the OTU relative abundances of five samples taken at the centre and the corner of each plot (83 plots from 18 sites) to obtain one community measure per plot. We further analysed the influence of variation in geographical distance, abiotic environmental factors, plant biomass and plant communities on the global-scale variation in bacterial and fungal community composition using multiple regression on distance matrices (MRM) in the *ecodist* package (Goslee & Urban, 2007). Environmental variables included: two climatic variables (MAT and MAP), N deposition, plant biomass and 15 soil variables (pH, total N, total P, available P, CEC, BS, BD, % sand, % clay, % silt, SOM, Ca, Mg, K and C:N ratio). All variables (except pH and BS) were transformed using square root transformation, centred and scaled to reduce positive skewness and to allow for the comparison of effect sizes. Community data (fungi, bacteria plants) were transformed with Hellinger transformation using the *decostand* function in the *vegan* package. Distance matrices for communities were created using the Bray-Curtis (BC) distances, and distance matrices for each of the environmental factors (and plant biomass) were created using Euclidean distances.

For the global-scale analysis, a multiple regression model on distance matrices was first fit using bacterial/fungal distances as response variables and all variables (except plant community composition) as independent variables. Subsequently, a backward model selection was performed, where the variables that did not significantly contribute to the model were sequentially removed and the final model included only the remaining variables that had a significant effect ($P < 0.05$). In this way, all available environmental variables that explained unique variation in global community composition were taken into account. To test if and how much plant communities can add to the variation explained by geographic distances and the environment, we added plant community distances to the best model.
In order to partition the variance explained between predominantly broad-scale predictors, ecosystemic environmental predictors and community-related predictors we created three groups of variables: i) climate + N deposition + geographical distance; ii) soil parameters + plant biomass; iii) plant community composition; and assessed how much of unique and shared variation each of these three groups explained.

To examine if the observed global-scale relationships (across all the plots and all the sites) persist on the local scale (i.e. between the plots within each site; these share the same climate and are part of the regional species pool), we created a common variable that represents the influence of the selected ecosystemic environmental factors (soil and plant biomass) by first multiplying each selected variable by its coefficient in the best MRM model and then summing them into one variable. The within-site (Euclidean) distances in the environmental variables

were then plotted against the within-site distances in bacterial and fungal communities. In the same way, we plotted the within-site microbial distances against the within-site plant community distances to examine how well plant community dissimilarities can predict microbial community dissimilarities at the local scale. To assess the consistency of these relationships (environment – bacteria, plants – bacteria, environment – fungi, plants – fungi) across sites that contained more than three plots, we calculated the variance in their slope values and we reported their mean R^2 values as well as the standard deviations.

Microbial community composition at different local relative productivity levels

Our local productivity gradients allowed us to test whether there is a general difference between relatively low-productivity and high-productivity grasslands replicated at a global scale. For this analysis, the dataset was divided into two parts: one containing two plots with low productivity and the other containing two plots with high productivity from each site. 11 sites that had a clear productivity gradient were considered. These sites were selected because they had a strong difference in plant biomass between the plots of low and high productivity (where the average of two plots with the highest productivity was at least 100% higher than the average of two plots with the lowest productivity). This yielded two datasets, each containing 22 plots.

In order to examine if bacterial and fungal communities differed significantly between the two productivity levels with a consistent pattern across sites distributed all over the globe, we performed PERMANOVA analysis using the *adonis* function in *vegan* adding 'site' as *strata* to account for site differences. We used multidimensional scaling (MDS) ordination to visualise the BC distance in bacterial and fungal communities at different productivity levels after removing the effect of 'site' differences (to control for inherent community differences between sites) using the *dbrda* function in *vegan*. To examine if the best predictors of bacterial and fungal community composition differed at different productivity levels, we repeated model selection described above (using the MRM function) for microbial communities at each of the productivity levels.

Furthermore, we examined if there was a significant change ($P < 0.01$) in the relative abundances of bacterial and fungal taxonomic/functional groups and bacterial and fungal abundances (number of gene copies) at low compared to high productivity levels using the *lme* function in *nlme* package with 'site' as a random effect.

Finally, we examined whether the correlation networks between microbial groups/microbial abundances, plant functional groups and soil properties across different sites differed between low and high productivity levels. To this end, we analysed the pairwise correlations (using *corr.test* in the *'psych'* package) between the three most dominant bacterial taxa, three most dominant fungal functional groups, plant aggregated to three plant functional groups (grasses, forbs, legumes), fungal and bacterial abundances (including fungal to bacterial ratio), plant

biomass and the most important, non-redundant soil properties (SOM, CEC, BS, pH, total N, C:N, total P, available P and % sand) for low and high productivity datasets. Only the correlations with Spearman $r > 0.5$ and P-value < 0.01 were retained and visualised in the form of correlation networks.

All statistical analyses were performed using the R software, version 3.6.1 (R Core Team 2019). The results were visualized using the *ggplot2* package in R (Wickham, 2016).

6.4 Results

Global-scale predictors of microbial community composition

Environmental variables and geographical distance explained 65% of the global-scale variation in bacterial community composition, where the strongest individual predictors were base saturation and pH (Table 6.1). When plant distances were added to this model, they increased the variance explained from 65% to 72%. Overall, in the model including plant community composition, ecosystemic environmental factors (particularly soil base saturation and pH) were the best predictors of bacterial community composition explaining 35% of the unique variation (Figure 6.2). The best model with environmental variables and geographical distances explained 44% of the variation in fungal community composition, where broad-scale predictors (climate, N deposition and geographical distances) played the most important role (Table 6.1). Adding plant communities to this model increased the explained variation to 59% (Figure 6.2). Plant community distances alone explained more variation in fungal communities than all other factors combined (51% compared to 44%) and they explained a relatively high amount of unique variation in the model together with environmental factors and geographical distances (15%).

Table 6.1 Global-sale predictors of bacterial and fungal community distances. The upper part of the table shows the broad-scale variables and ecosystemic variables that were selected in the best model. R^2 values of the individual relationships between the variables and bacterial/fungal community composition are shown together with the coefficients and P values for each of the variables in the model. The coefficients of variables shaded in grey were used to weigh these variables before summing them to obtain a single representative ecosystemic environmental variable for bacteria and fungi in Figure 6.3. The lower part of the table shows the variation explained by community-related variable alone.

		Bacterial communities Best model R^2 = 0.65			**Fungal communities** Best model R^2 = 0.44		
		Individual R^2	Coeff. (model)	P-value (model)	Individual R^2	Coeff. (model)	P-value (model)
Broad-scale variables	Geo. dist.	0.04	0.02	0.002	0.15	0.003	0.001
	MAT	0.03	0.01	0.003	0.13	0.02	0.001
	MAP	-	-	-	0.03	0.006	0.001
	N deposition	0.18	0.04	0.001	0.13	0.02	0.001
Ecosystemic variables	Plant biomass	0.02	0.01	0.001	0.05	0.006	0.008
	pH	0.27	0.04	0.001	0.09	0.02	0.001
	CEC	0.16	0.02	0.001	0.06	0.01	0.001
	N total	0.10	0.02	0.001	-	-	-
	C:N	-	-	-	0.02	0.005	0.013
	BS	0.34	0.04	0.001	-	-	-
	Sand	0.05	0.01	0.001	0.02	0.006	0.008
Community-related variable	Plant community composition	0.27	NA	NA	0.51	NA	NA

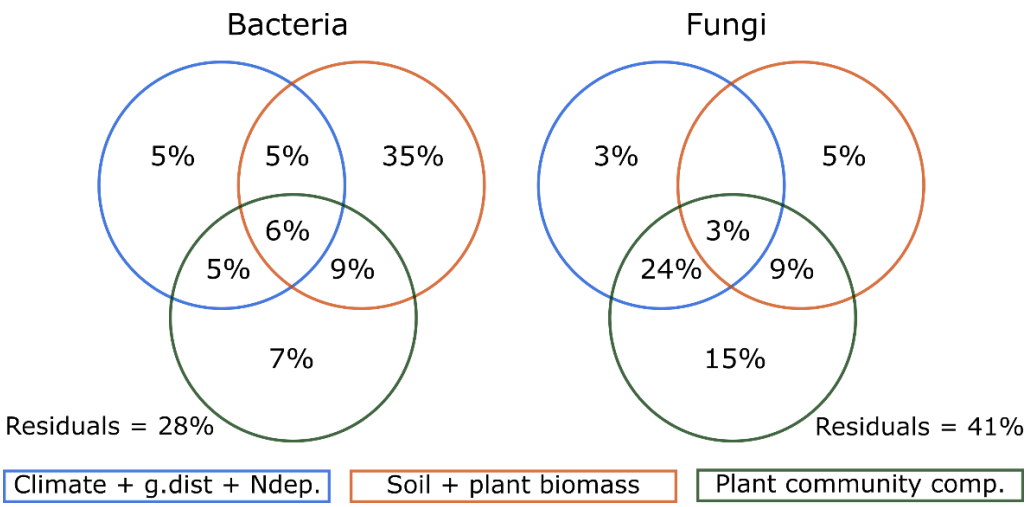


Figure 6.2 Variance partitioning between selected variables in the best model explaining bacterial and fungal community distances. The variables were grouped in 3 categories: i) predominantly broad-scale drivers (climate, N deposition and geographical distance); ii) ecosystemic environmental drivers (soil properties and biomass) and iii) plant-community composition.

Local-scale predictors in microbial community composition

To examine to what extent global scale predictors can predict microbial community composition at different grassland sites, the within-sites distances in the environmental variables (created using the most important ecosystemic environmental predictors in the global model for bacteria/fungi, as presented in Table 6.1) and plant community composition were regressed against the within-site distance in bacterial/fungal community composition. We found positive relationships between the environmental variables and bacterial and fungal communities in all the sites (Figure 6.3a and Figure 6.3c), but these relationships were much more consistent for bacteria than for fungi (slope variance values were 0.05 versus 0.16, respectively). The average amount of variation explained by the environment was 58% (sd = 32%) and 50% (sd = 32%), for bacteria and fungi, respectively. Moreover, the relationships between plant communities and bacterial/fungal communities were also positive for majority of the sites (Figure 6.3b and Figure 6.3d), but they were slightly more consistent for fungi than for bacteria (slope variance values were 0.05 versus 0.06, respectively). The average amount of variation explained by plant community composition was 64% in both cases (sd = 28% and sd = 26%; for bacteria and fungi, respectively).

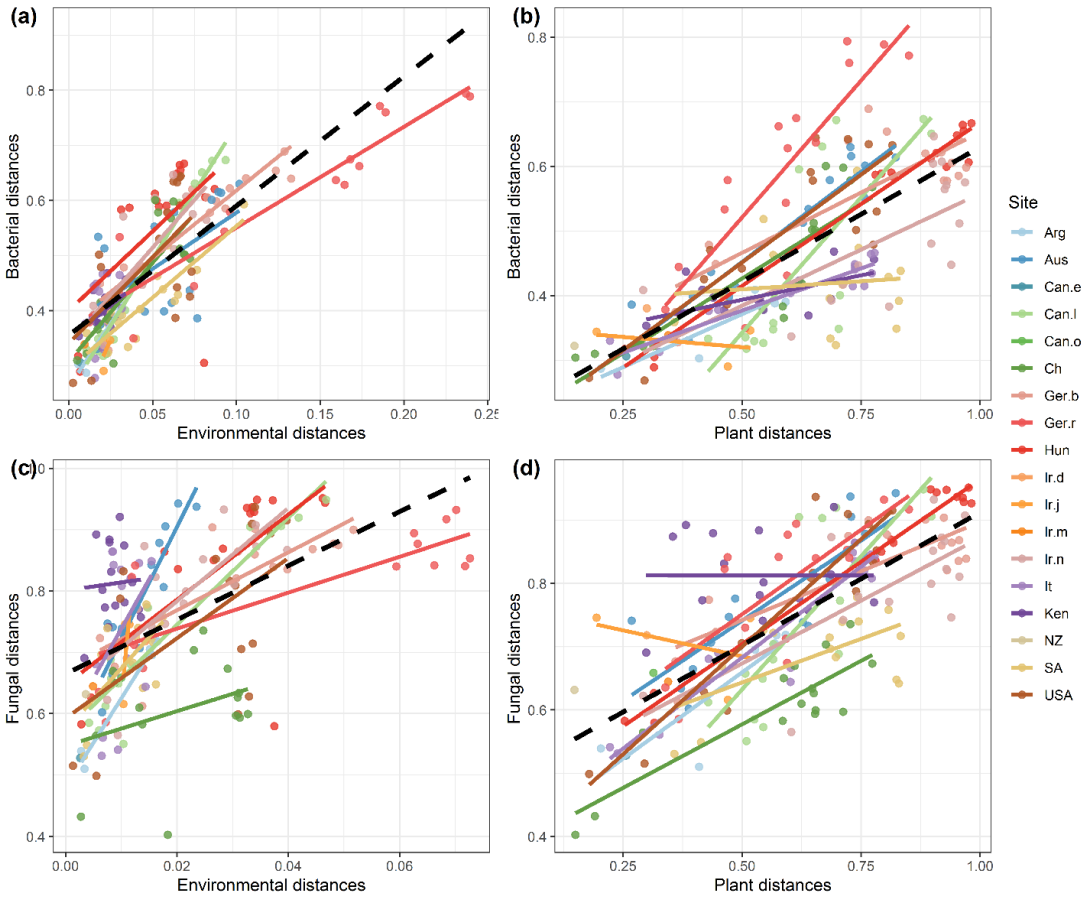


Figure 6.3 a) Within-site Euclidean environmental distances (pH, BS, N, CEC, sand and plant biomass weighted by their coefficient in the best MRM model and summed) against within-site bacterial BC distances; **b**) within-site plant BC distances against within-site bacterial BC distances; **c**) Within-sites Euclidean environmental distances (pH, C:N, CEC, sand and plant biomass weighted by their coefficient in the best MRM model and summed) and within-site fungal BC distances; **d**) within-site plant BC distances against within-site fungal BC distances. Colours of points and corresponding regression lines correspond to 18 different sites. Dashed lines represent general regression lines. The relationship between geographical distances and bacterial/fungal distances per site are shown in Figure S6.2.

Microbial community composition patterns at different plant productivity levels

Across sites, bacterial and fungal community composition differed significantly between low and high productivity levels ($P < 0.001$; Figure 6.4). Nonetheless, the predictors of bacterial and fungal community composition at low and high productivity level were similar to each other and to those selected for the entire dataset; i.e. environmental factors (particularly base saturation and pH) were the most important predictors of bacterial community composition, whereas fungal community composition was best predicted by plant community composition (Table S6.2, Appendix 6.1). Therefore, while distinct microbial communities were associated with different productivity levels, their relationship with plant communities and (a)biotic environment was largely consistent at different productivity levels.

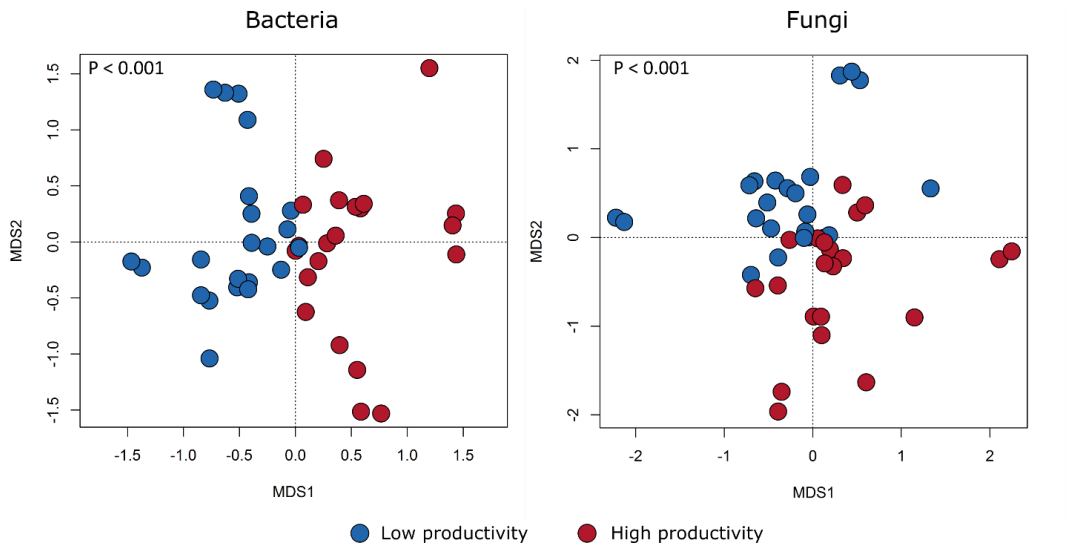


Figure 6.4 The MDS ordination showing bacterial and fungal BC distances (corrected for the effect of the site) at the high and the low productivity levels.

Dominant microbial groups and their correlations with plant groups/biomass and soil properties at different productivity levels

The most abundant bacterial phyla in the dataset were: Actinobacteria (42%), Firmicutes (16%), Proteobacteria (14%), Chloroflexi (8%), Acidobacteria (6%), Verrucomicrobia (3%), Bacteroidetes (3%), Planctomycetes (2%), and other phyla each contained less than 2% of the total number of bacterial sequences (Figure S6.3, Table S6.3). Saprotrophs were the most dominant fungal functional group with 54% of sequences followed by 14% of potential plant pathogens, 7% of arbuscular mycorrhizal fungi (AMF) and the other groups (such as lichens, ectomycorrhizal fungi, animal parasites, mycoparasites) together accounted for 4% of the total number of sequences (Figure S6.3).

When considering the three most dominant fungal and bacterial groups, at the high productivity level, there was a significantly higher relative abundance of Firmicutes and a lower relative abundance of Actinobacteria compared to the low productivity level (Figure 6.5b). The relative abundances of Proteobacteria, saprotrophs and AMF did not differ significantly between the two productivity levels.
Total fungal abundance was, however, significantly lower at the high productivity level, while bacterial abundance and fungal to bacterial ratio did not differ significantly.

The correlation networks between the three most dominant bacterial and fungal groups with plant functional groups, soil properties and total fungal and bacterial abundance differed strongly between two productivity levels. At high productivity, there were only a few correlations; e.g between C:N and both Actinobacteria and Proteobacteria, base saturation and bacterial/fungal abundance and a relationship between legumes and Firmicutes (Figure 6.5a). On the other hand, the number of associations was much higher at the low productivity level (Figure 6.5a) where different soil properties (including SOM, N, P and CEC) were related to fungal and bacterial groups. Moreover, there were negative correlations between putative plant pathogens and forbs as well as between Firmicuets and bacterial and fungal abundances.

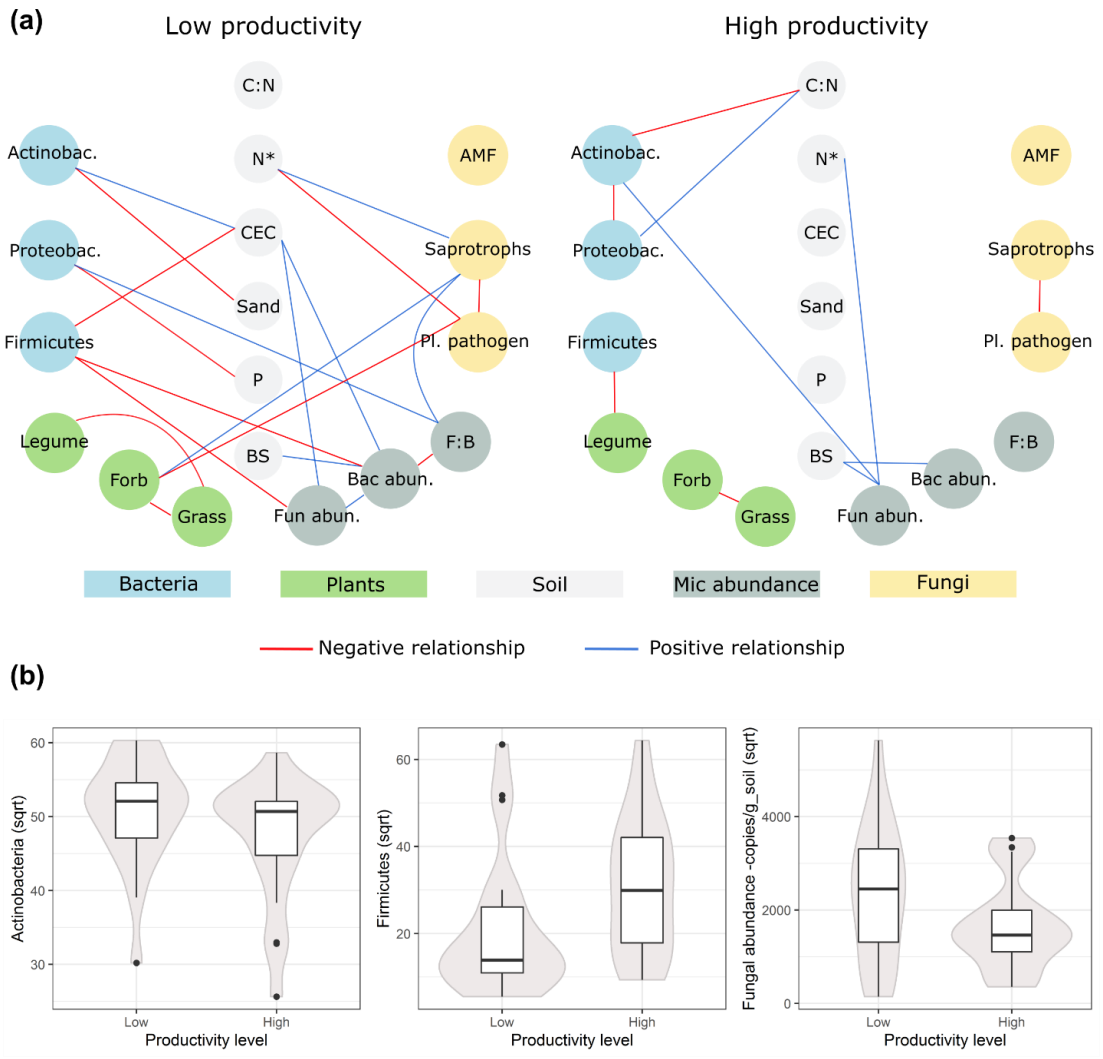


Figure 6.5 a) Correlation networks between three most dominant bacterial phyla (Actinobacteria, Firmicutes, Proteobacteria), three most dominant fungal functional groups (saprotrophs, putative plant pathogens, AMF), three main plant functional groups (grasses, forbs, legumes), total bacterial/fungal abundance (number of copies per g soil) and their ratio and soil properties at high and low productivity. Only the soil variables that had at least one significant correlation were shown. The red lines demonstrate significant negative correlations, while the blue lines demonstrate significant positive correlations ($P < 0.01$ and Spearman $r > 0.5$). Soil variables included C:N (carbon to nitrogen ratio), N (total nitrogen), CEC (cation exchange capacity), percentage sand, P (available phosphorus), BS (base saturation). N* = the same links as for total N were observed for total P and SOM, which were all strongly correlated to each and therefore only one of them was shown in the figure. The correlations between soil properties were not of interest and were therefore not included. **b**) Boxplots show the mean values of the three microbial variables that differed significantly between the productivity levels (Actinobacteria, Firmicutes and fungal abundances). The grey area depicts the distribution of samples.

6.5 Discussion

Despite a considerable amount of literature describing the most important predictors of microbial community composition in grassland soils, it remains unclear whether any general predictors across different spatial scales and environmental contexts can be identified or whether the context-dependency prevails, rendering microbial community composition largely unpredictable. In this study, we demonstrate that there is generality in the way bacterial and fungal communities are shaped across two spatial scales and two plant productivity levels in globally-distributed grasslands.

Generality in the predictors of microbial community composition at different spatial scales

The strongest predictors of bacterial community composition globally were soil abiotic factors, particularly base saturation and pH. While pH is known to be a major driver of bacterial community composition (Lauber *et al.*, 2008; Rousk *et al.*, 2010), the effect of base saturation is less often investigated, but was earlier also found to be one of the most important predictors of bacterial community composition across three different land-use types (Zheng *et al.*, 2019). On the other hand, plant community composition was the strongest predictor of fungal community composition, explaining more variation than all other predictors combined; a finding in line with that from another global-scale study in grasslands (Prober *et al.*, 2015). Similar patterns were found on the local scale, where environmental distances were consistently related with the variation in bacterial community composition across sites while fungal community composition had a more consistent relationship with plant community composition. These relationships are thus not just a matter of coincident community turnover at global scales between fungi and plants, but rather indicate a direct influence on each other and/or a high similarity in ecological niches.

Therefore, contrary to what has previously been found for bacterial communities in salt marsh sediments (Martiny *et al.*, 2011) and wheat fields (Shi *et al.*, 2018), root-associated fungal communities in salt marshes (Alzarhani *et al.*, 2019) and microbial communities in pastures along an elevation gradient (Chalmandrier *et al.*, 2019), our findings imply that the drivers of soil bacterial and fungal community composition in globally distributed grasslands do not strongly depend on the spatial scale (excluding broad-scale drivers, such as climate that only operate at large spatial scales) and that they are consistent across different grassland sites.

Moreover, our results overall indicate that plant community composition is strongly related to microbial community composition (particularly fungal, but also bacterial) across different grassland sites. Plant communities can affect soil microorganisms both directly by providing a diverse set of hosts for mutualistic and antagonistic microorganisms, and indirectly by altering edaphic factors and providing different quantity and quality of root exudates and litter (Wardle *et al.*, 2004; van der Heijden *et al.*, 2008; Berg & Smalla, 2009). For instance, productive, fast-growing plant communities create litter rich in labile carbon that is preferentially decomposed by acquisitive groups of microbes such as many heterotropihc bacteria and certain types of saprotrophic fungi (Marschner *et al.*, 2011) to the detriment of microbes specialized in the decomposition of more recalcitrant material (e.g. many other groups of fungi) which are predominantly associated with low-productive plant communities. Local experiments have confirmed that plant community composition can shape microbial communities (Schlatter *et al.*, 2015; Reese *et al.*, 2018; Heinen *et al.*, 2020) and that plant-soil biotic interactions might play a central role in microbial community assembly processes (Wubs *et al.*, 2019; Radujković *et al.*, 2020). It is thus likely that plant communities exert important control over bacterial and fungal community composition regardless of the spatial context.

Universal influence of plant productivity on microbial community composition

Consistent with the previous conclusions, we found that bacterial and fungal community composition differed significantly between low and high productivity levels when site-specific differences were taken into account. This suggests that plant productivity, as an integrator of a myriad of factors that are related to it (including soil fertility, plant diversity, plant-soil interactions (Craven *et al.*, 2016; Delgado-Baquerizo *et al.*, 2017; Guerrero-Ramírez *et al.*, 2019), etc.) shapes community composition in a predictable manner regardless of differences in climate and the variability of grassland types (e.g. xeric, mesic and hydric) within this study. Several experimental studies have previously suggested that soil microbial community composition often exhibits consistent responses to environmental changes across different systems. For instance, it has been shown that soil microbial communities responded in a consistent manner to land-use changes in tropical forests (Petersen *et al.*, 2019), to nutrient inputs in contrasting systems (Ramirez *et al.*, 2010) and nutrient inputs across different grasslands (Leff *et al.*, 2015). Our results imply that the drivers of overall soil microbial community composition in grassland soils are likely universal across different environmental contexts.

The differences in bacterial community composition between the two productivity levels are corroborated by a higher relative abundance of Firmicutes and lower relative abundance of Actinobacteria (two dominant bacterial phyla) at high compared to low productivity, with a similar total bacterial abundance between these productivity levels. Both Firmicutes (Ramirez *et al.*, 2010; Wakelin *et al.*, 2013; Yao *et al.*, 2014; Ling *et al.*, 2017) and Actinobacteria (Ramirez *et al.*, 2010; Wakelin *et al.*, 2013; Leff *et al.*, 2015) were previously found to increase

under elevated nutrient inputs suggesting that they might be more dominant in fertile soils with higher plant productivity. It is possible that in our study, Firmicutes were more strongly dependent on high soil fertility conditions than Actinobacteria; hence, the increased relative abundance of Actinobacteria at the low productivity level might be caused by a substantially lower abundance of Firmicutes.

As for fungal communities, the relative abundance of the three dominant functional groups (saprotrophs, AMF and potential plant-pathogens) did not differ significantly between productivity levels, indicating that the shift in the overall fungal community composition is a result of changes in OTU abundances within these functional groups. However, total fungal abundance was significantly higher at low compared to high productivity levels. Higher fungal abundance is often found in less fertile soils (Bardgett & McAlister, 1999; Innes *et al.*, 2004), which typically have low plant productivity, where fungi are favoured over bacteria as the predominant decomposers due to the higher recalcitrance of plant litter (Marschner *et al.*, 2011). Moreover, plant reliance upon, and allocation to AMF is often higher to provide sufficient P, N and other nutrients (Johnson *et al.*, 2013). We did not observe a significantly higher relative abundance of symbiotrophs at low compared to high productivity, but since total fungal abundance was higher at low productivity levels, the same should be true for symbiotrophs.

The correlation networks between microbial groups and the environment vary with plant productivity

The correlations among microbial groups and abundances and their associations with plant functional groups and soil properties differed markedly at different productivity levels and they were much more numerous at low compared to the high productivity level. For instance, at low productivity, the relative abundance of putative plant pathogens was negatively associated with the abundance of forbs and tended to increase with increasing grass abundance, possibly because grass species typically host a larger variety of pathogens than forbs (Heinen *et al.*, 2020). Also, their denser fine root system may offer less defence and higher abundance of substrate to pathogens than the less dense roots of forbs do (Laliberté *et al.*, 2015). The increase in the relative abundance of putative plant pathogens with decreasing total N and SOM, i.e. in the soils with lower fertility, could also (partly) be due to the increased relative abundance of saprotrophic fungi, which were negatively correlated with plant pathogens. At the high productivity level, plant pathogens and saprotrophs were not correlated with other groups of biota or with soil properties, presumably because nutrients were generally ample.

These examples suggest that, due to sufficient amounts of resources, microbial groups at high productivity might not be substantially affected by a further increase in resource availability and they might be forming fewer consistent interactions (symbiotic or competitive) with each other or with plant groups. This has been demonstrated in agricultural settings where soil fertilization reduced rhizosphere microbiome dependency on plant-derived carbon leading to

simpler plant-microbe associations (Ai *et al.*, 2015). Similarly, it has been shown that 150 years of soil fertilization has weakened the complexity of plant-microbiome networks in a managed grassland (Huang *et al.*, 2019). We here show that these tendencies also hold for natural grasslands on a global scale. Therefore, unlike for the overall bacterial and fungal community compositions, the drivers of higher bacterial taxonomic groups and fungal functional groups in grassland ecosystems may be strongly dependent on resource availability.

6.6 Conclusion

Several studies have argued that there are perhaps very few if any general drivers of microbial community composition suggesting that microbial communities are not structured in a predictable manner under different conditions. If that is the case, predictions on the fate of soil microbial communities and processes they drive under altered environmental conditions derived from one system or spatial scale cannot be extrapolated to another. Our findings, however, imply that although the drivers of microbial abundance and relative abundance of high taxonomic/functional groups might depend on the context, the factors that shape overall microbial community composition in grasslands operate in a consistent manner, regardless of the spatial scale, productivity levels or exposure to very different historical legacies and varying climatic conditions. Omitting some of the crucial drivers of microbial community assembly in the analyses may lead to apparent context-dependency and our study emphasises the need to consider the combination of plant community and edaphic factors. We thus argue that there is a substantial universality in the way in which microbial community assembly processes operate across different contexts in grassland systems. This has important implications for modelling of soil microbial community composition in global grasslands under environmental changes which is, according to the findings of this study, a feasible task.

General discussion

CHAPTER VII

CHAPTER VII

General discussion

7.1 Part I: Soil properties and nutrients as predictors of global grassland productivity

7.1.1 Predictors of grassland productivity – synthesis

Understanding the patterns of grassland productivity has long been an important topic of ecological research. While much is known about soil properties as drivers of plant productivity, the predictions of grassland productivity over broad spatial ranges have predominantly been based on climatic factors (Sala *et al.*, 1988; Huxman *et al.*, 2004; Hovenden *et al.*, 2014, 2019). The primary reason for this is that climatic data have been easier to collect than the data on soil properties. Moreover, it could be assumed that soil properties would not necessarily add much to the variation already explained by climate which is thought to be a principal factor affecting both soil properties and plant productivity. It is argued in this work that soil properties determining nutrient availability and overall soil fertility can contribute substantially to explaining global-scale patterns in grassland productivity on top of the effect of climatic predictors. Our two studies on predictors of grassland productivity both support and corroborate this hypothesis.

The role of soil nutrient availability in predicting global patterns of grassland biomass production

Chapter II (the study conducted within the NutNet experiment) enabled us to investigate the importance of soil properties as predictors of productivity patterns across a wide range of globally-distributed grassland sites focusing on factors determining nutrient availability: SOM, CEC, pH and concentrations of different nutrients (besides common macronutrients: N (C:N), P and K, other macronutrients Ca, Mg, Na, S, and micronutrients Fe, Mn, Zn, Cu and B were also included). We found that common fertility indicators – SOM and CEC, are important predictors of biomass productivity, but our findings also imply that much of the effect of SOM could be indirect through other soil properties and nutrients – CEC, C:N and particularly Zn. While Zn limitation is known to be widespread in agricultural systems (Alloway, 2009; Shukla *et al.*, 2014), Zn has rarely been investigated as a potentially limiting nutrient in (semi)natural grasslands (but see e.g. Fay et al. (2015) who demonstrated that a combination of different nutrients, Zn included, can limit productivity in different grasslands), nor has it been used a

predictor of global grassland productivity. In this study, we found indications that when N and P are ample, Zn might be an important limiting factor in grasslands.

Soil physicochemical properties as key predictors of local and global variability in grassland biomass production

Chapter III (the study conducted within the HerbDivNet experiment) was focused on analysing the role of soil properties, many of which are well-known indicators of soil fertility (Bünemann *et al.*, 2018), as predictors of both within-site (local-scale) and between-site (global-scale) plant productivity in worldwide distributed grasslands. The results demonstrated that the variation in soil properties (soil organic matter, percentage sand, calcium content and fine soil bulk density) consistently explained the local-scale variation in plant biomass across different grassland sites ($R^2 = 0.52$) suggesting that the effect of soil properties was not strongly dependent on climatic conditions. Interestingly, these same soil factors were selected as best predictors of global-scale variation in AGB, together with growing season precipitation ($R^2_{mean} = 0.77$). In this model, a large proportion of the variation was shared between precipitation and soil properties (25%), but soil properties could still explain 45% of non-shared variation compared to 7% of unique variation explained by precipitation. Even if all available climatic factors in this study (mean annual precipitation, growing season precipitation, mean annual temperature, potential evapotranspiration and aridity) would be included in the model with the four selected soil properties, climate would explain only 13% of unique variation compared to 38% explained by the soil properties (with 32% of shared variation). This is another line of evidence that soil properties are stronger predictors of plant productivity than climatic factors in the HerbDivNet dataset.

Comparison between the two studies and general conclusions

Both studies examined the role of soil properties on global biomass production, but their findings were not entirely aligned, possibly due to the differences in site locations, available soil data and experimental setups. For instance, while HerbDivNet experiment contained sites with local productivity gradients where the variation in biomass could not be explained by climate, the NutNet experiments contained no such gradient in productivity, but it contained more sites with better global distribution.

One notable difference between the results of the two studies was that CEC was positively associated with biomass in the NutNet study, while Ca (which was tightly correlated with CEC) had a negative effect in the model explaining biomass productivity in the HerbDivNet study. In general, CEC is a measure of soil fertility, given that it indicates the capacity of soil to exchange cations that are important for plant growth (Ross & Ketterings, 1995). Therefore, a positive relationship between CEC/Ca and plant productivity would be expected. However, the HerbDivNet dataset included a higher proportion of sites with calcareous soils (indicated by high CEC, pH and the amount of exchangeable Ca) than the NutNet dataset. For instance, 36% of HerbDivNet sites had CEC > 25 meq/100g which is considered to be high (Hazelton & Murphy, 2019) compared to 14% of NutNet sites. Calcareous soils are known to be relatively infertile, among others due to the low P availability and/or decreased plant uptake of soil micronutrients (FAO, 2020). Interestingly, Zn availability is shown to be particularly low in calcareous soils, given that calcium carbonate precipitates Zn and makes it unavailable to plants (Chen *et al.*, 2017). Therefore, low availability and uptake of Zn or other nutrients, such as P (Larcher, 2003) in Ca-rich soil could explain the negative effect of Ca (CEC) on biomass production in the HerbDivNet grasslands.

Moreover, while in the HerbDiveNet study SOM was identified as the best predictor of aboveground biomass, the results of NutNet study indicate that its effect may be indirect, through other nutrients and particularly through soil Zn. Indeed, SOM determines nutrient availability by retaining and providing different nutrients (Kononova, 1966; Oldfield *et al.*, 2018) and is known to play a key role in the availability and transformation of soil Zn (Chen *et al.*, 2017). Given that available Zn was not measured in the HerbDivNet study, it was not possible to test if it would also emerge as an important predictor in this dataset.

Finally, a common pattern that emerges in both studies is that soil properties and nutrients explained a significant portion of the variation in global grassland productivity in addition to the variation explained by climate. In fact, soil properties explained more variation than climate in both cases (if the shared variation between them would be attributed to climate: 45% vs 32% for soil properties vs climate in the HerbDivNet study and 32% vs 16% in the NutNet study).

Overall, these two studies identified the soil properties that can best explain the variation in plant productivity across globally-distributed grasslands (i.e. soil organic matter, soil texture, Ca as an indicator of calcareous soil, soil Zn and CEC) suggesting that they likely play a crucial role in driving global grassland productivity, even beyond the effect of climate. Taking these soil properties into account could thus significantly improve our capacity to understand global grassland productivity patterns and to predict the consequences of environmental changes on grassland ecosystems.

7.1.2 Limitations of the studies and future research

The studies presented here are probably the first to test a comprehensive set of soil properties as potential predictors of plant productivity across a wide range of grassland sites. However, like all syntheses based on observations, our studies also come with limitations that need to be tackled in complementary research efforts. First, the correlations between biomass productivity and certain soil properties do not necessarily imply causation. For instance, as discussed in Chapter III, SOM which was identified as one of the best predictors of plant productivity, can both affect plant productivity and be affected by it. However, our structural equation models demonstrated that SOM likely had a stronger influence on biomass than the other way around. Secondly, given that many soil properties are correlated to each other, the one that emerges as the best predictor of biomass production model might not necessarily be the strongest causal driver. For example, while our models suggest that the effect of SOM may be indirect through Zn, it is not possible to fully disentangle these two effects in an observational study.

Fertilization experiments are needed to demonstrate and fully unravel the effect of organic matter and different nutrients on plant productivity. While there have been numerous fertilization experiments in (semi)natural grasslands (Grogan & Chapin, 2000; Fay *et al.*, 2015; You *et al.*, 2017; van Dobben *et al.*, 2017; Liu *et al.*, 2018) there are very few that applied additions of individual macronutrients (other than main macronutrients N, P and K) and particularly micronutrients. This is because main macronutrients are widely recognized as the most limiting nutrients (Elser *et al.*, 2007; Harpole *et al.*, 2011) while other nutrients are expected to play a relatively minor individual role or to only have an effect in combination with other nutrients. However, one recent study, for example, found that boron was the main factor limiting productivity in Cerrado grassland in Brazil (Lannes *et al.*, 2020) demonstrating that one single micronutrient can limit productivity in natural grasslands. Our findings, which point to a potentially important role of Zn, can be used to guide future experimental studies. For instance, based on our results, experiments could test the role of SOM additions versus Zn addition (or other nutrients) to assess and potentially disentangle their effect on grassland productivity. Moreover, N/P fertilization could be combined with Zn fertilization to test if grasslands with ample N/P indeed become Zn limited. It could also be interesting to explore thresholds of Zn deficiency for soil microorganisms and interactions between microorganisms and plants in Zn deficient grassland soils, which received little attention thus far (and this is poorly explored even in agricultural systems (Khan & Joergensen, 2010, 2015)). For instance, it has been demonstrated that AMF could further reduce plant-available Zn in Zn deficient soils (Tran *et al.*, 2019).

Finally, our studies emphasize the importance of collecting and reporting the data on different soil properties and nutrients (particularly those identified as important predictors) in experiments in grasslands and other ecosystems. For instance, when conducting fertilization

experiments, in addition to reporting the response of plant productivity, it would be interesting to measure and correlate the concentrations of soil nutrients with plant biomass. If these different measurements would be more commonly reported in studies, it would help to create comprehensive datasets covering currently underrepresented regions (e.g. the southern hemisphere) and test the generality of the current findings through synthesis analyses.

7.1.3 Implications

The importance of soil in predicting grassland carbon budget under global change

Numerous studies have tried to estimate whether the changes in precipitation, CO_2 levels or atmospheric N deposition will lead to an increase or decrease in grassland productivity (Hovenden *et al.*, 2014, 2019; Stevens *et al.*, 2015; Chang *et al.*, 2017; Zheng *et al.*, 2019) and how this will be reflected in total grassland carbon balance (Scurlock & Hall, 1998). For instance, the impact of elevated CO_2 on grassland productivity might be strongly contingent on N availability (Reich *et al.*, 2006). It is thought that increased N fertilization through anthropogenic N deposition could alleviate N limitation (Stevens *et al.*, 2004) and promote the positive effect of elevated CO_2 on grassland productivity (Hovenden *et al.*, 2014). Indeed, Stevens et al. (2015) showed that already now, 16% of the variation in grassland productivity worldwide could be explained by N deposition and similar results were found in our study from the NutNet dataset. However, our findings also indicate that even if N limitation would be alleviated, overall plant productivity might not increase substantially if grassland productivity is or becomes limited by other nutrients (such as Zn) or by the amount of organic matter that affects multiple nutrient availability. For example, the fertilization study by Fay et al. (2015), already demonstrated that grasslands are often co-limited by multiple nutrients, and N additions alone were not able to increase productivity in several grasslands unless other nutrients were added. Besides nutrients, another important limiting factor for plant growth is water availability. The global-scale studies investigating the effect of precipitation changes on plant productivity, however, often neglect the fact the amount of plant-available water is not only determined by precipitation but also by soil texture and soil organic matter. Overlooking soil properties can thus lead to large uncertainties in the model predictions (Folberth *et al.*, 2016). Therefore, understanding the role of soil in driving global grassland productivity is essential to improve the predictions of future carbon-storing potential in grasslands.

Implications for the management of grassland biodiversity

Grasslands can sustain large biodiversity and it has recently been urged that old-growth grasslands, in particular, should become a conservation priority (Nerlekar & Veldman, 2020). The factors that influence grassland productivity can also have an important effect on

biodiversity, e.g. N deposition is expected to increase productivity by favouring fast-growing plant species on the detriment of slow-growing species leading to a decrease in overall plant diversity (Stevens *et al.*, 2004). However, it has been shown that increased amounts of N might cause a decrease in plant diversity only when P is not limiting (van Dobben *et al.*, 2017). Further research in unravelling limitations by N and other possibly important nutrients such as Zn could be used to inform grassland management efforts aiming to protect or restore grassland biodiversity. For instance, knowing which nutrients are limiting in particular grasslands would help focus the efforts on strictly controlling the amount of these nutrients to prevent a decline in plant diversity. Moreover, this would help to estimate which management strategies are sufficient to restore and conserve grassland biodiversity by reducing the nutrient levels: e.g. mowing or top-soil removal.

7.2 Part II – Soil microbial community assembly processes

The importance of soil microbes cannot be overstated. They underpin all Earth's biogeochemical cycles (Falkowski *et al.*, 2008) and perform functions that impact every component of an ecosystem, including humans (Widder *et al.*, 2016). While recent research, driven by the development of high-throughput sequencing technology, has significantly contributed to clarifying the mechanisms that determine microbial community assembly, many important questions still remain to be answered. In this work, we tackle two emerging questions: one concerning the role of different mechanisms structuring microbial community assembly **(Chapter V)** and the other examining whether there is generality in predictors of microbial community composition across different conditions **(Chapter VI)**.

7.2.1 Mechanisms driving the development of soil fungal community composition in restored heathlands

Dispersal constraints, environmental filters, and biotic interactions (Belyea & Lancaster, 1999; Lortie *et al.*, 2004) are thought to be the main processes structuring microbial community composition. However, there have been two opposing views on the importance of these processes: one assuming that biotic interactions only operate after environmental filtering has taken place (Belyea & Lancaster, 1999; Raevel *et al.*, 2013), while the other suggests that priority effects determine biotic interactions which in turn significantly mediate species' responses to the environment and determine the strength and extent of the environmental filter (Wisz *et al.*, 2013; Cadotte & Tucker, 2017; Aguilar-Trigueros *et al.*, 2017).

Our study in the heathland restoration experiment simultaneously investigated the importance of these three filters and demonstrated that fungal and plant communities from the soils inoculated by heathland soil and plant propagules formed similar links across differing abiotic conditions leading to the convergence of fungal community composition. We, therefore, argue that the early stage presence of heathland communities and the interactions they form can reinforce the development of a heathland system and alleviate the abiotic filter. If the system is exposed to natural dispersal, however, other incoming plant and fungal species establish their own, alternative interactions leading to an altered community trajectory that is more strongly influenced by the abiotic context.

Moreover, the relationships observed between plant communities and both bacterial and fungal communities in globally distributed grasslands suggest that plant-microbial interactions might be universally important in shaping microbial community composition, particularly in the case of fungi. While these relationships could partially result from common environmental drivers (shared environmental niches) and shared evolutionary histories, we show that plant community composition explained more variation in fungal community composition than the most important environmental factors and geographical distances.

Collectively, these findings support the theory that the influence of environmental filters, biotic interactions and historical legacies are not necessarily hierarchical and that microbial community assembly processes are much more complex than suggested by the common principle "the environment selects" (Figure 7.1) given that biotic interactions can alter the way in which the environment affects microbial community composition (Kraft *et al.*, 2015; Cadotte & Tucker, 2017; Aguilar-Trigueros *et al.*, 2017).

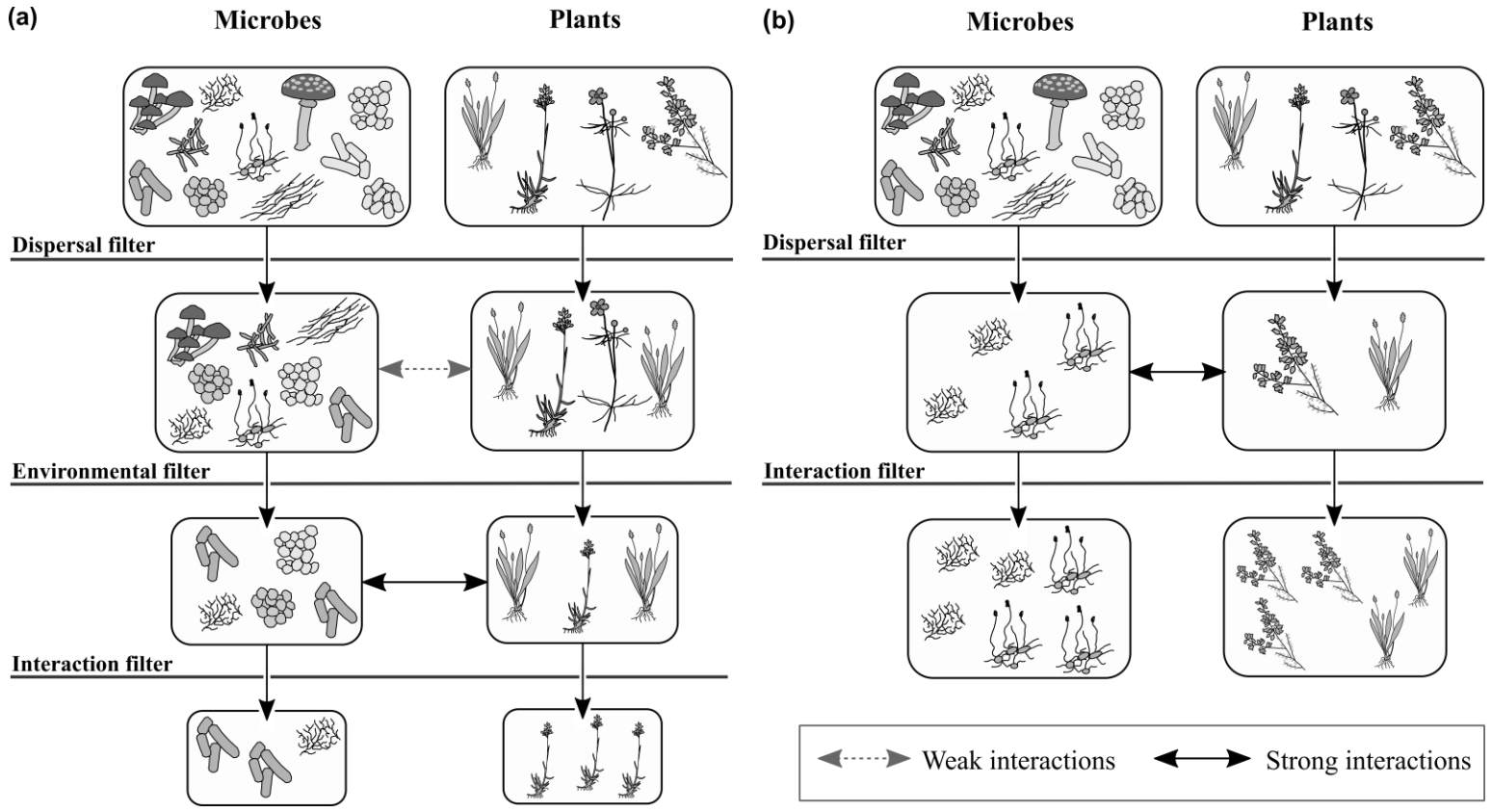


Figure 7.1 Conceptual model of processes structuring microbial and plant community assembly based on our study in heathlands. **a**) When different incoming microbial and plant species are not strongly associated and they form weak interactions, the environment (i.e. abiotic filters) selects the species that can thrive and form interactions to produce final communities. **b**) When incoming species are more strongly associated, they form stronger interactions from the beginning, reinforcing each other's development and overriding the environmental constraints. Drawing credit: Miguel Portillo-Estrada.

Implications and future research

The findings of this study have clear implications for the restoration of heathland systems. In that context, we argue that plant-soil interactions formed at the beginning of system development can have long-term legacies on the further development of the system and that inoculation with plant and soil material from well-developed heathlands is the fastest way to steer the restoration of heathland systems on bare soil. On the other hand, leaving the system to natural colonization by dispersal leads to slow development of plant and fungal communities with an uncertain outcome which might largely depend on the dominant surrounding vegetation and soil abiotic conditions.

Although heathlands are relatively simple systems, these mechanisms may apply to other systems, such as grasslands, as demonstrated in the soil inoculation experiment by Wubs et al. (2016). We suggest that similar inoculation experiments in different systems could reveal if plant-soil feedbacks are universally important in driving the early development of plant communities and their corresponding microbial communities. Longer-term monitoring would be needed to assess if the communities in the plots with and without inoculations would eventually converge or develop into completely distinct communities.

7.2.2 Generality in predictors of soil microbial community composition in grasslands

Context dependency (i.e. variability in the way processes operate under different conditions) appears to be common in nature (Maestre *et al.*, 2005; Chamberlain *et al.*, 2014; Tedersoo *et al.*, 2016; Song *et al.*, 2020) and it is thought to be a rule rather than the exception in community ecology (Lawton, 1999). Particularly in microbial ecology, with highly complex microbial communities and the interactions they form, it could be assumed the mechanisms that shape community composition are predominantly dependent on the context. Indeed, several studies investigating the drivers of microbial community composition under different contexts found a wide range of site- and spatial-scale specific predictors (Hendershot *et al.*, 2017; Alzarhani *et al.*, 2019; Gao *et al.*, 2019; Chalmandrier *et al.*, 2019) meaning that findings from one type of habitat or spatial scale cannot be successfully extrapolated to other environmental or spatial contexts. If there is no generality in drivers of microbial community composition, as these studies suggest, it is also impossible to predict the effect of environmental changes on microbial community composition.

Contrary to this, the results of our study in different grasslands across the globe suggest considerable generality in the way microbial communities are structured. More specifically, our results for two different spatial scales (local and global) and two different plant productivity levels (low and high) consistently indicated a dominant role of soil base saturation and pH on

bacterial community compositions at both spatial scales, as well as at two plant productivity levels. In the case of fungal communities, the strongest and the most consistent relationships were found with plant community composition. These patterns are not surprising since it has previously been demonstrated that, for instance, bacterial communities are more strongly influenced by abiotic factors such as pH than fungal communities (Cassman *et al.*, 2016; Zheng *et al.*, 2019). However, the validity of these findings under different contexts over large spatial ranges has never been demonstrated before. Moreover, our study is the first to demonstrate a consistent relationship between plant productivity and soil microbial community composition across grasslands worldwide, where total fungal abundance was consistently increased and the relative abundance of Firmicutes was significantly decreased at low compared to high productivity level.

While the "real" context-dependency of microbial community assembly processes clearly exists, as shown e.g. in our heathland experiment where abiotic filters played an important role in shaping fungal communities in the case where plant and soil communities were not simultaneously added (Figure 7.1) but had no influence when soil and plant propagules were present from the beginning (i.e. due to differences in historical legacies); in many studies, the "apparent" context dependency could occur because some important drivers of microbial community composition were not measured. Since we measured numerous parameters in our grassland study (including different soil properties, plant biomass and plant community composition), it was possible to identify general predictors of microbial community composition. We thus argue that, despite numerous processes that could produce context-dependency, such as historical legacies (Fukami, 2015), stochastic effects and ecological drift (Beck *et al.*, 2015) and dispersal limitation (Peay *et al.*, 2010), microbial communities can exhibit predictable patterns across different contexts.

Implications and future research

The majority of experimental studies, which are the only ones able to disentangle the mechanisms that underpin microbial community assembly processes, are conducted at one particular site and are rarely replicated under different contexts. If there would be strong context-dependency in drivers of microbial communities, the findings of each local-scale experiment would likely not be valid under different conditions (Alzarhani *et al.*, 2019). However, we argue in this work that there is generality in the way microbial community assembly processes operate across different grassland systems and that the findings from one study system in grasslands can likely be generalised and used to predict microbial community composition in other grassland systems. These findings have very important implications for modelling of grassland soil microbial community composition under environmental changes over larger spatial ranges.

Further research should be done to clarify whether this 'universality' in the way microbial communities are shaped holds also across contrasting systems: such as temperate and tropical forests, peatlands, heathlands, etc. As demonstrated with our grassland experiment, within-site gradients (e.g. soil fertility gradient) along which the changes in community composition can be estimated, are very valuable for testing both local-scale and global scale predictors. The observational studies would ideally be accompanied by experiments where the influence of soil abiotic factors and plant communities on microbial community assembly could be disentangled. These experiments could, for example, include manipulations of plant community composition under the same environmental conditions across different sites. Given that the grasslands included in our study were already very diverse, with different climates, hydrology, historical legacies, etc. it is expected that some general patterns would also be observed across different systems.

7.2.3 Limitations of the studies

As discussed previously, we cannot infer causality from the correlations between soil factors/plant communities and microbial community composition. For instance, plants can structure microbial communities by selecting for particular species (Bezemer *et al.*, 2010; Wubs & Bezemer, 2018), but soil microbial communities can in turn influence plants by providing limiting nutrients and helping seedlings to establish or by decreasing plants' fitness through parasitism (van der Heijden & Hartmann, 2016; Wubs *et al.*, 2016). Therefore, plant-microbe relationships are likely a result of complex feedback processes between them. This is also clear from our experimental study in heathlands where we concluded that plant-fungal interactions probably drive both plant and fungal community assembly. Nonetheless, the existence of the relationships between plant and microbial communities can indicate an important role of plants in shaping microbial community composition, either directly or through feedback loops.

Secondly, in our grassland study, it was not possible to fully disentangle the effect of the soil properties and plant community composition because much of the variation they explained was shared. This could be either because microbial and plant communities are both driven by the same factors (following the environmental filter theory) or because plant-microbe interactions change with the change in the environmental factors (Cadotte & Tucker, 2017). Disentangling the exact mechanisms requires experiments such as the one performed in the heathland study where it was demonstrated that fungal community composition was strongly driven by plant-soil interactions. However, even if the mechanisms that drive microbial community composition in grasslands were not fully disentangled, it was possible to discern important general patters; e.g. that soil abiotic factors affect bacterial community composition more strongly and more consistently than fungal, be it through pure environmental filtering or through a mix of community structuring mechanisms.

7.3 Conclusion

Plant productivity shapes microbial communities in soil which in turn can influence plant biomass production. This feedback loop between plant productivity and microbial communities is strongly affected by the balance of plant-derived organic matter inputs to the soil and degradation of organic matter to nutrients available for plant growth (van der Heijden *et al.*, 2008; Waldrop *et al.*, 2017) i.e. by the amount of soil resources. In this work, we showed that microbial groups at low plant productivity might be more dependent on soil resources and more tightly connected with plant functional groups than the communities at higher productivity levels which typically have access to ample resources. This could lead to a mutual reinforcement between plants at low productive grasslands (typically slow-growing, specialized for the acquisition of limited water and nutrients) and soil microbes with comparable traits (e.g. many fungi specialized in the decomposition of recalcitrant plant material and microbes forming symbiotic interactions with plants) resulting in microbial communities that are clearly distinct from those from high-productive grasslands. This is in line with the results by Delgado-Baquerizo et al. (2018) showing that there is a particular cluster of globally dominant soil bacterial taxa significantly associated with low-productivity regions around the globe. Based on our results, low-productive grasslands could be associated with increased overall fungal abundance while the members of Firmicutes bacterial phylum might be a good indicator of high productive grasslands.

These findings collectively suggest that considering key soil physicochemical properties determining soil fertility and nutrient availability (soil texture, soil organic matter, the amount of certain soil nutrients) when forecasting the effect of environmental changes on grassland productivity, would not only improve grassland productivity models but could also help predict the shifts in soil microbial community composition and/or microbial abundances. Further research is now needed to pinpoint which specific dominant taxa might be the best indicators of low vs high productive grasslands while examining physiological attributes of these taxa is crucial to comprehend microbial controls of soil processes and feedbacks to plant productivity. A better understanding of which soil microbial taxa and functions are associated with low vs high grassland productivity across the globe could help tackle long-standing questions regarding future C budget in grasslands; such as how environmental changes will alter the balance between productivity and decomposition and hence grassland C sequestration.

The tractability of this task is critically aided by the observation that some of the factors and interactions predicting both grassland plant productivity and microbial community composition were found to be universal across contrasting climates, meaning that predictions of processes going on in unmeasured sites are possible even if there is quite large environmental variation. Similar general patterns might also hold in other systems. While this remains to be explored, the universal patterns observed in this study provide hope that making general predictions

regarding future changes in plant productivity, soil microbial community composition and resulting changes in ecosystem functioning is a less daunting task than often assumed.

References

Abarenkov K, Henrik Nilsson R, Larsson K-H, Alexander IJ, Eberhardt U, Erland S, Høiland K, Kjøller R, Larsson E, Pennanen T, et al. 2010. The UNITE database for molecular identification of fungi - recent updates and future perspectives. *New Phytologist* 186: 281–285.

Ackerman D, Chen X, Millet D. 2018. Global nitrogen deposition (2°×2.5° grid resolution) simulated with GEOS-Chem for 1984-1986, 1994-1996, 2004-2006, and 2014-2016. *Retrieved from the Data Repository for the University of Minnesota*.

Adl SM, Habura A, Eglit Y. 2014. Amplification primers of SSU rDNA for soil protists. Soil Biology and *Biochemistry* 69: 328–342.

Agler MT, Ruhe J, Kroll S, Morhenn C, Kim S-T, Weigel D, Kemen EM. 2016. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLOS Biology* 14: e1002352.

Ågren GI, Wetterstedt JÅM, Billberger MFK. 2012. Nutrient limitation on terrestrial plant growth - modeling the interaction between nitrogen and phosphorus. *New Phytologist* 194: 953–960.

Aguilar-Trigueros CA, Rillig MC, Ballhausen M-B. 2017. Environmental filtering is a relic. A response to Cadotte and Tucker. *Trends in ecology & evolution* 32: 882–884.

Ai C, Liang G, Sun J, Wang X, He P, Zhou W, He X. 2015. Reduced dependence of rhizosphere microbiome on plant-derived carbon in 32-year long-term inorganic and organic fertilized soils. *Soil Biology and Biochemistry* 80: 70–78.

Alberti G, Vicca S, Inglima I, Belelli-Marchesini L, Genesio L, Miglietta F, Marjanovic H, Martinez C, Matteucci G, D'Andrea E, et al. 2015. Soil C:N stoichiometry controls carbon sink partitioning between above-ground tree biomass and soil organic matter in high fertility forests. *iForest - Biogeosciences and Forestry* 8: 195–206.

Albright MBN, Martiny JBH. 2018. Dispersal alters bacterial diversity and composition in a natural community. *ISME Journal* 12.

Allison SD, Martiny JBH. 2008. Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences of the United States of America* 105: 11512–11519.

Alloway BJ. 2008. Micronutrients and crop production: an introduction. In: *Micronutrient deficiencies in global crop production*. Dordrecht: Springer Netherlands, 1–39.

Alloway BJ. 2009. Soil factors associated with zinc deficiency in crops and humans. *Environmental Geochemistry and Health* 31: 537–548.

Alzarhani AK, Clark DR, Underwood GJC, Ford H, Cotton TEA, Dumbrell AJ. 2019. Are drivers of root-associated fungal community structure context specific? *ISME Journal* 13: 1330–1344.

Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26: 32–46.

Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL, Sanders NJ, Cornell H V., Comita LS, Davies KF, et al. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters* 14: 19–28.

Anderson MJ, Ellingsen KE, McArdle BH. 2006. Multivariate dispersion as a measure of beta diversity. *Ecology Letters* 9: 683–693.

Andreini C, Banci L, Bertini I, Rosato A. 2006. Zinc through the three domains of life. *Journal of Proteome Research* 5: 3173–3178.

Andrianarisoa KS, Zeller B, Dupouey JL, Dambrine E. 2009. Comparing indicators of N status of 50 beech stands (Fagus sylvatica L.) in northeastern France. *Forest Ecology and Management* 257: 2241–2253.

Archibold O. 1995. *Ecology of world vegetation*. New York: Chapman & Hali.

Baas Becking L. 1934. Geobiologie of inleiding tot de milieukunde. W. P. Van Stockum & Zoon, The Hague, the Netherlands.

Banerjee S, Kirkby CA, Schmutter D, Bissett A, Kirkegaard JA, Richardson AE. 2016. Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. *Soil Biology and Biochemistry* 97: 188–198.

Banerjee S, Schlaeppi K, van der Heijden MGA. 2018. Keystone taxa as drivers of microbiome structure and functioning. *Nature Reviews Microbiology* 16: 567–576.

Barberán A, Bates ST, Casamayor EO, Fierer N. 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME Journal* 6: 343–351.

Bardgett RD, McAlister E. 1999. The measurement of soil fungal:bacterial biomass ratios as an indicator of ecosystem self-regulation in temperate meadow grasslands. *Biology and Fertility of Soils* 29: 282–290.

Bartram AK, Lynch MDJ, Stearns JC, Moreno-Hagelsieb G, Neufeld JD. 2011. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end Illumina reads. *Applied and Environmental Microbiology* 77: 3846–3852.

Bashan A, Gibson TE, Friedman J, Carey VJ, Weiss ST, Hohmann EL, Liu YY. 2016. Universality of human microbial dynamics. *Nature* 534: 259–262.

Bauer JT, Mack KML, Bever JD. 2015. Plant-soil feedbacks as drivers of succession: evidence from remnant and restored tallgrass prairies. *Ecosphere* 6: art158.

Beck S, Powell JR, Drigo B, Cairney JWG, Anderson IC. 2015. The role of stochasticity differs in the assembly of soil- and root-associated fungal communities. *Soil Biology and Biochemistry* 80: 18–25.

Op De Beeck M, Lievens B, Busschaert P, Declerck S, Vangronsveld J, Colpaert J V. 2014. Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. *PloS one* 9: e97629.

Belyea LR, Lancaster J. 1999. Assembly rules within a contingent ecology. *Oikos* 86: 402–416.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*. Series B (Methodological) 57: 289–300.

Berg G, Smalla K. 2009. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiology Ecology* 68: 1–13.

Bever JD, Mangan SA, Alexander HM. 2015. Maintenance of plant species diversity by pathogens. *Annual Review of Ecology, Evolution, and Systematics* 46: 305–325.

Bezemer TM, Fountain MT, Barea JM, Christensen S, Dekker SC, Duyts H, Van Hal R, Harvey JA, Hedlund K, Maraun M, et al. 2010. Divergent composition but similar function of soil food webs of individual plants: Plant species and community effects. *Ecology* 91: 3027–3036.

van der Bij AU, Pawlett M, Harris JA, Ritz K, van Diggelen R. 2017. Soil microbial community assembly precedes vegetation development after drastic techniques to mitigate effects of nitrogen deposition. *Biological Conservation* 212: 476–483.

van der Bij AU, Weijters MJ, Bobbink R, Harris JA, Pawlett M, Ritz K, Benetková P, Moradi J, Frouz J, van Diggelen R. 2018. Facilitating ecosystem assembly: plant-soil interactions as a restoration tool. *Biological Conservation* 220: 272–279.

Blaalid R, Kumar S, Nilsson RH, Abarenkov K, Kirk PM, Kauserud H. 2013. ITS1 versus ITS2 as DNA metabarcodes for fungi. *Molecular Ecology Resources* 13: 218–224.

Blair J, Nippert J, Briggs J. 2014. Grassland ecology. In: *Ecology and the environment*. New York, NY: Springer New York, 389–423.

Blakemore LC, Searle PL, Daly B. K. 1987. Methods for chemical analysis of soils. *NewZealand Soil Bureau Scientific Report* 80.

Borer ET, Harpole WS, Adler PB, Lind EM, Orrock JL, Seabloom EW, Smith MD. 2014. Finding generality in ecology: A model for globally distributed experiments. *Methods in Ecology and Evolution* 5: 65–73.

Bot A, Benites J. 2005. The importance of soil organic matter: key to drought-resistant soil and sustained food production. In: *FAO Soils Bulletin, 80. Food & Agriculture Org*.

Bray, R.H. and Kurtz LT. 1945. Determination of total organic and available forms of phosphorus in soils. *Soil Science* 59: 39–45.

Brewer S, Techtmann SM, Mahmoudi N, Niang D, Pfiffner S, Hazen TC. 2015. Co-extraction of DNA and PLFA from soil samples. *Journal of Microbiological Methods* 115: 64–66.

Broadley M, Brown P, Cakmak I, Rengel Z, Zhao F. 2011. Function of nutrients: Micronutrients. In: *Marschner's Mineral Nutrition of Higher Plants: Third Edition*. Elsevier Inc., 191–248.

Brown IC. 1943. Rapid method of determining exchangeable hydrogen and total exchangeable bases of soils. *Soil Science* 56: 353–358.

Bruno JF, Stachowicz JJ, Bertness MD. 2003. Inclusion of facilitation into ecological theory. *Trends in Ecology & Evolution* 18: 119–125.

Bünemann EK, Bongiorno G, Bai Z, Creamer RE, De Deyn G, de Goede R, Fleskens L, Geissen V, Kuyper TW, Mäder P, et al. 2018. Soil quality – A critical review. *Soil Biology and Biochemistry* 120: 105–125.

Cadotte MW, Tucker CM. 2017. Should environmental filtering be abandoned? *Trends in Ecology & Evolution* 32: 429–437.

Cakmak I. 2000. Tansley Review No. 111 Possible roles of zinc in protecting plant cells from damage by reactive oxygen species. *New Phytologist* 146: 185–205.

Cakmak I. 2008. Enrichment of cereal grains with zinc: Agronomic or genetic biofortification? *Plant and Soil* 302: 1–17.

Cakmak I, Kalayci M, Kaya Y, Torun AA, Aydin N, Wang Y, Arisoy Z, Erdem H, Yazici A, Gokmen O, et al. 2010. Biofortification and localization of zinc in wheat grain. *Journal of Agricultural and Food Chemistry* 58: 9092–9102.

Calcagno V, Mazancourt C de. 2010. glmulti : An R Package for easy automated model selection with (Generalized) Linear Models. *Journal of Statistical Software* 34: 1–29.

Callahan BJ, McMurdie PJ, Holmes SP. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal* 11: 2639–2643.

Callaway RM, Walker LR. 1997. Competition and facilitation: A synthetic approach to interactions in plant communities. *Ecology* 78: 1958–1965.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pẽa AG, Goodrich JK, Gordon JI, et al. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7: 335–336.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America* 108: 4516–22.

Cassman NA, Leite MF, Pan Y, de Hollander M, van Veen JA, Kuramae EE. 2016. Plant and soil fungal but not soil bacterial communities are linked in long-term fertilized grassland. *Sci Rep* 6: 23680.

Castro HF, Classen AT, Austin EE, Norby RJ, Schadt CW. 2010. Soil microbial community responses to multiple experimental climate change drivers. *Applied and Environmental Microbiology* 76: 999–1007.

Cech PG, Kuster T, Edwards PJ, Olde Venterink H. 2008. Effects of herbivory, fire and N2-fixation on nutrient limitation in a humid African savanna. *Ecosystems* 11: 991–1004.

Chalmandrier L, Pansu J, Zinger L, Boyer F, Coissac E, Génin A, Gielly L, Lavergne S, Legay N, Schilling V, et al. 2019. Environmental and biotic drivers of soil microbial β-diversity across spatial and phylogenetic scales. *Ecography* 42: 2144–2156.

Chamberlain SA, Bronstein JL, Rudgers JA. 2014. How context dependent are species interactions? *Ecology Letters* 17: 881–890.

Chang J, Ciais P, Viovy N, Soussana JF, Klumpp K, Sultan B. 2017. Future productivity and phenology changes in European grasslands for different warming levels: Implications for grassland management and carbon balance. *Carbon Balance and Management* 12: 1–21.

Chapin FS. 1980. The mineral nutrition of wild plants. *Annual Review of Ecology and Systematics* 11: 233–260.

Chen Y, Barak P. 1982. Iron nutrition of plants in calcareous soils. *Advances in Agronomy* 35: 217–240.

Chen Y, Cui J, Tian X, Zhao A, Li M, Wang S, Li X, Jia Z, Liu K. 2017. Effect of straw amendment on soil Zn availability and ageing of exogenous water-soluble Zn applied to calcareous soil. *PLoS ONE* 12: e0169776.

Chen D, Mi J, Chu P, Cheng J, Zhang L, Pan Q, Xie Y, Bai Y. 2015. Patterns and drivers of soil microbial communities along a precipitation gradient on the Mongolian Plateau. *Landscape Ecology* 30: 1669–1682.

Chung H, Zak DR, Reich PB, Ellsworth D. 2007. Plant species richness, elevated CO2 , and atmospheric nitrogen deposition alter soil microbial community composition and function. *Global Change Biology* 13: 980–989.

Clemmensen KE, Bahr A, Ovaskainen O, Dahlberg A, Ekblad A, Wallander H, Stenlid J, Finlay RD, Wardle DA, Lindahl BD. 2013. Roots and associated fungi drive long-term carbon sequestration in boreal forest. *Science* 339: 1615–8.

Cleveland CC, Liptzin D. 2007. C:N:P stoichiometry in soil: Is there a 'Redfield ratio' for the microbial biomass. *Biogeochemistry* 85: 235–252.

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. 2014. Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research* 42: D633–D642.

Cotrufo MF, Wallenstein MD, Boot CM, Denef K, Paul E. 2013. The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: Do labile plant inputs form stable soil organic matter? *Global Change Biology* 19: 988–995.

Craine JM, Jackson RD. 2010. Plant nitrogen and phosphorus limitation in 98 North American grassland soils. *Plant and Soil* 334: 73–84.

Craven D, Isbell F, Manning P, Connolly J, Bruelheide H, Ebeling A, Roscher C, van Ruijven J, Weigelt A, Wilsey B, et al. 2016. Plant diversity effects on grassland productivity are robust to both nutrient enrichment and drought. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 371: 20150277-.

Daniel R. 2005. The metagenomics of soil. Nature reviews. *Microbiology* 3: 470–478.

Davidson EA, Janssens IA. 2006. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* 440: 165–173.

Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett RD, Maestre FT, Singh BK, Fierer N. 2018a. A global atlas of the dominant bacteria found in soil. *Science* 359: 320–325.

Delgado-Baquerizo M, Powell JR, Hamonts K, Reith F, Mele P, Brown M V., Dennis PG, Ferrari BC, Fitzgerald A, Young A, et al. 2017. Circular linkages between soil biodiversity, fertility and plant productivity are limited to topsoil at the continental scale. *New Phytologist* 215: 1186–1196.

Delgado-Baquerizo M, Reith F, Dennis PG, Hamonts K, Powell JR, Young A, Singh BK, Bissett A. 2018b. Ecological drivers of soil microbial diversity and soil biological networks in the Southern Hemisphere. *Ecology* 99: 583–596.

DeMalach N, Zaady E, Kadmon R. 2016. Light asymmetry explains the effect of nutrient enrichment on grassland diversity. *Ecology Letters*: 1–10.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and environmental microbiology* 72: 5069–72.

Díaz A, Green I, Evans D. 2011. Heathland restoration techniques: Ecological consequences for plant-soil and plant-animal interactions. *ISRN Ecology* 2011: 1–8.

Díaz A, Green I, Tibbett M. 2008. Re-creation of heathland on improved pasture using top soil removal and sulphur amendments: Edaphic drivers and impacts on ericoid mycorrhizas. *Biological Conservation* 141: 1628–1635.

Dickie IA, Fukami T, Wilkie JP, Allen RB, Buchanan PK. 2012. Do assembly history effects attenuate from species to ecosystem properties? A field test with wood-inhabiting fungi. *Ecology Letters* 15: 133–141.

Dimkpa CO, Bindraban PS. 2016. Fortification of micronutrients for efficient agronomic production: a review. *Agronomy for Sustainable Development* 36: 7.

van Dobben HF, Wamelink GWW, Slim PA, Kamiński J, Piórkowski H. 2017. Species-rich grassland can persist under nitrogen-rich but phosphorus-limited conditions. *Plant and Soil* 411: 451–466.

Drenovsky RE, Steenwerth KL, Jackson LE, Scow KM. 2010. Land use and climatic factors structure regional patterns in soil microbial communities. Global Ecology and *Biogeography* 19: 27–39.

Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* 10: 996–998.

Edgar RC, Flyvbjerg H. 2015. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 31: 3476–3482.

Eisenhauer N, Milcu A, Sabais ACW, Bessler H, Brenner J, Engels C, Klarner B, Maraun M, Partsch S, Roscher C, et al. 2011. Plant Diversity Surpasses Plant Functional Groups and Plant Productivity as Driver of Soil Biota in the Long Term. *PLoS ONE* 6: e16055.

Elser JJ, Bracken MES, Cleland EE, Gruner DS, Harpole WS, Hillebrand H, Ngai JT, Seabloom EW, Shurin JB, Smith JE. 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology Letters* 10: 1135–1142.

Encinas-Viso F, Alonso D, Klironomos JN, Etienne RS, Chang ER. 2016. Plant-mycorrhizal fungus co-occurrence network lacks substantial structure. *Oikos* 125: 457–467.

Englander L, Hull RJ. 1980. Reciprocal transfer of nutrients between ericaceous plants and A Clavaria SP. *New Phytologist* 84: 661–667.

Erenoglu EB, Kutman UB, Ceylan Y, Yildiz B, Cakmak I. 2011. Improved nitrogen nutrition enhances root uptake, root-to-shoot translocation and remobilization of zinc (65Zn) in wheat. *New Phytologist* 189: 438–448.

Fageria NK, Baligar VC, Clark RB. 2002. Micronutrients in crop production. In: *Advances in Agronomy*. Academic Press, 185–268.

Falkowski PG, Fenchel T, Delong EF. 2008. The microbial engines that drive Earth's biogeochemical cycles. *Science* 320: 1034–1039.

FAO. 2020. Calcareous soils. Food and Agriculture Organization of the United Nations. *FAO Soils Bulletin* 21.

Fay PA, Carlisle JD, Knapp AK, Blair JM, Collins SL. 2003. Productivity responses to altered rainfall patterns in a C4-dominated grassland. *Oecologia* 137: 245–251.

Fay PA, Kelley AM, Procter AC, Hui D, Jin VL, Jackson RB, Johnson HB, Polley HW. 2009. Primary productivity and water balance of grassland vegetation on three soils in a continuous CO_2 gradient: Initial results from the lysimeter CO_2 gradient experiment. *Ecosystems* 12: 699–714.

Fay PA, Prober SM, Stanley Harpole W, Knops J. MHJ., Bakker JD, Borer ET, Lind EM, MacDougall AS, Seabloom EW, Wragg PD, et al. 2015. Grassland productivity limited by multiple nutrients. *Nature Plants* 1: 15080.

Fernández-Illescas CP, Porporato A, Laio F, Rodríguez-Iturbe I. 2001. The ecohydrological role of soil texture in a water limited ecosystem. *Water resources research* 37: 2863–2872.

Fernández-Martínez M, Vicca S, Janssens IA, Sardans J, Luyssaert S, Campioli M, Chapin III FS, Ciais P, Malhi Y, Obersteiner M, et al. 2014. Nutrient availability as the key regulator of global forest carbon balance. *Nature Climate Change* 4: 471–476.

Fierer N, Jackson RB. 2006. The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America* 103: 626–31.

Folberth C, Skalský R, Moltchanova E, Balkovič J, Azevedo LB, Obersteiner M, Van Der Velde M. 2016. Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. *Nature Communications* 7: 1–13.

Frank DA, Kuns MM, Guido DR. 2002. Consumer control of grassland plant production. *Ecology* 83: 602–606.

Fraser LH, Henry H Al, Carlyle CN, White SR, Beierkuhnlein C, Cahill JF, Casper BB, Cleland E, Collins SL, Dukes JS, et al. 2013. Coordinated distributed experiments: An emerging tool for testing global hypotheses in ecology and environmental science. *Frontiers in Ecology and the Environment* 11: 147–155.

Fraser LH, Pither J, Jentsch A, Sternberg M, Zobel M, Askarizadeh D, Bartha S, Beierkuhnlein C, Bennett JA. 2015. Worldwide evidence of a unimodal relationship between productivity and plant species richness. *Science* 349: 302–306.

Frich P, Alexander L, Della-Marta P, Gleason B, Haylock M, Klein Tank A, Peterson T. 2002. Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research* 19: 193–212.

Frostegård Å, Tunlid A, Bååth E. 2011. Use and misuse of PLFA measurements in soils. *Soil Biology and Biochemistry* 43: 1621–1625.

Frostegård Å, Tunlid A, Bååth E, Frostegård A, Tunlid A, Bååth E. 1993. Phospholipid fatty acid composition, biomass, and activity of microbial communities from two soil types experimentally exposed to different heavy metals. *Applied and environmental microbiology* 59: 3605–17.

Fuhrman JA. 2009. Microbial community structure and its functional implications. Nature 459: 193–199.

Fukami T. 2015. Historical contingency in community assembly: Integrating niches, species pools, and priority effects. *Annual Review of Ecology, Evolution, and Systematics* 46: 1–23.

Fukami T, Martijn Bezemer T, Mortimer SR, Putten WH. 2005. Species divergence and trait convergence in experimental plant community assembly. *Ecology Letters* 8: 1283–1290.

Gao Q, Yang Y, Feng J, Tian R, Guo X, Ning D, Hale L, Wang M, Cheng J, Wu L, et al. 2019. The spatial scale dependence of diazotrophic and bacterial community assembly in paddy soil. *Global Ecology and Biogeography* 28: geb.12917.

Gardes M, Bruns TD. 1993. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Molecular Ecology* 2: 113–118.

Gee GW, Bauder JW. 1986. Particle-size analysis. In: In: Methods of soil analysis: Part 1 - Physical and mineralogical methods. *Soil Science Society of America, American Society of Agronomy*, 383–411.

Gerz M, Guillermo Bueno C, Ozinga WA, Zobel M, Moora M. 2018. Niche differentiation and expansion of plant species are associated with mycorrhizal symbiosis. *Journal of Ecology* 106: 254–264.

Gimingham CH. 1989. Heather and heathlands. *Botanical Journal of the Linnean Society* 101: 263–268.

Goslee SC, Urban DL. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software* 22: 1–19.

Grace JB, Anderson TM, Seabloom EW, Borer ET, Adler PB, Harpole WS, Hautier Y, Hillebrand H, Lind EM, Pärtel M, et al. 2016. Integrative modelling reveals mechanisms linking productivity and plant species richness. *Nature* 529: 390–393.

Grace JB, Anderson TM, Smith MD, Seabloom E, Andelman SJ, Meche G, Weiher E, Allain LK, Jutila H, Sankaran M, et al. 2007. Does species diversity limit productivity in natural grassland communities? *Ecology Letters* 10: 680–689.

Grogan P, Chapin FS. 2000. Nitrogen limitation of production in a Californian annual grassland: The contribution of arbuscular mycorrhizae. *Biogeochemistry* 49: 37–51.

Guerrero-Ramírez NR, Reich PB, Wagg C, Ciobanu M, Eisenhauer N. 2019. Diversity-dependent plant–soil feedbacks underlie long-term plant diversity effects on primary productivity. *Ecosphere* 10: e02704.

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, Boutte C, Burgaud G, de Vargas C, Decelle J, et al. 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic acids research* 41: 597–604.

Guo M, Gao G, Ding G, Zhang Y. 2020. Drivers of ectomycorrhizal fungal community structure associated with Pinus sylvestris var. mongolica differ at regional vs. local spatial scales in Northern China. *Forests* 11: 323.

Gupta N, Ram H, Kumar B. 2016. Mechanism of Zinc absorption in plants: uptake, transport, translocation and accumulation. *Reviews in Environmental Science and Biotechnology* 15: 89–109.

Hale VL, Tan CL, Knight R, Amato KR. 2015. Effect of preservation method on spider monkey (Ateles geoffroyi) fecal microbiota over 8 weeks. *Journal of Microbiological Methods* 113: 16–26.

Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH. 2012. Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology* 10: 497–506.

Harpole WS, Ngai JT, Cleland EE, Seabloom EW, Borer ET, Bracken MES, Elser JJ, Gruner DS, Hillebrand H, Shurin JB, et al. 2011. Nutrient co-limitation of primary producer communities. *Ecology Letters* 14: 852–862.

Harrison JG, John Calder W, Shuman B, Alex Buerkle C. 2020. The quest for absolute abundance: The use of internal standards for DNA-based community ecology. *Molecular Ecology Resources*.

Harry M, Gambier B, Garnier-Sillam E. 2000. Soil conservation for DNA preservation for bacterial molecular studies. *European Journal of Soil Biology* 36: 51–55.

Havlin JL. 2004. Fertility. In: *Encyclopedia of soils in the environment*. Elsevier Inc., 10–19.

Hayden HL, Mele PM, Bougoure DS, Allan CY, Norng S, Piceno YM, Brodie EL, DeSantis TZ, Andersen GL, Williams AL, et al. 2012. Changes in the microbial community structure of bacteria, archaea and fungi in response to elevated CO2 and warming in an Australian native grassland soil. *Environmental Microbiology* 14: 3081–3096.

Hazelton P, Murphy B. 2019. *Interpreting Soil Test Results*. Collingwood: CSIRO publishing.

He D, Wang E. 2019. On the relation between soil water holding capacity and dryland crop productivity. *Geoderma* 353: 11–24.

Hector A, Schmid B, Beierkuhnlein C, Caldeira MC, Diemer M, Dimitrakopoulos PG, Finn JA, Freitas H, Giller PS, Good J, et al. 1999. Plant diversity and productivity experiments in European grasslands. *Science* 286: 1123–1127.

van der Heijden MGA, Bakker R, Verwaal J, Scheublin TR, Rutten M, Van Logtestijn R, Staehelin C. 2006. Symbiotic bacteria as a determinant of plant community structure and plant productivity in dune grassland. *FEMS Microbiology Ecology* 56: 178–187.

van der Heijden MGA, Bardgett RD, Van Straalen NM. 2008. The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters* 11: 269–310.

van der Heijden MGA, Hartmann M. 2016. Networking in the plant microbiome. *PLoS Biology* 14: e1002378.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T, Sanders IR, Wiemken A. 1998. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* 396: 69–72.

Heinen R, Hannula SE, De Long JR, Huberty M, Jongen R, Kielak A, Steinauer K, Zhu F, Bezemer TM. 2020. Plant community composition steers grassland vegetation via soil legacy effects. *Ecology Letters* 23: 973–982.

Heiri O, Lotter AF, Lemcke G. 2001. Loss on ignition as a method for estimating organic and carbonate content in sediments: Reproducibility and comparability of results. *Journal of Paleolimnology* 25: 101–110.

Hendershot JN, Read QD, Henning JA, Sanders NJ, Classen AT. 2017. Consistently inconsistent drivers of microbial diversity and abundance at macroecological scales. *Ecology* 98: 1757–1763.

Hepler PK. 2005. Calcium: A central regulator of plant growth and development. Plant Cell 17: 2142–2155.

Herlemann DP, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF. 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME journal* 5: 1571–9.

Heuscher SA, Brandt CC, Jardine PM. 2005. Using soil physical and chemical properties to estimate bulk density. *Soil Science Society of America Journal* 69: 51–56.

Hooper D, Coughlan J, Mullen M. 2008. Structural equation modelling: Guidelines for determining model fit. *Electronic Journal on Business Research Methods* 6: 53–60.

Hovenden MJ, Leuzinger S, Newton PCD, Fletcher A, Fatichi S, Lüscher A, Reich PB, Andresen LC, Beier C, Blumenthal DM, et al. 2019. Globally consistent influences of seasonal precipitation limit grassland biomass response to elevated CO2. *Nature Plants* 5: 167–173.

Hovenden MJ, Newton PCD, Wills KE. 2014. Seasonal not annual rainfall determines grassland biomass response to carbon dioxide. *Nature* 511: 583–586.

Hsu JS, Powell J, Adler PB. 2012. Sensitivity of mean annual primary production to precipitation. *Global Change Biology* 18: 2246–2255.

Huang R, McGrath SP, Hirsch PR, Clark IM, Storkey J, Wu L, Zhou J, Liang Y. 2019. Plant–microbe networks in soil are weakened by century-long use of inorganic fertilizers. *Microbial Biotechnology* 12: 1464–1475.

Huston M. 1995. *Biological diversity: The coexistence of species*. Cambridge University Press.

Huxman TE, Smith MD, Fay PA, Knapp AK, Shaw MR, Loik ME, Smith SD, Tissue DT, Zak JC, Weltzin JF, et al. 2004. Convergence across biomes to a common rain-use efficiency. *Nature* 429: 651–654.

Ihrmark K, Bödeker ITM, Cruz-Martinez K, Friberg H, Kubartova A, Schenck J, Strid Y, Stenlid J, Brandström-Durling M, Clemmensen KE, et al. 2012. New primers to amplify the fungal ITS2 region - evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiology Ecology* 82: 666–677.

Innes L, Hobbs PJ, Bardgett RD. 2004. The impacts of individual plant species on rhizosphere microbial communities in soils of different fertility. *Biology and Fertility of Soils* 40: 7–13.

Janda JM, Abbott SL. 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. *Journal of Clinical Microbiology* 45: 2761–2764.

Jobbágy EG, Sala OE, Paruelo JM. 2002. Patterns and controls of primary production in the patagonian steppe: a remote sensing approach. *Ecology* 83: 307–319.

Johnson NC, Angelard C, Sanders IR, Kiers ET. 2013. Predicting community and ecosystem outcomes of mycorrhizal responses to global change. *Ecology Letters* 16: 140–153.

Johnson NC, Rowland DL, Corkidi L, Allen EB. 2008. Plant winners and losers during grassland n-eutrophication differ in biomass allocation and mycorrhizas. *Ecology* 89: 2868–2878.

Jones JB. 1990. Universal soil extractants: Their composition and use. *Communications in Soil Science and Plant Analysis* 21: 1091–1101.

Kardol P, Cornips NJ, van Kempen MML, Bakx-Schotman JMT, van der Putten WH. 2007. Microbe-mediated plant–soil feedback causes historical contigency effects in plant community assembly. *Ecological Monographs* 77: 147–162.

Kardol P, Martijn Bezemer T, van der Putten WH. 2006. Temporal variation in plant-soil feedback controls succession. *Ecology Letters* 9: 1080–1088.

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. 2017. Climatologies at high resolution for the earth's land surface areas. *Scientific Data* 4: 170122.

Kauserud H, Kumar S, Brysting AK, Nordén J, Carlsen T. 2012. High consistency between replicate 454 pyrosequencing analyses of ectomycorrhizal plant root samples. *Mycorrhiza* 22: 309–315.

Kazenel MR, Debban CL, Ranelli L, Hendricks WQ, Chung YA, Pendergast TH, Charlton ND, Young CA, Rudgers JA, Rudgers JA. 2015. A mutualistic endophyte alters the niche dimensions of its host plant. *AoB PLANTS* 7: plv005.

Kennedy PG, Cline LC, Song Z. 2018. Probing promise versus performance in longer read fungal metabarcoding. *New Phytologist* 217: 973–976.

Kerley SJ, Read DJ. 1998. The biology of mycorrhiza in the Ericaceae. XX. Plant and mycorrhizal necromass as nitrogenous substrates for the ericoid mycorrhizal fungus Hymenoscyphus ericae and its host. *New Phytologist* 139: 353–360.

Khan KS, Joergensen RG. 2010. Effects of Zn and P addition on the microbial biomass in a Zn deficient calcareous soil amended with glucose. *Plant and Soil* 335: 493–499.

Khan KS, Joergensen RG. 2015. Response of white mustard (Sinapis alba) and the soil microbial biomass to P and Zn addition in a greenhouse pot experiment. *Journal of Plant Nutrition and Soil Science* 178: 834–840.

Kirkham MB. 2005. *Principles of Soil and Plant Water Relations*. Elsevier.

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glockner FO. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research* 41: e1.

Kõljalg U, Larsson K-H, Abarenkov K, Nilsson RH, Alexander IJ, Eberhardt U, Erland S, Høiland K, Kjøller R, Larsson E, et al. 2005. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. *New Phytologist* 166: 1063–1068.

Kononova M. 1966. *Soil Organic Matter*. Elsevier.

Koziol L, Bever JD. 2017. The missing link in grassland restoration: arbuscular mycorrhizal fungi inoculation increases plant diversity and accelerates succession. *Journal of Applied Ecology* 54: 1301–1309.

Kraft NJB, Adler PB, Godoy O, James EC, Fuller S, Levine JM. 2015. Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology* 29: 592–599.

Krüger M, Stockinger H, Krüger C, Schüssler A. 2009. DNA-based species level detection of Glomeromycota: one PCR primer set for all arbuscular mycorrhizal fungi. *New Phytologist* 183: 212–23.

Kryza M, Dore AJ, Błaś M, Sobik M. 2011. Modelling deposition and air concentration of reduced nitrogen in Poland and sensitivity to variability in annual meteorology. *Journal of Environmental Management* 92: 1225–1236.

Kuffner M, Hai B, Rattei T, Melodelima C, Schloter M, Zechmeister-Boltenstern S, Jandl R, Schindlbacher A, Sessitsch A. 2012. Effects of season and experimental warming on the bacterial community in a temperate mountain forest soil assessed by 16S rRNA gene pyrosequencing. *FEMS Microbiology Ecology* 82: 551–562.

Kutman UB, Yildiz B, Ozturk L, Cakmak I. 2010. Biofortification of durum wheat with zinc through soil and foliar applications of nitrogen. *Cereal Chemistry* 87: 1–9.

Laliberté E, Lambers H, Burgess TI, Wright SJ. 2015. Phosphorus limitation, soil-borne pathogens and the coexistence of plant species in hyperdiverse forests and shrublands. *New Phytologist* 206.

Lannes LS, Bustamante MMC, Edwards PJ, Olde Venterink H. 2016. Native and alien herbaceous plants in the Brazilian Cerrado are (co-)limited by different nutrients. *Plant and Soil* 400: 231–243.

Lannes LS, Olde Venterink H, Leite MR, Silva JN, Oberhofer M. 2020. Boron application increases growth of Brazilian Cerrado grasses. *Ecology and Evolution*.

Larcher W. 2003. *Physiological plant ecology (fourth edition)*. Berlin, Heidelberg, New York: Springer-Verlag.

Lauber CL, Hamady M, Knight R, Fierer N. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and environmental microbiology* 75: 5111–20.

Lauber CL, Strickland MS, Bradford MA, Fierer N. 2008. The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biology and Biochemistry* 40: 2407–2415.

Lauenroth WK, Wade AA, Williamson MA, Ross BE, Kumar S, Cariveau DP. 2006. Uncertainty in calculations of net primary production for grasslands. *Ecosystems* 9: 843–851.

Lawson CS, Ford MA, Mitchley J, Warren JM. 2004. The establishment of heathland vegetation on ex-arable land: the response of Calluna vulgaris to soil acidification. *Biological Conservation* 116: 409–416.

Lawton JH. 1999. Are there general laws in ecology? *Oikos* 84: 177–192.

LeBauer D, Treseder K. 2008. Nitrogen limitation of net primary productivity. Ecology 89: 371–379.

Lee J, Lee S, Young JPW. 2008. Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiology Ecology* 65: 339–349.

Leff JW, Jones SE, Prober SM, Barberán A, Borer ET, Firn JL, Harpole WS, Hobbie SE, Hofmockel KS, Knops JMH, et al. 2015. Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proceedings of the National Academy of Sciences of the United States of America* 112: 10967–10972.

Lehmann J, Kern D, Glaser B, Woods W. 2003. *Amazonian dark earths: Origin, properties, management*. Kluwer Academic Publishers.

Lehmann J, Schroth G. 2005. Nutrient leaching. In: Schroth G, Sinclair F, eds. *Trees, crops and soil fertility: Concepts and research methods*. Wallingford, Oxon: Center for Agriculture and Biosciences International (CABI) Publishing, 151–166.

Lekberg Y, Bever JD, Bunn RA, Callaway RM, Hart MM, Kivlin SN, Klironomos J, Larkin BG, Maron JL, Reinhart KO, et al. 2018. Relative importance of competition and plant-soil feedback, their synergy, context dependency and implications for coexistence. *Ecology Letters* 21: 1268–1281.

Leopold DR. 2016. Ericoid fungal diversity: Challenges and opportunities for mycorrhizal research. *Fungal Ecology* 24: 114–123.

Von Liebig J. 1840. Die organische chemie in ihrer anwendung auf agrikultur und physiologi. Braunschweig: Friedrich Vieweg Sohn.

Lindahl BD, Nilsson RH, Tedersoo L, Abarenkov K, Carlsen T, Kjøller R, Kõljalg U, Pennanen T, Rosendahl S, Stenlid J, et al. 2013. Fungal community analysis by high-throughput sequencing of amplified markers - a user's guide. *New Phytologist* 199: 288–299.

Ling N, Chen D, Guo H, Wei J, Bai Y, Shen Q, Hu S. 2017. Differential responses of soil bacterial communities to long-term N and P inputs in a semi-arid steppe. *Geoderma* 292: 25–33.

Liu X, Liang M, Etienne RS, Gilbert GS, Yu S. 2016. Phylogenetic congruence between subtropical trees and their associated fungi. *Ecology and Evolution* 6: 8412–8422.

Liu C, Liu Y, Guo K, Qiao X, Zhao H, Wang S, Zhang L, Cai X. 2018. Effects of nitrogen, phosphorus and potassium addition on the productivity of a karst grassland: Plant functional group and community perspectives. *Ecological Engineering* 117: 84–95.

Loneragan JF, Webb MJ. 1993. Interactions between zinc and other nutrients affecting the growth of plants. In: *Zinc in soils and plants*. Springer Netherlands, 119–134.

Lortie CJ, Brooker RW, Choler P, Kikvidze Z, Michalet R, Pugnaire FI, Callaway RM. 2004. Rethinking plant community theory. *Oikos* 107: 433–438.

Ma W, Liu Z, Wang Z, Wang W, Liang C, Tang Y, He JS, Fang J. 2010. Climate change alters interannual variation of grassland aboveground productivity: Evidence from a 22-year measurement series in the Inner Mongolian grassland. *Journal of Plant Research* 123: 509–517.

Maestre FT, Valladares F, Reynolds JF. 2005. Is the change of plant-plant interactions with abiotic stress predictable? A meta-analysis of field results in arid environments. *Journal of Ecology* 93: 748–757.

Mahé F, de Vargas C, Bass D, Czech L, Stamatakis A, Lara E, Singer D, Mayor J, Bunge J, Sernaker S, et al. 2017. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nature Ecology & Evolution* 1: 0091.

Mangan SA, Schnitzer SA, Herre EA, Mack KML, Valencia MC, Sanchez EI, Bever JD. 2010. Negative plant–soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* 466: 752–755.

Männistö MK, Tiirola M, Häggblom MM. 2007. Bacterial communities in Arctic fjelds of Finnish Lapland are stable but highly pH-dependent. *FEMS microbiology ecology* 59: 452–65.

Marcon E, Herault B. 2015. entropart: An R package to measure and partition diversity. *Journal of Statistical Software* 67: 1–26.

Maron JL, Marler M, Klironomos JN, Cleveland CC. 2011. Soil fungal pathogens and the relationship between plant diversity and productivity. *Ecology Letters* 14: 36–41.

Marrs RH, Snow CSR, Owen KM, Evans CE. 1998. Heathland and acid grassland creation on arable soils at Minsmere: identification of potential problems and a test of cropping to impoverish soils. *Biological Conservation* 85: 69–82.

Marschner P, Umar S, Baumann K. 2011. The microbial community composition changes rapidly in the early stages of decomposition of wheat residue. *Soil Biology and Biochemistry* 43: 445–451.

Martiny JBH, Eisen JA, Penn K, Allison SD, Horner-Devine MC. 2011. Drivers of bacterial β-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences of the United States of America* 108: 7850–7854.

Mehlich A. 1984. Mehlich 3 soil test extractant: A modification of Mehlich 2 extractant. *Communications in Soil Science and Plant Analysis* 15: 1409–1416.

Miechówka A, Drewnik M. 2018. Rendzina soils in the Tatra mountains, central Europe : a review. *Soil Science Annual* 69: 88–100.

Millard P, Singh BK. 2010. Does grassland vegetation drive soil microbial diversity? *Nutrient Cycling in Agroecosystems* 88: 147–158.

Moreno-Jiménez E, Plaza C, Saiz H, Manzano R, Flagmeier M, Maestre FT. 2019. Aridity and reduced soil micronutrient availability in global drylands. *Nature Sustainability* 2: 371–377.

Morriën E, Hannula SE, Snoek LB, Helmsing NR, Zweers H, de Hollander M, Soto RL, Bouffaud M-L, Buée M, Dimmers W, et al. 2017. Soil networks become more connected and take up more carbon as nature restoration progresses. *Nature Communications* 8: 14349.

Nemergut DR, Costello EK, Hamady M, Lozupone C, Jiang L, Schmidt SK, Fierer N, Townsend AR, Cleveland CC, Stanish L, et al. 2011. Global patterns in the biogeography of bacterial taxa. *Environmental Microbiology* 13: 135–144.

Nemergut DR, Townsend AR, Sattin SR, Freeman KR, Fierer N, Neff JC, Bowman WD, Schadt CW, Weintraub MN, Schmidt SK. 2008. The effects of chronic nitrogen fertilization on alpine tundra soil microbial communities: Implications for carbon and nitrogen cycling. *Environmental Microbiology* 10: 3093–3105.

Nerlekar AN, Veldman JW. 2020. High plant diversity and slow assembly of old-growth grasslands. *Proceedings of the National Academy of Sciences*: 201922266.

Nguyen NH, Song Z, Bates ST, Branco S, Tedersoo L, Menke J, Schilling JS, Kennedy PG. 2016. FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecology* 20: 241–248.

Nobel I, Slatyer R. 1977. Post-fire succession of plants in Mediterranean ecosystems. In: *Proceedings of the symposium on the environmental consequences of fire and fuel management in Mediterranean ecosystems*. Palo Alto, California, USA: United States Forest Service, 27–36.

Novozamsky I, Houba VJG, van Eck R, van Vark W. 1983. A novel digestion technique for multi-element plant analysis. *Communications in Soil Science and Plant Analysis* 14: 239–248.

Obrador A, Novillo J, Alvarez JM. 2003. Mobility and availability to plants of two zinc sources applied to a calcareous soil. *Soil Science Society of America Journal* 67: 564–572.

Oehri J, Schmid B, Schaepman-Strub G, Niklaus PA. 2017. Biodiversity promotes primary productivity and growing season lengthening at the landscape scale. *Proceedings of the National Academy of Sciences of the United States of America* 114: 10160–10165.

Oksanen J, et al. 2015. Vegan: Community ecology package. R package version 2.4-0.

Olde Venterink H, Van Der Vliet RE, Wassen MJ. 2001a. Nutrient limitation along a productivity gradient in wet meadows. *Plant and Soil* 234: 171–179.

Olde Venterink H, Wassen MJ, Belgers JDM, Verhoeven JTA. 2001b. Control of environmental variables on species density in fens and meadows: importance of direct effects and effects through community biomass. *Journal of Ecology* 89: 1033–1040.

Oldfield EE, Wood SA, Bradford MA. 2018. Direct effects of soil organic matter on productivity mirror those observed with organic amendments. *Plant and Soil* 423: 363–373.

Olsen S, Cole C, Watanabe F DL. 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. USDA Circular Nr 939, US Gov. Print. Office, Washington, D.C 939.

Öpik M, Davison J, Moora M, Zobel M. 2014. DNA-based detection and identification of Glomeromycota : the virtual taxonomy of environmental sequences. *Botany* 92: 135–147.

Orwin KH, Buckland SM, Johnson D, Turner BL, Smart S, Oakley S, Bardgett RD. 2010. Linkages of plant traits to soil properties and the functioning of temperate grassland. *Journal of Ecology* 98: 1074–1083.

Ovaskainen O, Tikhonov G, Norberg A, Guillaume Blanchet F, Duan L, Dunson D, Roslin T, Abrego N. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters* 20: 561–576.

Peay KG. 2016. The mutualistic niche: Mycorrhizal symbiosis and community dynamics. *Annual Review of Ecology, Evolution, and Systematics* 47: 143–164.

Peay KG. 2018. Timing of mutualist arrival has a greater effect on Pinus muricata seedling growth than interspecific competition. *Journal of Ecology* 106: 514–523.

Peay KG, Bidartondo MI, Elizabeth Arnold A. 2010. Not every fungus is everywhere: scaling to the biogeography of fungal-plant interactions across roots, shoots and ecosystems. *New Phytologist* 185: 878–882.

Petersen IAB, Meyer KM, Bohannan BJM. 2019. Meta-analysis reveals consistent bacterial responses to land use change across the tropics. *Frontiers in Ecology and Evolution* 7: 391.

Phoenix GK, Emmett BA, Britton AJ, Caporn SJM, Dise NB, Helliwell R, Jones L, Leake JR, Leith ID, Sheppard LJ, et al. 2012. Impacts of atmospheric nitrogen deposition: responses of multiple plant and soil parameters across contrasting ecosystems in long-term field experiments. *Global Change Biology* 18: 1197–1215.

Poisot T, Bever JD, Nemri A, Thrall PH, Hochberg ME. 2011. A conceptual framework for the evolution of ecological specialisation. *Ecology Letters* 14: 841–851.

Pons P, Latapy M. 2005. Computing communities in large networks using random walks (long version). arXiv:physics/0512106 [physics.soc-ph].

Prado-Fiedler R. 1990. On the relationship between precipitation amount and wet deposition of nitrate and ammonium. *Atmospheric Environment. Part A. General Topics* 24: 3061–3065.

Prober SM, Leff JW, Bates ST, Borer ET, Firn J, Harpole WS, Lind EM, Seabloom EW, Adler PB, Bakker JD, et al. 2015. Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecology Letters* 18: 85–95.

Prosser JI, Bohannan BJM, Curtis TP, Ellis RJ, Firestone MK, Freckleton RP, Green JL, Green LE, Killham K, Lennon JJ, et al. 2007. The role of ecological theory in microbial ecology. Nature reviews. Microbiology 5: 384–392.

van der Putten WH. 2017. Belowground drivers of plant diversity. *Science* 355: 134–135.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41: D590–D596.

R Core Team. 2015. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*.

Radujković D, Diggelen R, Bobbink R, Weijters M, Harris J, Pawlett M, Vicca S, Verbruggen E. 2020. Initial soil community drives heathland fungal community trajectory over multiple years through altered plant–soil interactions. *New Phytologist* 225: 2140–2151.

Radujković D, Verbruggen E, Sigurdsson BD, Leblans NIW, Janssens IA, Vicca S, Weedon JT. 2018. Prolonged exposure does not increase soil microbial community compositional response to warming along geothermal gradients. *FEMS Microbiology Ecology* 94: fix174.

Raevel V, Munoz F, Pons V, Renaux A, Martin A, Thompson JD. 2013. Changing assembly processes during a primary succession of plant communities on Mediterranean roadcuts. *Journal of Plant Ecology* 6: 19–28.

Ramirez KS, Geisen S, Morriën E, Snoek BL, van der Putten WH. 2018. Network analyses can advance above-belowground ecology. *Trends in Plant Science* 23: 759–768.

Ramirez K, Lauber C, Knight R, Bradford M, Fierer N. 2010. Consistent effects of nitrogen fertilization on soil bacterial communities in contrasting systems. *Ecology* 91: 3463–3470.

Rashid A, Ryan J. 2004. Micronutrient constraints to crop production in soils with mediterranean-type characteristics: A review. *Journal of Plant Nutrition* 27: 959–975.

Reed HE, Martiny JBH. 2007. Testing the functional significance of microbial composition in natural communities. *FEMS Microbiology Ecology* 62: 161–170.

Reed HE, Martiny JBH. 2013. Microbial composition affects the functioning of estuarine sediments. *ISME Journal* 7: 868–879.

Reese AT, Lulow K, David LA, Wright JP. 2018. Plant community and soil conditions individually affect soil microbial community assembly in experimental mesocosms. *Ecology and Evolution* 8: 1196–1205.

Reeuwijk L. 2002. Technical paper 09: Procedures for soil analysis (6th Ed). *International Soil Reference and Information Centre, Wageningen, The Netherlands*.

Reeves DW. 1997. The role of soil organic matter in maintaining soil quality in continuous cropping systems. *Soil and Tillage Research* 43: 131–167.

Reich PB, Hungate BA, Luo Y. 2006. Carbon-nitrogen interactions in terrestrial ecosystems in response to rising atmospheric carbon dioxide. *Annual Review of Ecology, Evolution, and Systematics* 37: 611–636.

Rinnan R, Michelsen A, Bååth E, Jonasson S. 2007. Fifteen years of climate change manipulations alter soil microbial communities in a subarctic heath ecosystem. *Global Change Biology* 13: 28–39.

Robe P, Nalin R, Capellano C, Vogel TM, Simonet P. 2003. Extraction of DNA from soil. *European Journal of Soil Biology* 39: 183–190.

Rochelle PA, Cragg BA, Fry JC, John Parkes R, Weightman AJ. 1994. Effect of sample handling on estimation of bacterial diversity in marine sediments by 16S rRNA gene sequence analysis. *FEMS Microbiology Ecology* 15: 215–225.

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4: e2584.

Rosenzweig M. 1968. Net Primary productivity of terrestrial communities: Prediction from climatological data on JSTOR. *The American Naturalist* 102: 67–74.

Rosling A, Cox F, Cruz-Martinez K, Ihrmark K, Grelet G-A, Lindahl BD, Menkis A, James TY. 2011. Archaeorhizomycetes: Unearthing an ancient class of ubiquitous soil fungi. *Science* 333: 876–879.

Rosling A, Timling I, Taylor DL. 2013. Archaeorhizomycetes: Patterns of distribution and abundance in soil. *Genomics of Soil-and Plant-Associated Fungi*, Springer, Berlin, Heidelberg, 333–349.

Ross DS, Ketterings Q. 1995. Recommended methods for determining soil cation exchange capacity. *Recommended soil testing procedures for the northeastern United States* 2: 62–70.

Rosseel Y. 2012. lavaan: An R package for structural equation modeling. *Journal of Statistical Software* 48: 1–36.

Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG, Knight R, Fierer N. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. *Isme J*. 4: 1340–1351.

Roux KH. 2009. Optimization and troubleshooting in PCR. *Cold Spring Harbor protocols*: pdb.ip66.

Sahrawat KL, Wani SP, Pardhasaradhi G, Murthy KVS. 2010. Diagnosis of secondary and micronutrient deficiencies and their management in rainfed agroecosystems: Case study from Indian semi-arid tropics. Communications in Soil Science and Plant Analysis 41: 346–360.

Sala OE, Gherardi LA, Reichmann L, Jobbágy E, Peters D. 2012. Legacies of precipitation fluctuations on primary production: Theory and data synthesis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 367.

Sala OE, Parton WJ, Joyce LA, Lauenroth WK. 1988. Primary production of the central grassland region of the United States. *Ecology* 69: 40–45.

Santamaria M, Fosso B, Consiglio A, De Caro G, Grillo G, Licciulli F, Liuni S, Marzano M, Alonso-Alemany D, Valiente G, et al. 2012. Reference databases for taxonomic assignment in metagenomics. *Briefings in Bioinformatics* 13: 682–695.

Sayer EJ, Oliver AE, Fridley JD, Askew AP, Mills RTE, Grime JP. 2017. Links between soil microbial communities and plant traits in a species-rich grassland under long-term climate change. *Ecology and evolution* 7: 855–862.

Schadt CW, Martin AP, Lipson DA, Schmidt SK, Young JPW, Lindahl BD, Menkis A, James TY. 2003. Seasonal dynamics of previously unknown fungal lineages in tundra soils. *Science* 301: 1359–1361.

Schlaeppi K, Bender SF, Mascher F, Russo G, Patrignani A, Camenzind T, Hempel S, Rillig MC, van der Heijden MGA. 2016. High-resolution community profiling of arbuscular mycorrhizal fungi. *New Phytologist* 212: 780–791.

Schlatter DC, Bakker MG, Bradeen JM, Kinkel LL. 2015. Plant community richness and microbial interactions structure bacterial communities in soil. *Ecology* 96: 134–142.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, et al. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology* 75: 7537–41.

Schmidt MWI, Torn MS, Abiven S, Dittmar T, Guggenberger G, Janssens IA, Kleber M, Kögel-Knabner I, Lehmann J, Manning DAC, et al. 2011. Persistence of soil organic matter as an ecosystem property. *Nature* 478: 49–56.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium FB, Fungal Barcoding Consortium Author List FBCA, Bolchacova E, et al. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America* 109: 6241–6.

Scurlock JMO, Hall DO. 1998. The global carbon sink: a grassland perspective. *Global Change Biology* 4: 229–233.

Scurlock JMO, Johnson K, Olson RJ. 2002. Estimating net primary productivity from grassland biomass dynamics measurements. *Global Change Biology* 8: 736–753.

Shi Y, Li Y, Xiang X, Sun R, Yang T, He D, Zhang K, Ni Y, Zhu YG, Adams JM, et al. 2018. Spatial scale affects the relative role of stochasticity versus determinism in soil bacterial communities in wheat fields across the North China Plain. *Microbiome* 6: 27.

Shi R, Zhang Y, Chen X, Sun Q, Zhang F, Römheld V, Zou C. 2010. Influence of long-term nitrogen fertilization on micronutrient density in grain of winter wheat (Triticum aestivum L.). *Journal of Cereal Science* 51: 165–170.

Shokralla S, Spall JL, Gibson JF, Hajibabaei M. 2012. Next-generation sequencing technologies for environmental DNA research. *Molecular Ecology* 21: 1794–1805.

Shukla AK, Tiwari PK, Prakash C. 2014. Micronutrients deficiencies vis-a-vis food and nutritional security of India. *Indian Journal of Fertilisers* 10: 94–112.

Sigurdsson BD, Leblans NIW, Dauwe S, Guðmundsdóttir E, Gundersen P, Gunnarsdóttir GE, Holmstrup M, Ilieva-Makulec K, Kätterer T, Marteinsdóttir B, et al. 2016. Geothermal ecosystems as natural climate change experiments: The ForHot research site in Iceland as a case study. *Icel.agric.sci* 29: 53–71.

Sillanpää M. 1982. Micronutrients and the nutrient status of soil: a global study. *FAO soils bulletin* 48.

Sillanpää M. 1990. Micronutrient assessment at the country level: an international study. *FAO soils bulletin***.**

Simard SW. 2009. The foundational role of mycorrhizal networks in self-organization of interior Douglas-fir forests. *Forest Ecology and Management* 258: S95–S107.

Smets W, Leff JW, Bradford MA, Mcculley RL, Lebeer S, Fierer N. 2016. A method for simultaneous measurement of soil bacterial abundances and community composition via 16S rRNA gene sequencing. *Soil Biology and Biochemistry* 96: 145–151.

Smith DP, Peay KG. 2014. Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *PLoS ONE* 9: e90234.

Smith SE, Read DJ. 2008. *Mycorrhizal symbiosis (3rd Ed.)*. London: Academic Press.

Song C, Von Ahn S, Rohr RP, Saavedra S. 2020. Towards a probabilistic understanding about the context-dependency of species interactions. *Trends in Ecology and Evolution* 35: 384–396.

Song SJ, Amir A, Metcalf JL, Amato KR, Xu ZZ, Humphrey G, Knight R. 2016. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. *mSystems* 1: e00021-16.

Squires V, Dengler J, Hua L, Feng H. 2018. *Grasslands of the world: diversity, management and conservation*. CRC Press.

Steidinger BS, Crowther TW, Liang J, Van Nuland ME, Werner GDA, Reich PB, Nabuurs G, de-Miguel S, Zhou M, Picard N, et al. 2019. Climatic controls of decomposition drive the global biogeography of forest-tree symbioses. *Nature* 569: 404–408.

Stevens CJ, Dise NB, Mountford JO, Gowing DJ. 2004. Impact of nitrogen deposition on the species richness of grasslands. *Science* 303: 1876–1879.

Stevens CJ, Lind EM, Hautier Y, Harpole WS, Borer ET, Hobbie S, Seabloom EW, Ladwig L, Bakker JD, Chu C, et al. 2015. Anthropogenic nitrogen deposition predicts local grassland primary production worldwide. *Ecology* 96: 1459–1465.

Stirzaker R, Passioura J, Wilms Y. 1996. Soil structure and plant growth : Impact of bulk density and biopores. *Plant and Soil* 185: 151–162.

Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner HW, Richards TA. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology* 19: 21–31.

Strickland MS, Lauber C, Fierer N, Bradford MA. 2009. Testing the functional significance of microbial community composition. *Ecology* 90: 441–451.

Sun J, Cheng GW, Li WP. 2013. Meta-analysis of relationships between environmental factors and aboveground biomass in the alpine grassland on the Tibetan Plateau. *Biogeosciences* 10: 1707–1715.

Van Sundert K, Horemans JA, Stendahl J, Vicca S. 2018. The influence of soil properties and nutrients on conifer forest growth in Sweden, and the first steps in developing a nutrient availability metric. *Biogeosciences* 15: 3475–3496.

Van Sundert K, Radujković D, Cools N, De Vos B, Etzold S, Fernández-Martínez M, Janssens IA, Merilä P, Peñuelas J, Sardans J, et al. 2019. Towards comparable assessment of the soil nutrient status across scales—Review and development of nutrient metrics. *Global Change Biology* 26: 392–409.

Tedersoo L, Bahram M, Cajthaml T, Põlme S, Hiiesalu I, Anslan S, Harend H, Buegger F, Pritsch K, Koricheva J, et al. 2016. Tree diversity and species identity effects on soil fungi, protists and animals are context dependent. *The ISME Journal* 10: 1–17.

Tedersoo L, Bahram M, Põlme S, Kõljalg U, Yorou NS, Wijesundera R, Ruiz LV, Vasco-Palacios AM, Thu PQ, Suija A, et al. 2014. Global diversity and geography of soil fungi. *Science* 346: 1052–1053.

Tedersoo L, Nilsson RH, Abarenkov K, Jairus T, Sadam A, Saar I, Bahram M, Bechem E, Chuyong G, Kõljalg U. 2010. 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist* 188: 291–301.

Thebault E, Fontaine C. 2010. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science* 329: 853–856.

Tilman D, Lehman CL, Thomson KT. 1997. Plant diversity and ecosystem productivity: Theoretical considerations. *Proceedings of the National Academy of Sciences of the United States of America* 94: 1857–1861.

Tilman D, Reich PB, Knops J, Wedin D, Mielke T, Lehman C. 2001. Diversity and productivity in a long-term grassland experiment. *Science* 294: 843–845.

Toju H, Peay KG, Yamamichi M, Narisawa K, Hiruma K, Naito K, Fukuda S, Ushio M, Nakaoka S, Onoda Y, et al. 2018. Core microbiomes for sustainable agroecosystems. *Nature Plants* 4: 247–257.

Torrez V, Ceulemans T, Mergeay J, de Meester L, Honnay O. 2016. Effects of adding an arbuscular mycorrhizal fungi inoculum and of distance to donor sites on plant species recolonization following topsoil removal. *Applied Vegetation Science* 19: 7–19.

Tran BTT, Cavagnaro TR, Watts-Williams SJ. 2019. Arbuscular mycorrhizal fungal inoculation and soil zinc fertilisation affect the productivity and the bioavailability of zinc and iron in durum wheat. *Mycorrhiza*: 1–13.

Tylianakis JM, Martínez-García LB, Richardson SJ, Peltzer DA, Dickie IA. 2018. Symmetric assembly and disassembly processes in an ecological network. *Ecology Letters* 21: 896–904.

van der Valk AG. 1981. Succession in wetlands: A gleasonian approach. *Ecology* 62: 688–696.

Vannette RL, Fukami T, Wootton T. 2014. Historical contingency in species interactions: Towards niche-based predictions. *Ecology letters* 17: 115–24.

Verbruggen E, Sheldrake M, Bainard LD, Chen B, Ceulemans T, De Gruyter J, Van Geel M. 2018. Mycorrhizal fungi show regular community compositions in natural ecosystems. *The ISME journal* 12: 380–385.

Vicca S, Luyssaert S, Peñuelas J, Campioli M, Chapin FS, Ciais P, Heinemeyer A, Högberg P, Kutsch WL, Law BE, et al. 2012. Fertile forests produce biomass more efficiently. *Ecology Letters* 15: 520–526.

Vicca S, Stocker BD, Reed S, Wieder WR, Bahn M, Fay PA, Janssens IA, Lambers H, Peñuelas J, Piao S, et al. 2018. Using research networks to create the comprehensive datasets needed to assess nutrient availability as a key determinant of terrestrial carbon cycling. *Environmental Research Letters* 13: 125006.

Vitousek PM. 2015. Complexity of nutrient constraints. *Nature Plants* 1: 1–2.

de Vries FT, Bloem J, van Eekeren N, Brusaard L, Hoffland E. 2007. Fungal biomass in pastures increases with age and reduced N input. *Soil Biology and Biochemistry* 39: 1620–1630.

de Vries FT, Griffiths RI, Bailey M, Craig H, Girlanda M, Gweon HS, Hallin S, Kaisermann A, Keith AM, Kretzschmar M, et al. 2018. Soil bacterial networks are less stable under drought than fungal networks. *Nature Communications* 9: 3033.

de Vries FT, Manning P, Tallowin JRB, Mortimer SR, Pilgrim ES, Harrison KA, Hobbs PJ, Quirk H, Shipley B, Cornelissen JHC, et al. 2012. Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities. *Ecology Letters* 15: 1230–1239.

Wagg C, Bender SF, Widmer F, van der Heijden MGA. 2014. Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proceedings of the National Academy of Sciences of the United States of America* 111: 5266–70.

Wakelin SA, Barratt BIP, Gerard E, Gregg AL, Brodie EL, Andersen GL, DeSantis TZ, Zhou J, He Z, Kowalchuk GA, et al. 2013. Shifts in the phylogenetic structure and functional capacity of soil microbial communities follow alteration of native tussock grassland ecosystems. *Soil Biology and Biochemistry* 57: 675–682.

Wałaszek K, Kryza M, Dore AJ. 2013. The impact of precipitation on wet deposition of sulphur and nitrogen compounds. *Ecological Chemistry and Engineering* S 20: 733–745.

Waldrop MP, Holloway JM, Smith DB, Goldhaber MB, Drenovsky RE, Scow KM, Dick R, Howard D, Wylie B, Grace JB. 2017. The interacting roles of climate, soils, and plant production on soil microbial communities at a continental scale. *Ecology* 98: 1957–1967.

Wang E, Cresswell H, Xu J, Jiang Q. 2009. Capacity of soils to buffer impact of climate variability and value of seasonal forecasts. *Agricultural and Forest Meteorology* 149: 38–50.

Wang C, Wang X, Liu D, Wu H, Lü X, Fang Y, Cheng W, Luo W, Jiang P, Shi J, et al. 2014. Aridity threshold in controlling ecosystem nitrogen cycling in arid and semi-Arid grasslands. *Nature Communications* 5: 1–8.

Wardle DA, Bardgett RD, Klironomos JN, Setälä H, Van Der Putten WH, Wall DH. 2004. Ecological linkages between aboveground and belowground biota. *Science* 304: 1629–1633.

Waring BG, Adams R, Branco S, Powers JS. 2016. Scale-dependent variation in nitrogen cycling and soil fungal communities along gradients of forest composition and age in regenerating tropical dry forests. *New Phytologist* 209: 845–854.

Webb NR. 2008. The traditional management of European heathlands. Journal of Applied Ecology 35: 987–990.

Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, et al. 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5: 27.

White TJ, Bruns T, W.T LSJ, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications* 18: 315–322.

White RP, Murray S, Rohweder M, World Resources Institute. 2000. Pilot analysis of global ecosystems: Grassland ecosystems. *World Resources Institute*.

Wickham H. 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York.

Widder S, Allen RJ, Pfeiffer T, Curtis TP, Wiuf C, Sloan WT, Cordero OX, Brown SP, Momeni B, Shou W, et al. 2016. Challenges in microbial ecology: building predictive understanding of community function and dynamics. *The ISME journal* 10: 2557–2568.

Wilson H, Johnson BR, Bohannan B, Pfeifer-Meister L, Mueller R, Bridgham SD. 2016. Experimental warming decreases arbuscular mycorrhizal fungal colonization in prairie plants along a Mediterranean climate gradient. *PeerJ* 4: e2083.

Winsley T, van Dorst JM, Brown M V., Ferrari BC. 2012. Capturing greater 16S rRNA gene sequence diversity within the domain bacteria. *Applied and Environmental Microbiology* 78: 5938–5941.

Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, Dormann CF, Forchhammer MC, Grytnes J-A, Guisan A, et al. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews* 88: 15–30.

Woodward FI, Diament AD. 1991. Functional approaches to predicting the ecological effects of global change. *Functional Ecology* 5: 202.

Wubs ERJ, Bezemer TM. 2018. Plant community evenness responds to spatial plant–soil feedback heterogeneity primarily through the diversity of soil conditioning. *Functional Ecology* 32: 509–521.

Wubs ERJ, Putten WH, Mortimer SR, Korthals GW, Duyts H, Wagenaar R, Bezemer TM. 2019. Single introductions of soil biota and plants generate long-term legacies in soil and plant community assembly. *Ecology Letters* 22: 13271.

Wubs ERJ, van der Putten WH, Bosch M, Bezemer TM. 2016. Soil inoculation steers restoration of terrestrial ecosystems. *Nature Plants* 2: 16107.

Xia LC, Steele JA, Cram JA, Cardon ZG, Simmons SL, Vallino JJ, Fuhrman JA, Sun F. 2011. Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. *BMC systems biology 5(Suppl 2)*: S15.

Yang YH, Fang JY, Pan YD, Ji CJ. 2009. Aboveground biomass in Tibetan grasslands. *Journal of Arid Environments* 73: 91–95.

Yang G, Yang X, Zhang W, Wei Y, Ge G, Lu W, Sun J, Liu N, Kan H, Shen Y, et al. 2016. Arbuscular mycorrhizal fungi affect plant community structure under various nutrient conditions and stabilize the community productivity. *Oikos* 125: 576–585.

Yao M, Rui J, Li J, Dai Y, Bai Y, Heděnec P, Wang J, Zhang S, Pei K, Liu C, et al. 2014. Rate-specific responses of prokaryotic diversity and structure to nitrogen deposition in the Leymus chinensis steppe. *Soil Biology and Biochemistry* 79: 81–90.

You C, Wu F, Gan Y, Yang W, Hu Z, Xu Z, Tan B, Liu L, Ni X. 2017. Grass and forbs respond differently to nitrogen addition: a meta-analysis of global grassland ecosystems. Scientific reports 7: 1563.

Zak D, Holmes WE, White DC, Peacock AD, Ecology S, Aug N. 2014. Plant diversity, soil microbial communities, and ecosystem function: Are there any links? *Ecology* 84: 2042–2050.

Zhang Y, Dong S, Gao Q, Liu S, Zhou H, Ganjurjav H, Wang X. 2016. Climate change and human activities altered the diversity and composition of soil microbial community in alpine grasslands of the Qinghai-Tibetan plateau. *Science of The Total Environment* 562: 353–363.

Zhao X, Yang Y, Shen H, Geng X, Fang J. 2019. Global soil–climate–biome diagram: linking surface soil properties to climate and biota. *Biogeosciences* 16: 2857–2871.

Zheng Q, Hu Y, Zhang S, Noll L, Böckle T, Dietrich M, Herbold CW, Eichorst SA, Woebken D, Richter A, et al. 2019. Soil multifunctionality is affected by the soil environment and by microbial community composition and diversity. *Soil Biology and Biochemistry* 136: 107521.

Zijlstra JD, Van't Hof P, Baar J, Verkley GJM, Paradi I, Braakhekke WG, Berendse F. 2005. Diversity of symbiotic root endophytes of the Helotiales in ericaceous plants and the grass, Deschampsia flexuosa. *Studies in Mycology* 53: 147–162.

Zogg GP, Zak DR, Ringelberg DB, White DC, MacDonald NW, Pregitzer KS. 1997. Compositional and functional shifts in microbial communities due to soil warming. *Soil Science Society of America Journal* 61: 475–481.

Zomer RJ, Trabucco A, Bossio DA, Verchot L V. 2008. Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture, Ecosystems & Environment* 126: 67–80.

Author contributions

Chapters II – VI of this thesis are based on manuscripts (or published papers).
My contribution to these manuscripts was as follows:

Chapter II: Data collection was done by the NutNet members. I analysed and interpreted the data with the help of Sara Vicca (SV) and Erik Verbruggen (EV). I wrote the first draft of the manuscript with the input by SV and EV. The manuscript was further revised by Kevin Van Sundert, Michael Bahn, Matteo Campioli and the members of the NutNet team.

Chapter III: I collected part of the soil samples and carried out lab analyses, statistical analyses and data interpretation with the help of SV and EV. I drafted the first version of the manuscript with the input by SV, while EV, Ivan Janssens and Kevin Van Sundert contributed to the final version.

Chapter IV: I wrote the protocol together with EV.

Chapter V: I collected part of the soil samples and performed lab analyses. I performed statistical analyses and data interpretation with the help of EV and SV. I wrote the first draft of the manuscript with the help of EV and all co-authors contributed to the final version.

Chapter VI: I collected part of the soil samples for the study and carried out lab analyses. I performed statistical analyses and data interpretation with the help of EV. I wrote the first draft of the manuscript with the input by EV and the draft was further revised by SV.

Supplementary material

Chapter II

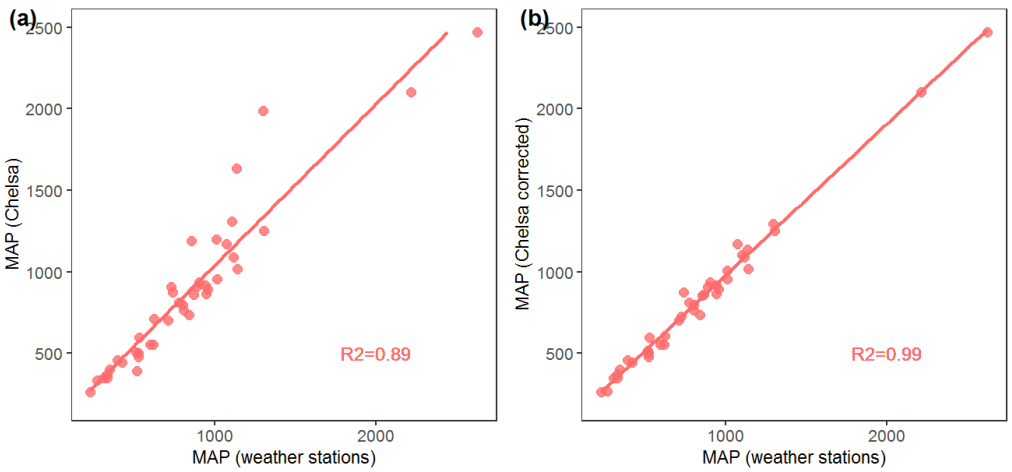


Figure S2.1 a) The correlation between MAP measured by local weather stations and MAP derived from the CHELSA database for 41 NutNet sites; b) The same correlation after correcting the values derived by CHELSA based on the values measured in the weather stations from 9 sites.

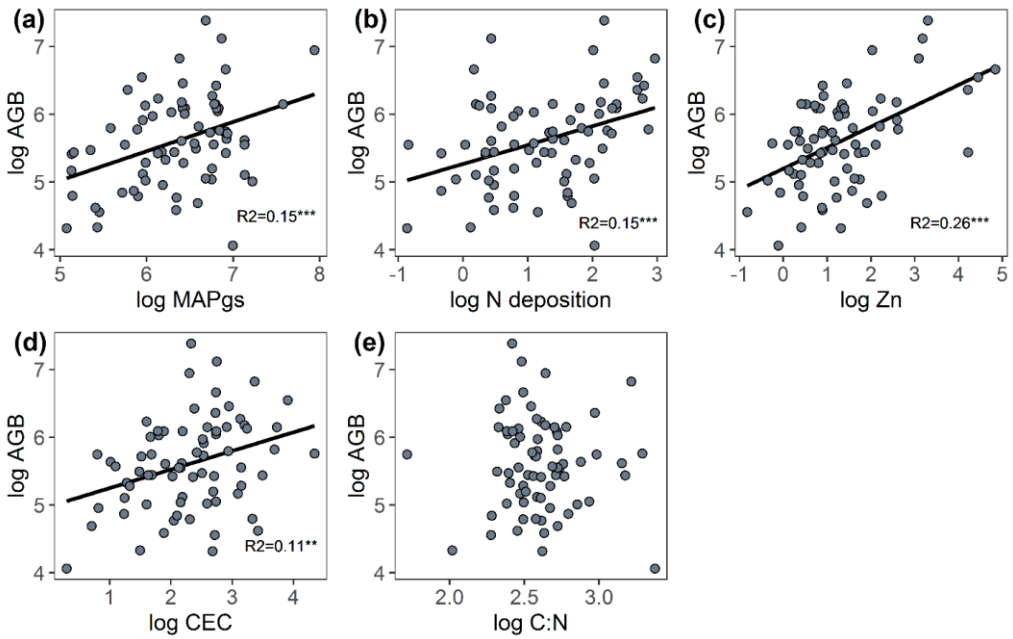


Figure S2.2 Linear regression analyses of AGB against the variables with the direct effect on AGB in the final SEM model: MAPgs, N deposition, Zn, CEC and C:N. Variables were log-transformed to improve normality and linearity.

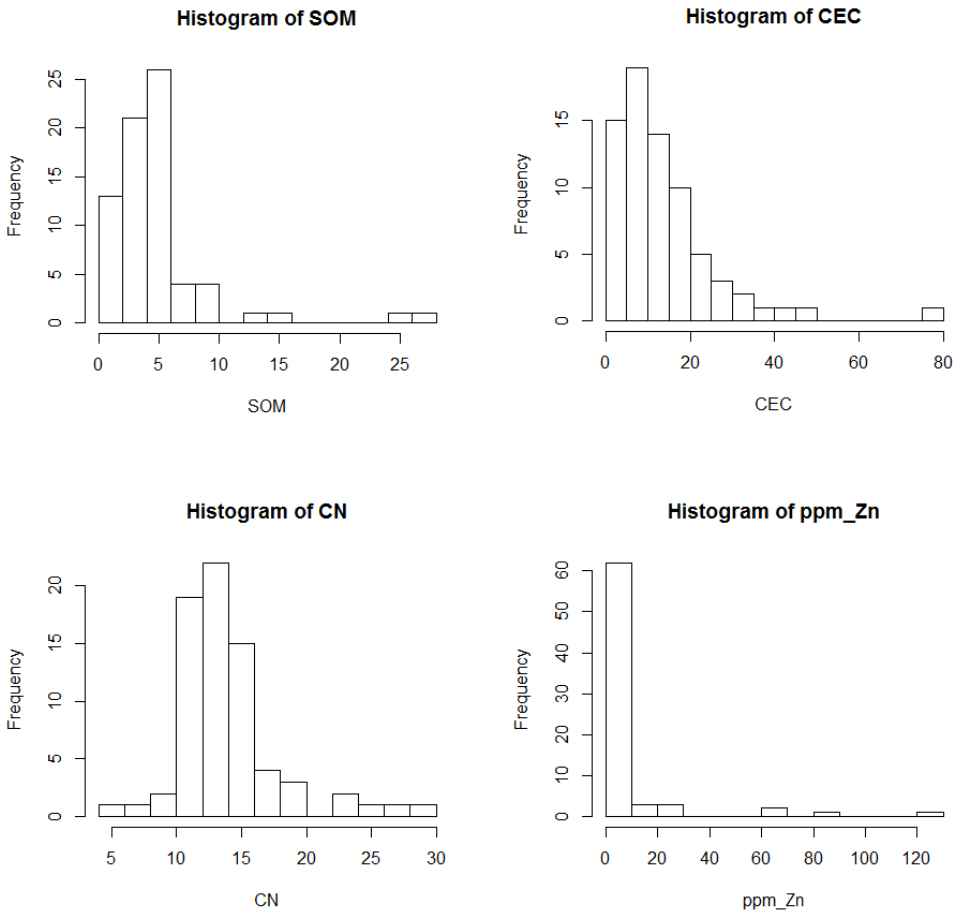


Figure S2.3 Histograms showing the distribution of values of four most important soil variables (SOM – soil organic matter [%], CEC – cation exchange capacity [meq/100g], C:N and available Zn [ppm]) across grassland sites.

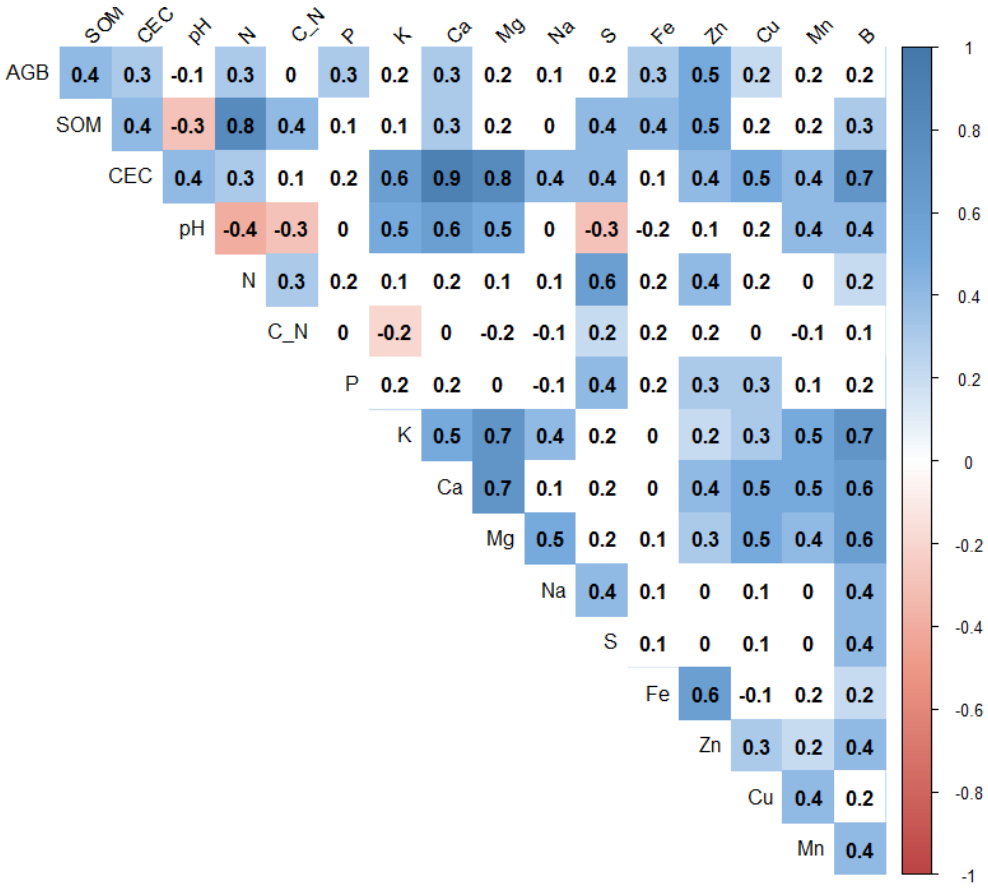


Figure S2.4 Pairwise correlation matrix between different soil factors and AGB. All variables, except pH, were log-transformed. The asterisks indicate the levels of significance (* P < 0.05, ** P < 0.01, *** P < 0.001) and the numbers indicate Pearson correlation coefficient.

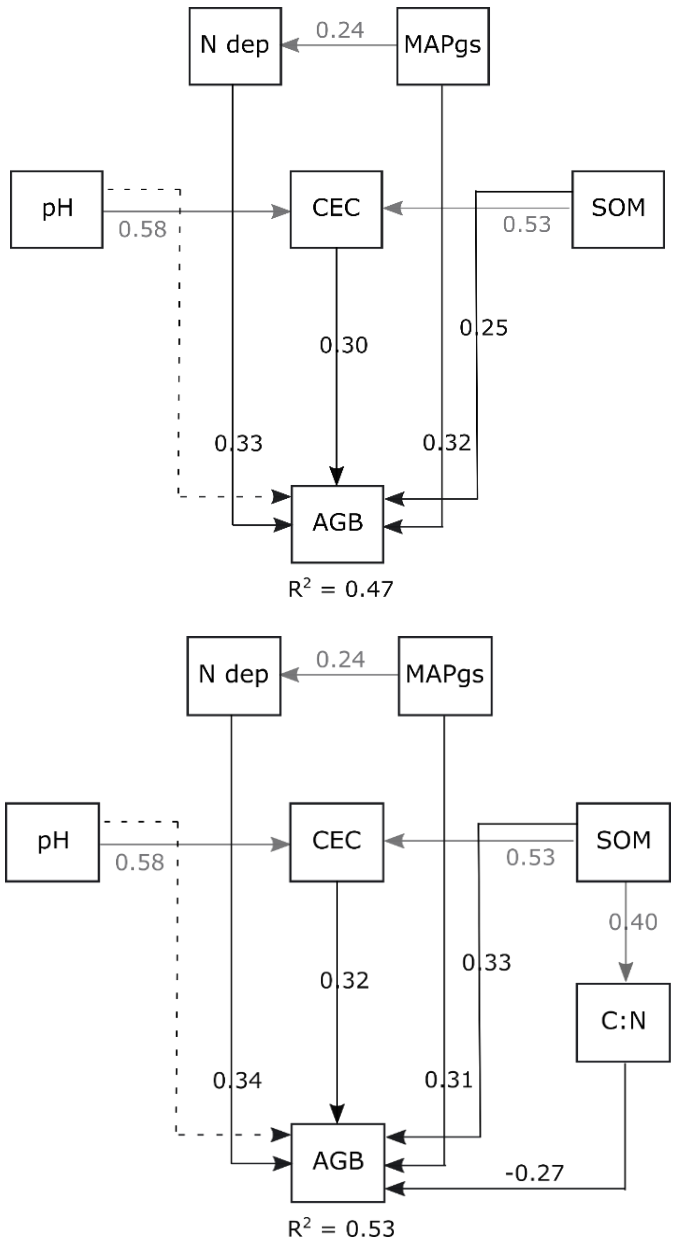


Figure S2.5 Top: 'Core' SEM depicting the direct (black lines) and indirect (grey lines) influence of different predictors of that were hypothesised to be the most important drivers of AGB. Full lines indicate significant paths and dashed non-significant paths. Factor loadings are indicated for each significant path. SRMR = 0.035, RMSEA = 0.000, CFI = 1.00, TLI = 1.024 , P (χ2) = 0.5, df = 4, R^2 (AGB) = 0.47. The model has a good fit based on each of the goodness-of-fit criteria. **Bottom:** The best model after the second step when the main macronutrients were added. SRMR = 0.062, RMSEA = 0.042, CFI = 0.99, TLI = 0.98, P (χ2) = 0.344, df = 8, R^2 (AGB) = 0.53.This model remained the best after the first step (addition of other macronutrients).

Table S2.1 72 NutNet sites included in this study: latitude, longitude, MAT (mean annual temperature), MAP (mean annual precipitation), the concentration of extractable Zn, AGB (aboveground biomass) and disturbance scores based on the presence of management practices (burning, mowing, grazing).

Site	Longitude	Latitude	Year	MAT [°C]	MAP [mm]	Zn [ppm]	AGB [g/m^2]	Disturbance score
amcamp.us	-123.01	48.47	2007	11.2	629	2.1	441.0	0
arch.us	-81.1833	27.15	2015*	22.8	1249	3.3	275.3	4
azi.cn	101.87	33.67	2007	1.8	492	13.3	369.2	2
badlau.de	11.88	51.39	2015	9.6	475	67.3	578.0	1
bari.ar	-71.1532	-41.0064	2015*	8.8	601	8.3	257.4	0
barta.us	-99.65	42.24	2007	9.2	598	1.6	205.8	0
bldr.us	-105.23	39.97	2008	9.2	390	1.3	167.0	1
bnch.us	-121.97	44.28	2007	6.5	2101	1.4	142.0	0
bogong.au	147.25	-36.87	2009	5.6	1636	3.5	415.7	0
bttr.us	-121.96	44.28	2007	6.2	2046	2.5	231.3	0
bunya.au	151.61	-26.89	2013	15.5	959	24.1	1233.2	0
burrawan.au	151.14	-27.73	2008	19.1	710	1.5	272.4	0
burren.ie	-8.99262	53.07202	2015*	9.7	1197	5.6	227.8	2
cbgb.us	-93.39	41.79	2009	9.8	895	1.8	243.2	1
cdcr.us	-93.21	45.43	2007	7.1	764	2.2	196.9	0
cdpt.us	-101.63	41.2	2007	10.4	442	1.6	120.4	0
cereep.fr	2.66	48.28	2012	11.4	614	9.2	508.7	1
chilcas.ar	-58.27	-36.28	2013	15.5	917	7.7	422.2	0
comp.pt	-8	38	2014	16.9	555	6.4	230.4	4
cowi.ca	-123.38	48.46	2007	11.1	862	2.5	528.6	0
derr.au	144.79	-37.81	2007	14.8	570	3.4	117.7	2
elliot.us	-117.052	32.875	2009*	17.9	334	4.9	328.7	0
ethass.au	138.4	-23.64	2013	24.9	228	1.5	75.9	2
fnly.us	-123.28	44.41	2007	12.3	1251	1.1	257.5	0
frue.ch	8.54	47.11	2008	7.4	1986	2.5	616.9	2
gilb.za	30.29	-29.28	2010	12.6	1016	2.0	280.5	2
glac.us	-123.03	46.87	2007	11.8	1255	1.5	164.7	0
hall.us	-86.7	36.87	2007	14.6	1261	2.3	441.1	1
hart.us	-119.5	42.72	2007	8.6	287	1.1	175.7	0
hnvr.us	-72.14	43.42	2007	7.6	992	3.7	442.2	0
hogone.us	-75.686	37.417	2017*	14.9	1071	1.4	313.3	0
hopl.us	-123.06	39.01275	2013*	12.4	1188	5.7	154.2	0
jena.de	11.53	50.93	2014	8.5	554	84.9	696.2	1
jorn.us	-106.81	32.54	2013	17.9	211	3.1	237.4	0
kbs.us	-85.39	42.41	2013	9.2	907	4.1	405.9	0
kidman.au	130.95	-16.11	2014	27.8	735	9.0	336.6	0
kilp.fi	20.83	69.05	2013	-2.7	503	4.0	224.7	1

kiny.au	143.75	-36.2	2007	15.7	398	0.7	151.5	0
koffler.ca	-79.5359	44.02405	2013*	7.2	968	4.3	636.5	0
lake.us	-95.1864	43.37913	2015*	7.7	728	13.3	482.4	1
lancaster.uk	-2.63	53.99	2008	8.6	1308	5.1	156.5	3
lead.us	-124.05	46.61	2007	11.6	1951	3.8	467.4	0
look.us	-122.13	44.21	2007	5.3	2469	2.4	98.1	0
mcla.us	-122.406	38.86427	2013*	12.9	875	5.2	258.0	0
mitch.au	143.3235	-22.4836	2013	24.9	397	2.1	459.3	0
mtca.au	117.6109	-31.7821	2008	17.9	347	4.8	130.2	0
nilla.au	146.007	-36.896	2017	14.3	699	0.8	262.3	4
pape.de	7.47	53.09	2007	9.7	766	22.0	917.8	0
pich.ec	-79	-0.08	2013	22.3	2813	7.6	1037.3	0
ping.au	116.97	-32.5	2013	16.5	459	2.3	226.8	3
pinj.au	152.92	-27.53	2013	20.4	1004	126.1	781.6	2
potrok.ar	-70.4074	-51.916	2015*	6.9	257	3.7	75.0	2
saana.fi	20.8433	69.0368	2014*	-1.8	499	67.7	230.1	0
sage.us	-120.24	39.43	2007	7.3	795	9.4	121.1	0
sava.us	-81.65	33.34	2007	17.6	1091	0.9	58.1	1
sedg.us	-120.02	34.7	2007	15.3	399	1.9	197.3	0
sereng.tz	34.51	-2.25	2008	22.5	839	3.0	308.1	1
sevi.us	-106.69	34.36	2007	14.7	262	0.4	95.4	0
sgs.us	-104.77	40.82	2007	9.2	370	0.9	127.0	0
shps.us	-112.2	44.24	2007	6.8	349	2.4	101.5	1
sier.us	-121.28	39.24	2007	16.8	860	4.3	180.9	0
smith.us	-122.62	48.21	2007	11.4	594	3.5	392.7	0
spin.us	-84.5	38.14	2007	13.1	1171	1.7	468.4	2
summ.za	30.72	-29.81	2010	18.0	1007	1.3	313.1	2
temple.us	-97.35	31.04	2007	19.9	936	2.5	317.3	0
trel.us	-88.829	40.075	2011*	11.5	956	27.0	1609.0	0
tyso.us	-90.56	38.52	2007	13.2	1104	4.0	440.7	0
ufrec.us	-81.92	27.43	2013	22.3	1368	3.2	149.5	1
ukul.za	30.4	-29.67	2009	17.2	882	1.5	469.0	1
unc.us	-79.02	36.01	2007	15.5	1114	2.4	304.2	0
valm.ch	10.37	46.63	2008	0.6	809	13.7	322.0	0
yarra.au	150.73	-33.61	2014	17.6	906	6.8	108.8	1

* For these sites, measurements were taken at control plots. For all others they were taken from pre-treatment plots.

Table S2.2 The results of the linear regression analyses between soil Zn concentrations and AGB under different C:N - N deposition levels (low-low; low-high; high-high, high-low; respectively) and different P availability levels (low and high).

C:N - N deposition level	**Slope**	**Standard error**	***t*-value**	***P*-value**	R^2
low-low	0.378	0.112	3.351	0.004	0.43
low-high	0.330	0.100	3.286	0.004	0.39
high-high	0.389	0.124	2.804	0.013	0.34
high-low	0.022	0.108	0.203	0.842	0.01
P availability level	**Slope**	**Standard error**	***t*-value**	***P*-value**	R^2
low	0.193	0.098	1.959	0.058	0.10
high	0.368	0.083	4.424	0.000	0.36

Table S2.3 The results of the linear regression between soil Zn concentrations and AGB under different C:N and N deposition levels (low- low; low-high; high-high, high-low). The threshold between the levels was chosen based on the mean values.

C:N - N dep. level	slope	st_error	t-value	P-value	R2
low-low	0.313	0.157	3.427	0.002	0.31
low-high	0.237	0.083	2.836	0.01	0.33
high-high	0.525	0.224	2.347	0.057	0.48
high-low	-0.038	0.092	-0.411	0.687	0.01
P availability level	slope	st_error	t-value	P-value	R2
low	0.201	0.097	2.078	0.044	0.09
high	0.363	0.078	4.621	0.000	0.48

Table S2.4 The relationship between Zn and AGB in the subset of NutNet sites for which the effect nutrient additions was assessed in the study by Fay et al. (2015). The relationship in the soils that were shown to be (co)limited by N was contrasted to the relationship in the soils that did not respond to N addition (alone or in the combination with P).

Response	R^2	p-value	slope	st_error	t-value
N limitation					
No (n = 9)	0.36	<0.001	0.46	0.11	3.88
Yes (n = 29)	0.02	0.73	-0.11	0.30	-0.37
NP co-limitation					
No (n = 23)	0.48	0.01	0.44	0.14	3.07
Yes (n = 15)	0.06	0.26	0.21	0.18	1.15

Chapter III

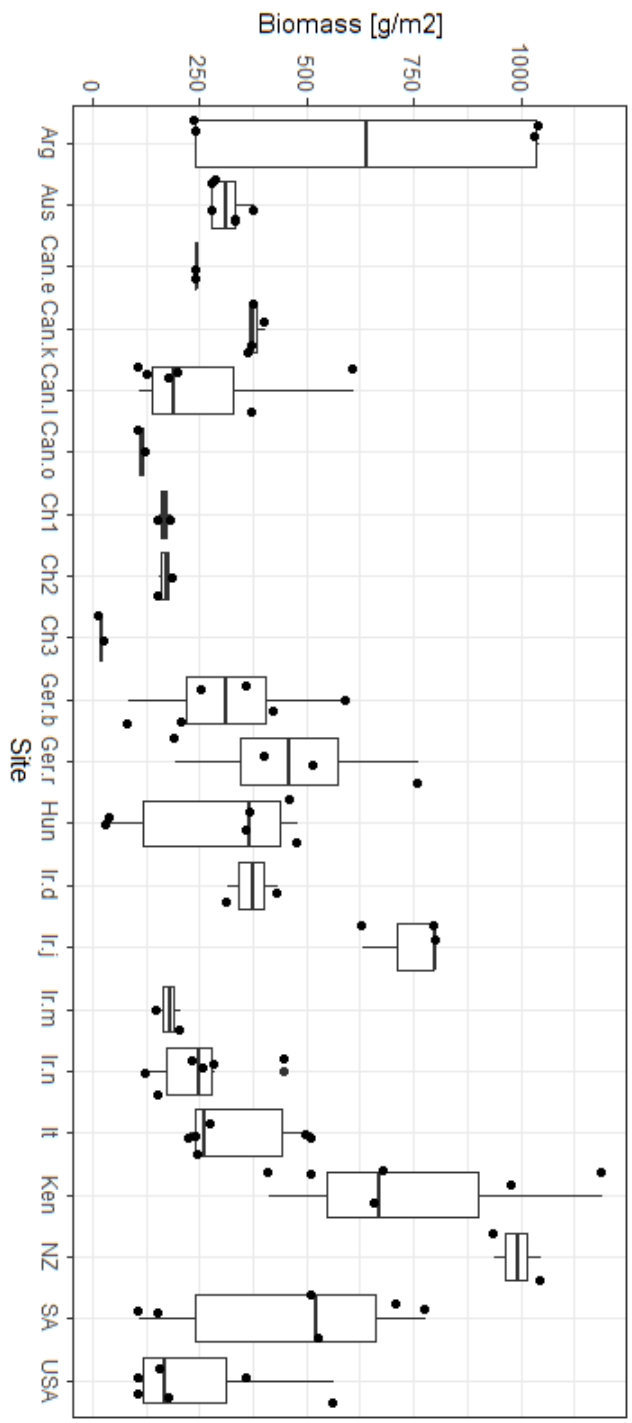


Figure S3.1 AGB variability within each of the 21 HerbDivNet sites. Points represent the 64 m^2 sampling plots within a site.

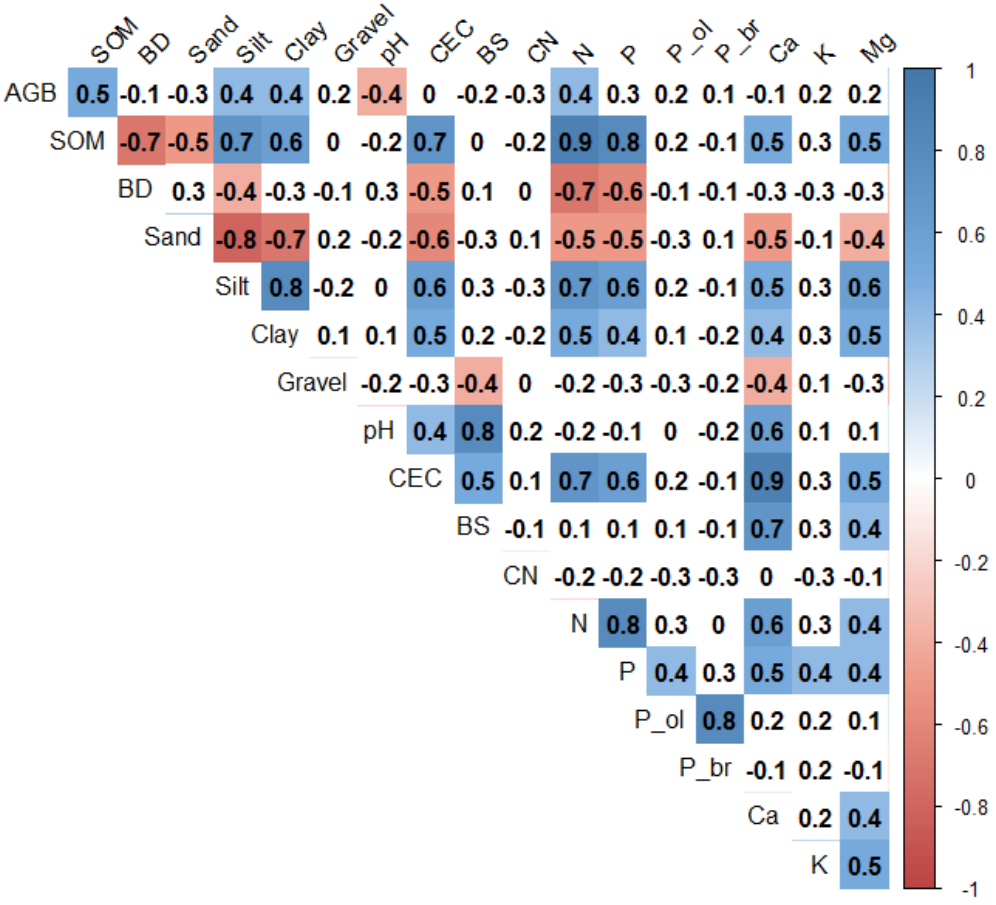


Figure S3.2 Correlations between different soil properties. Values indicate the Pearson correlation coefficient. Red indicates negative correlations, blue positive correlations and white no significant correlations ($P > 0.05$).

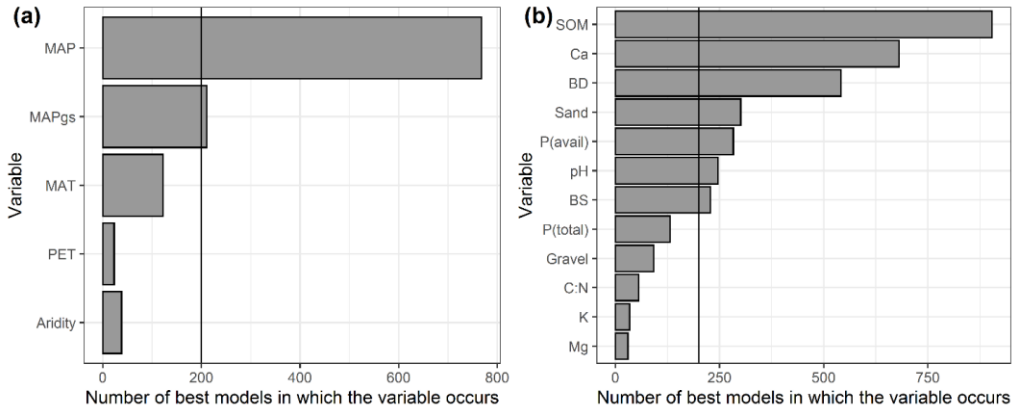


Figure S3.3 The number of the best models explaining global-scale variation in AGB (out of 1000 best models selected for each of the subsets of the dataset) in which different a) climatic variables occur, b) soil variables occur. The vertical line indicates the threshold above which the influence of variables on AGB was considered to be potentially important and these variables were included in further analyses.

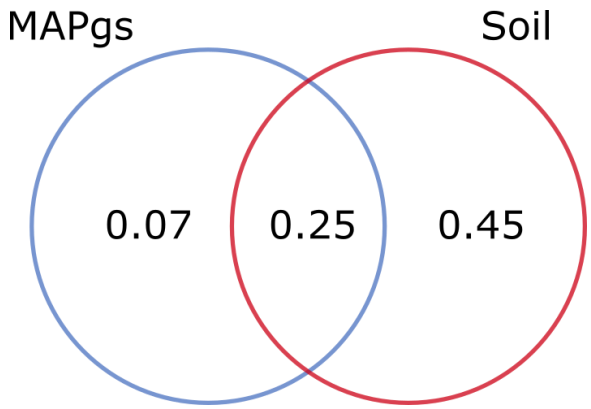


Figure S3.4 Variance partitioning between the most important climatic (MAPgs) and soil variables (SOM, Ca, % sand, BD) explaining the global-scale variation in AGB.

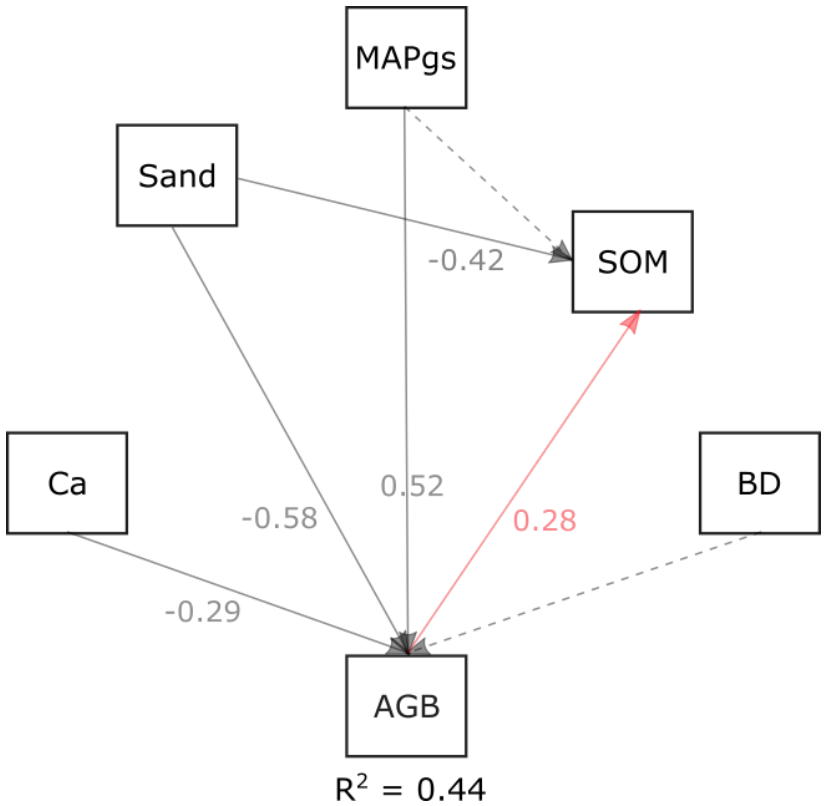


Figure S3.5 Structural equation model including the path from AGB to SOM (red arrow). Dashed lines indicate non-significant paths (P > 0.05). Indicators of model fit: SRMR = 0.15, RMSEA = 0.6, CFI = 0.55, TLI = -1.1, P (χ^2) = 0.00, df = 3, $R^2_{(AGB)}$ = 0.44, $R^2_{(SOM)}$ = 0.34. This model had a poor fit according to the goodness-of-fit criteria. AGB had a substantially lower effect on SOM than the effect of SOM on AGB (Figure 3.4).

Table S3.1 Information about 21 sites included in the analyses (MAP – mean annual precipitation, MAT – mean annual temperature). The mean values of MAT and MAP per site are reported along with the standard deviation.

site ID	*site name*	*country*	*# of plots*	*year: AGB sampling*	*year: soil sampling*	*mean MAP (sd)*	*mean MAT (sd)*
Aus	Pitztal	Austria	6	2012	2017	796 (50)	4.6 (0.9)
Arg	Córdoba	Argentina	4	2012	2017	995 (0)	12.9 (0)
Can.e	Elginfield Observatory	Canada	2	2012	2018	1001 (0)	7.9 (0)
Can.k	Kinsella	Canada	4	2017	2017	429 (16)	2.5 (0.1)
Can.l	Lac du Bois	Canada	6	2012	2018	379 (18)	6.5 (0.5)
Can.o	Onefour, AB	Canada	2	2012	2017	301 (10)	6.7 (0.1)
Ch1	Inner Mongolia	China	2	2017	2018	340 (0)	1.5 (0)
Ch2	Inner Mongolia	China	2	2017	2018	290 (0)	2.5 (0)
Ch3	Inner Mongolia	China	2	2017	2018	180 (0)	2.4 (0)
Ger.b	Bayreuth	Germany	6	2012	2017	721 (38)	8.5 (0.2)
Ger.r	Rostock	Germany	4	2015	2017	668 (0)	8.7 (0)
Hun	Soroksár, Fülöpháza, Battonya	Hungary	6	2012	2018	561 (5)	11.0 (0.4)
Ir.d	Damavand	Iran	2	2012	2017	435 (16)	7.0 (0.5)
Ir.j	Javaherdeh	Iran	3	2012	2017	1230 (13)	5.3 (1.3)
Ir.m	Masuleh	Iran	2	2012	2017	843 (54)	5.5 (0.1)
Ir.n	North Khorasan, Golestan	Iran	6	2012/2017	2017	398 (65)	8.6 (2.5)
It	Torricchio Nature Reserve	Italy	6	2012	2017	1057 (21)	8.3 (0.1)
Ken	Laikipia	Kenya	6	2012	2018	562 (67)	19.9 (0.3)
NZ	Waitati South Island	New Zealand	2	2012	2018	689 (0)	10.3 (0)
SA	Loskopdam	South Africa	6	2012	2017	699 (96)	16.6 (1)
USA	Fort Keogh	USA	6	2016	2017	333 (16)	8.9 (0.1)

Table S3.2 Aboveground biomass (AGB) [g/m^2] and the subset of important soil properties measured at each plot within different sites: pH (KCl), Ca [meq/100g], cation exchange capacity (CEC) [meq/100g], base saturation (BS) [%], available P (Olsen) [ppm], P total [ppm], C:N, soil organic matter (SOM) [%], fine-soil bulk density [kg/m^3], sand [%].

Site	*Plot*	*AGB*	*pH*	*Ca*	*CEC*	*BS*	*Pav*	*Ptot*	*C:N*	*SOM*	*BD*	*Sand*
Aus	A1	284.7	4.3	6.8	17.9	0.5	5.3	488	16.7	13.7	0.70	47
Aus	A2	278.4	4.2	8.6	22.0	0.5	4.7	654	15.9	15.7	0.60	73
Aus	A3	375.2	5.0	14.4	22.9	0.7	13.4	777	12.6	13.1	0.70	67
Aus	A4	331.1	4.6	13.0	23.2	0.7	8.0	391	13.0	15.1	0.60	57
Aus	A5	276.7	4.8	17.8	33.7	0.7	12.4	921	11.5	26.8	0.30	54
Aus	A6	333.1	4.7	22.6	44.7	0.6	25.7	2403	11.6	30.0	0.20	54
Arg	Ar1	240.8	4.9	9.0	15.1	0.8	3.2	341	11.5	6.4	0.90	73
Arg	Ar2	236.7	4.8	9.1	16.8	0.9	3.7	348	12.5	6.0	0.60	73
Arg	Ar3	1032.2	5.0	7.8	12.9	0.8	4.7	395	11.1	6.2	1.30	62
Arg	Ar4	1037.6	5.0	7.0	11.3	0.8	5.5	524	12.1	5.9	1.20	74
Can.e	C.e1	239.7	7.1	29.2	31.6	1.0	7.3	825	13.3	11.3	1.10	20
Can.e	C.e2	240.3	7.2	28.3	30.3	1.0	4.9	651	15.0	9.0	1.20	22
Can.k	Ca.k1	376.4	5.1	13.6	23.6	0.8	0.7	828	12.0	16.2	1.00	37
Can.k	Ca.k2	361.9	5.6	15.0	22.3	0.9	0.7	803	12.0	14.2	0.90	42
Can.k	Ca.k3	398.5	5.1	12.4	21.3	0.8	0.6	872	11.9	13.2	0.70	42
Can.k	Ca.k4	369.0	4.6	10.4	18.2	0.8	0.7	843	12.5	11.1	0.90	53
Can.l	Ca.l1	126.6	5.5	10.6	16.7	1.0	15.2	1107	12.2	7.2	0.80	52
Can.l	Ca.l2	607.9	7.7	23.9	53.7	1.0	35.6	1270	11.5	16.4	0.70	43
Can.l	Ca.l3	106.3	5.8	16.5	25.1	1.0	19.7	1464	11.6	10.9	0.60	50
Can.l	Ca.l4	371.1	5.9	17.8	27.6	1.0	25.9	1308	12.2	15.0	0.50	46
Can.l	Ca.l5	199.2	5.4	15.2	24.6	0.9	20.7	1418	12.1	11.0	0.60	50
Can.l	Ca.l6	175.9	5.6	20.5	28.8	0.9	26.1	1436	11.8	14.7	0.60	50
Can.o	Ca.o1	121.3	5.9	9.0	13.5	0.9	2.0	335	12.9	3.9	1.20	47
Can.o	Ca.o2	106.8	5.5	5.2	8.8	0.9	2.8	377	11.2	2.7	1.40	68
Ch1	Ch11	181.7	6.2	9.0	10.8	1.0	3.9	169	11.1	3.5	1.30	72
Ch1	Ch12	151.0	6.2	5.8	7.5	1.0	4.2	205	10.7	2.8	1.40	84
Ch2	Ch21	152.2	5.9	11.8	16.7	1.0	8.9	354	10.9	4.9	1.20	55
Ch2	Ch22	185.5	5.9	11.2	15.5	0.9	9.0	384	9.7	4.4	1.30	52
Ch3	Ch31	24.0	7.6	5.3	6.8	1.0	4.5	135	8.7	1.2	1.40	86
Ch3	Ch32	13.3	7.8	18.2	20.0	1.0	4.1	141	13.7	1.2	1.60	83
Ger.b	G.b1	421.4	5.9	18.7	26.5	0.9	15.2	787	12.7	10.9	0.60	22
Ger.b	G.b2	588.6	5.8	29.1	42.4	0.9	18.7	1042	13.1	19.4	0.50	9
Ger.b	G.b3	356.8	4.2	7.4	14.5	0.7	6.5	422	12.4	5.9	1.00	56
Ger.b	G.b4	254.4	4.5	5.1	10.1	0.7	26.3	442	11.7	3.5	1.20	73

Ger.b	G.b5	82.5	7.3	20.6	25.4	1.0	3.9	228	18.8	6.9	1.10	55
Ger.b	G.b6	207.7	4.6	3.3	6.5	0.6	12.8	480	11.7	3.6	1.00	83
Ger.r	G.r1	188.5	4.2	1.8	4.5	0.5	21.5	505	11.5	2.9	1.30	87
Ger.r	G.r3	756.1	4.5	3.0	6.5	0.6	16.2	488	13.9	4.1	1.10	84
Ger.r	G.r2	398.8	3.9	1.0	4.9	0.3	35.2	501	11.3	3.3	1.30	89
Ger.r	G.r4	513.4	4.6	3.8	7.8	0.6	14.0	574	11.6	5.2	1.20	78
Hun	H.bg1	476.3	7.7	31.9	37.0	1.0	16.2	699	19.1	14.4	1.30	30
Hun	H.bg2	459.8	7.5	32.3	37.3	1.0	14.4	626	20.2	13.4	1.50	28
Hun	H.s1	40.4	8.2	14.7	15.2	1.0	1.8	280	40.3	1.2	1.04	95
Hun	H.s2	30.2	8.4	15.1	15.9	1.0	2.0	224	36.0	1.1	1.21	94
Hun	H.s3	357.8	8.5	25.2	29.4	1.0	2.7	669	10.8	11.1	0.69	20
Hun	H.s4	366.8	6.2	18.1	22.7	1.0	4.1	739	12.1	10.7	0.76	26
Ir.d	I.d1	428.2	5.8	9.9	13.4	0.9	20.6	1035	10.9	6.8	0.85	11
Ir.d	I.d2	310.1	6.1	18.8	23.8	1.0	26.6	1261	6.7	7.2	0.92	18
Ir.j	I.j1	800.0	5.3	19.9	25.4	0.9	16.0	646	10.5	12.5	0.76	31
Ir.j	I.j2	793.7	5.7	17.3	21.5	1.0	13.0	626	10.5	9.7	0.83	10
Ir.j	I.j3	625.7	6.0	19.3	25.9	1.0	53.2	1438	10.1	12.6	0.80	36
Ir.m	I.m1	203.8	6.9	33.9	35.5	1.0	9.7	523	11.1	8.0	0.90	46
Ir.m	I.m2	149.3	6.8	30.1	31.9	1.0	13.5	622	10.1	8.5	0.86	32
Ir.n	I.n1	122.6	8.0	28.0	30.5	1.0	6.9	489	20.9	5.8	0.82	14
Ir.n	I.n2	152.8	7.8	22.3	23.8	1.0	5.5	375	10.6	4.1	0.73	23
Ir.n	I.n3	283.8	6.9	31.7	34.1	1.0	19.7	779	11.8	11.1	0.77	16
Ir.n	I.n4	257.4	7.4	32.7	36.1	1.0	12.4	505	11.4	8.8	0.82	11
Ir.n	I.r1	444.5	5.7	17.2	21.6	1.0	21.4	735	10.1	11.5	0.82	18
Ir.n	I.r2	230.8	6.0	19.1	24.0	1.0	18.9	710	9.8	11.4	0.78	19
It	It1	498.4	6.5	30.9	34.1	1.0	9.2	1080	10.8	15.0	0.74	32
It	It2	510.7	6.9	39.2	42.0	1.0	5.3	977	10.2	17.5	0.71	52
It	It3	274.2	6.8	50.1	53.3	1.0	8.4	1246	13.0	24.2	0.34	46
It	It4	243.7	6.8	42.9	45.9	1.0	7.0	1009	13.6	20.8	0.58	27
It	It5	224.9	6.3	31.5	36.6	1.0	5.5	909	12.5	20.0	0.54	40
It	It6	238.7	6.6	36.9	40.9	1.0	5.4	889	12.2	17.5	0.63	38
Ken	K1	1187.3	7.2	6.0	10.5	0.9	25.5	499	11.7	3.2	1.72	65
Ken	K2	974.8	6.7	4.1	8.7	0.9	8.0	258	10.8	3.9	1.99	64
Ken	K3	656.8	6.9	6.3	12.2	0.9	5.2	277	11.1	4.3	2.15	53
Ken	K4	676.4	6.7	2.4	6.0	0.8	6.4	219	10.9	3.8	1.76	74
Ken	K5	409.7	6.9	4.7	8.3	0.9	6.3	235	11.2	5.2	1.43	69
Ken	K6	509.2	6.8	5.5	9.5	0.9	9.0	267	10.2	3.6	1.73	67
NZ	NZ1	935.4	4.2	7.8	13.1	0.8	26.1	920	10.4	10.5	1.01	28
NZ	NZ2	1041.5	4.3	8.4	14.1	0.9	32.2	1165	10.5	7.0	0.69	20

SA	SA1	511.0	4.9	2.6	8.3	0.7	1.9	274	14.1	5.8	0.62	58
SA	SA2	524.5	4.4	2.0	7.5	0.6	1.8	273	16.0	5.9	0.56	58
SA	SA3	706.1	4.2	1.6	8.2	0.4	1.3	340	17.3	8.7	0.75	38
SA	SA4	776.2	4.1	1.1	6.3	0.4	0.8	372	16.4	6.8	0.93	41
SA	SA5	105.0	4.3	1.9	9.2	0.5	2.2	349	13.3	7.1	0.74	64
SA	SA6	150.3	4.4	2.2	9.6	0.5	3.4	383	14.7	8.6	0.47	55
USA	U1	560.7	5.8	28.7	52.0	1.0	20.2	706	15.6	11.8	0.50	8
USA	U2	356.3	7.0	19.1	28.2	1.0	3.5	731	16.2	4.5	0.87	33
USA	U3	104.3	7.3	24.4	28.1	1.0	8.9	691	16.1	4.6	0.89	58
USA	U4	103.9	7.3	23.4	27.7	1.0	4.8	598	15.4	4.2	1.00	54
USA	U5	177.4	5.4	6.7	12.5	0.9	8.6	554	13.8	4.7	0.90	49
USA	U6	154.5	5.7	6.3	10.4	0.9	5.6	500	15.0	3.4	0.96	60

Table S3.3 Parameters of the best multiple regression models (with 'site' as a random effect) using: 1) variables measured at soil depth 1 (0-10 cm); and, for comparison, the same model where high Ca plots are excluded; 2) variables measured at depth 1 where plots with very shallow soils (< 20 cm) were removed; 3) variables measured at depth 2 (10 – 20 cm).

Best model (depth 1)	Coeff.	P-value		High Ca soils removed	Coeff.	P-value	
SOM	0.72	< 0.001	n = 85	SOM	0.84	< 0.05	n = 35
Ca	-0.59	< 0.001	Sites = 21	Ca	-0.31	0.24	Sites = 12
BD	0.38	< 0.01	$^{*}R^2$ (m) = 0.51	BD	0.66	< 0.01	R^2 (m) = 0.40
Sand	-0.31	< 0.01	$^{**}R^2$ (c) = 0.74	Sand	-0.22	0.42	R^2 (c) = 0.62
P (total)	0.27	< 0.05		P (total)	0.38	0.14	
Best model (depth 1) - no shallow soils							
SOM	0.81	< 0.001	n = 68				
Ca	-0.56	< 0.001	Sites = 21				
Sand	-0.23	< 0.05	R^2 (m) = 0.51				
			R^2 (c) = 0.87				
Best model (depth 2)							
SOM	0.67	< 0.001	n = 84 ***				
Ca	-0.65	< 0.001	Sites = 21				
Sand	-0.34	< 0.01	R^2 (m) = 0.45				
			R^2 (c) = 0.79				

* R^2 (m) = marginal R^2
** R^2 (c) = conditional R^2
*** One plot contained a missing value, therefore n = 84

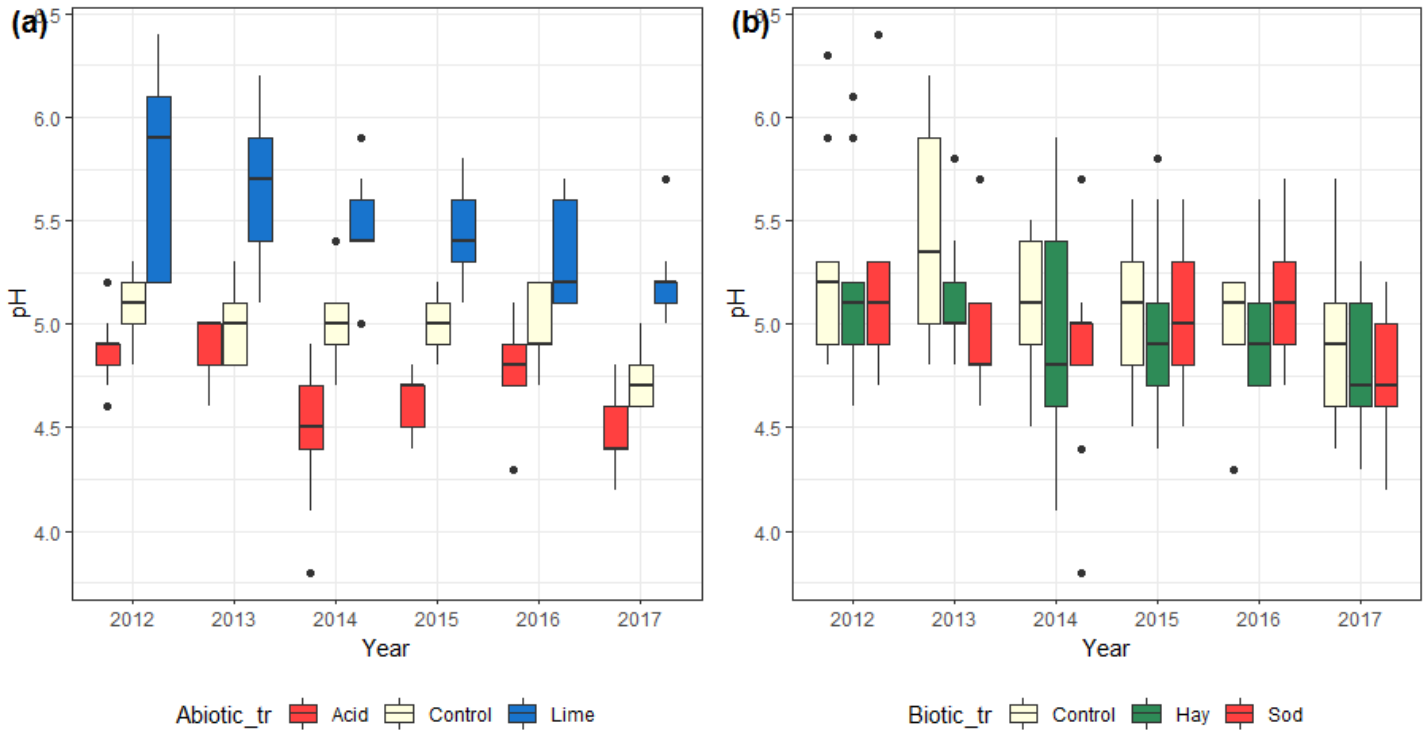


Figure S5.1 Soil pH (measured in 0.2 M NaCl) contrasted for abiotic (control, liming, acidification) and biotic (control, hay, sod additions) treatments from 2012 to 2017.

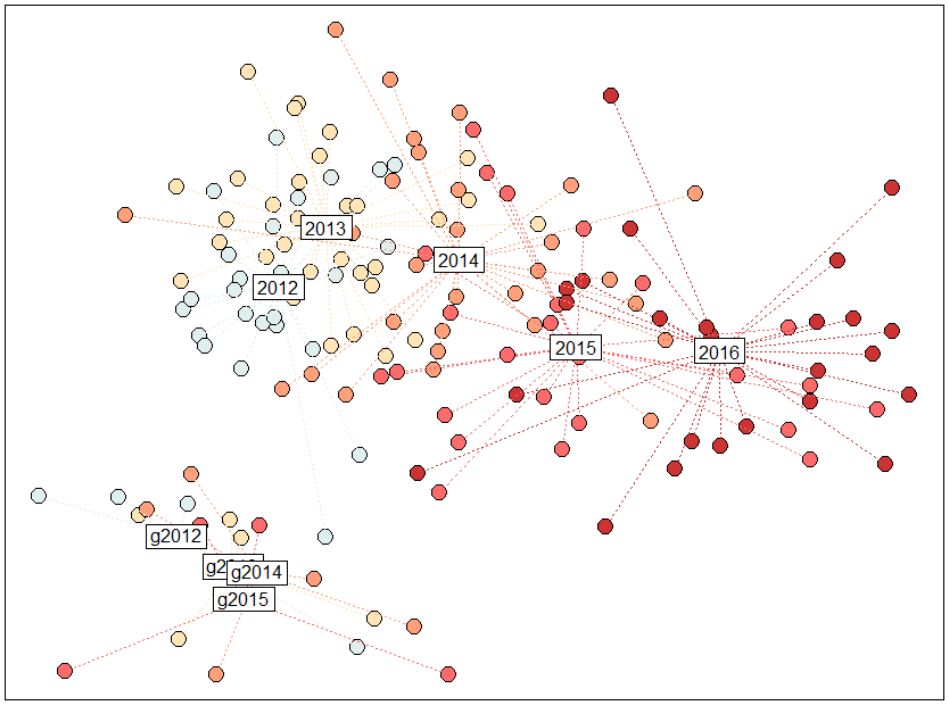


Figure S5.2 Fungal community composition sampled at the experimental site each year over five years (2012 – 2016) and stored before the DNA isolation in 2017 compared to the fungal community composition from five well-developed grasslands sampled each year over four years (g2012 – g2015) from which DNA was isolated simultaneously with heathland soils in 2017. Samples from the two sites were handled and stored in the same manner. There is a clear directional change in fungal community composition at the experimental site, while the is no difference between grassland communities from different years. This indicates that the storage effect was not a driver of the observed changes in fungal community composition at the experimental site.

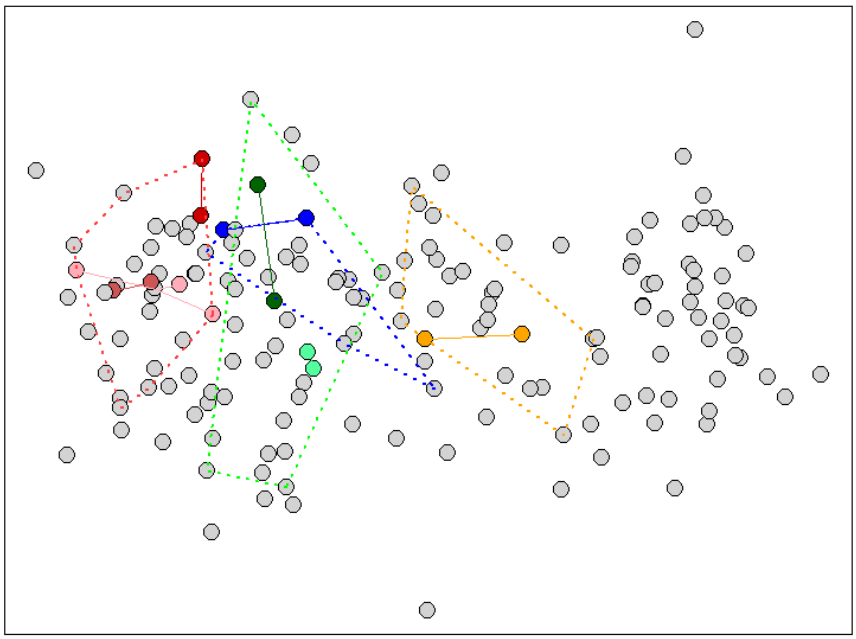


Figure S5.3 NMDS ordination (first and second axis) showing the distance between technical replicates (shown in different colours and connected with a line of the same colour). Grey dots represent other samples in the study. Dotted lines connect the samples within the group (combinations of biotic treatments and year) to which the replicates belong and their colour corresponds to the colour of the respective replicates. Relatively closely clustered replicate samples compared to the variation within the groups show the reproducibility of sample preparation and the sequencing procedure.

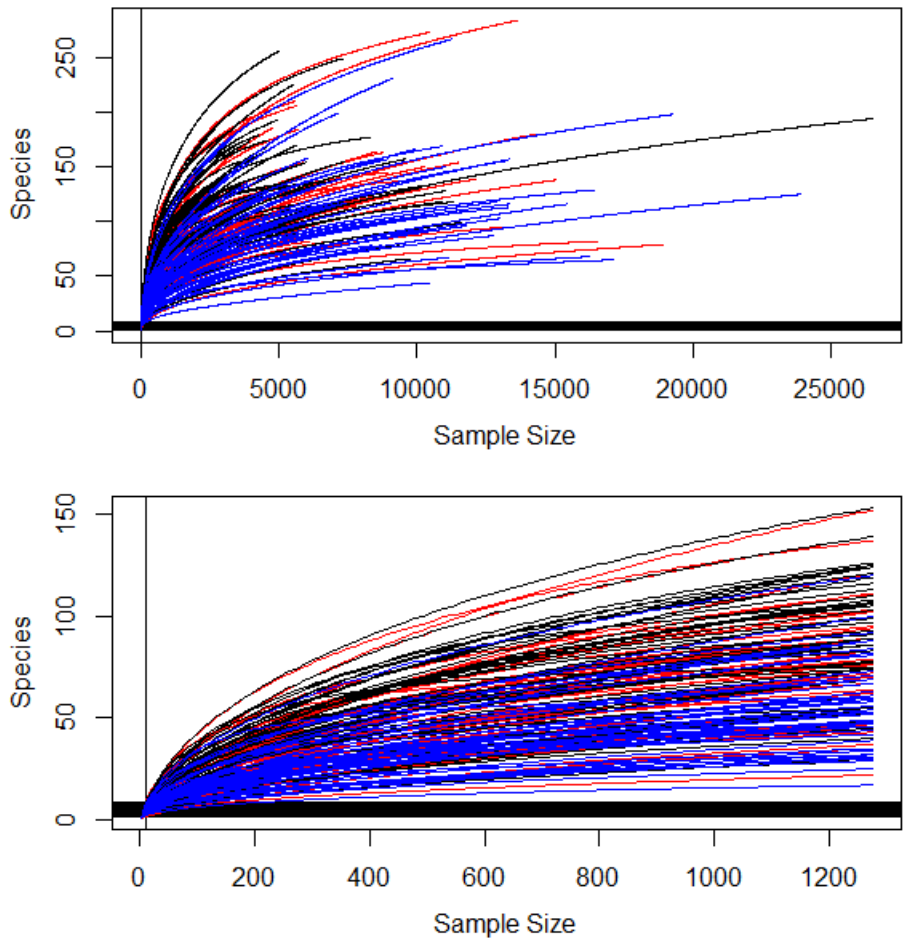


Figure S5.4 Rarefaction curves for all samples in the study; top – before rarefaction, bottom – after rarefaction to 1,275 reads. Different colours represent samples from different biotic treatments (black =control, red = hay additions, blue = sod additions).

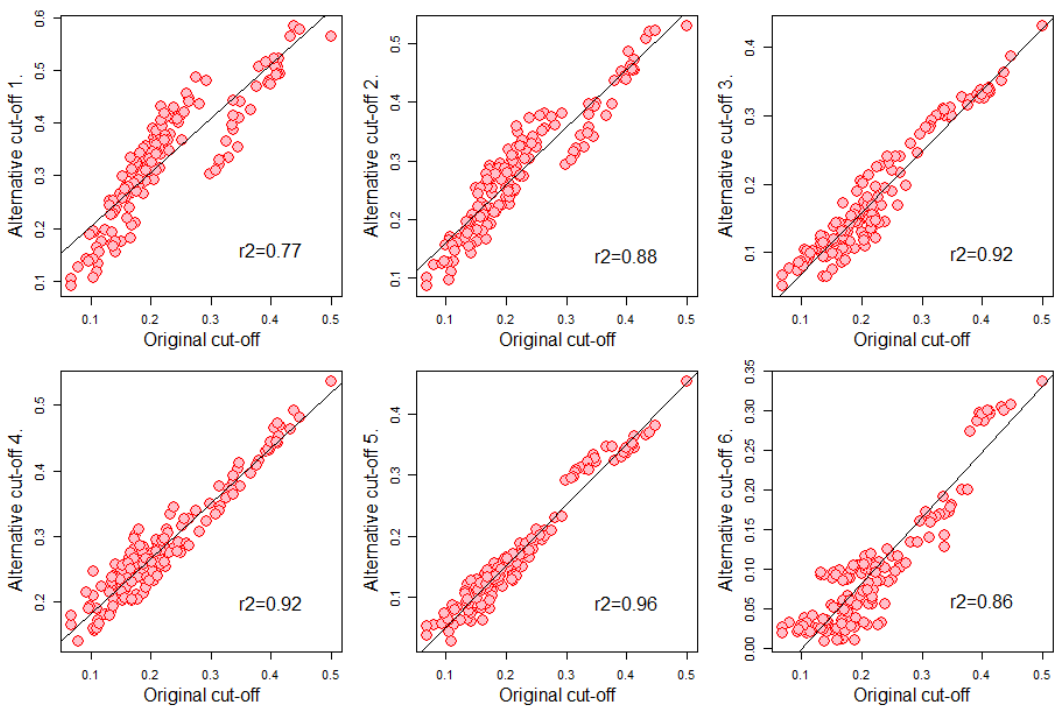


Figure S5.5 Correlations between pair-wise Euclidean distances of the full sample x sample matrix with the currently used cut-off (OTUs with more than 500 reads, 25 most abundant plant species, only the combination of taxa with Pearson r greater than 0.2) and the sample x sample distance matrix with alternative cut-offs. For each biotic treatment-year combination, we calculated the strength of the connection between taxa using the original cut-off and the alternative cut-offs. Then we calculated Euclidean distances between samples (based on the number and strength of connections of their taxa) to obtain sample x sample distance matrices, both for original and alternative cut-offs. The distance matrix with the original cut-offs was correlated with the distance matrices of different alternative cut-offs. Alternative cut-off 1: OTUs with more than 100 reads and all plant species in the study (36); cut-off 2: OTUs with more than 200 reads and all plant species; cut-off 3: OTUs with more than 1000 reads and 20 most abundant plant species; cut-off 4: no correlation criterion included; cut-off 5: Pearson r greater than 0.3; cut-off 6: Pearson r greater than 0.4 (if a higher Pearson r threshold was imposed, very few pairs of taxa would pass the filter).

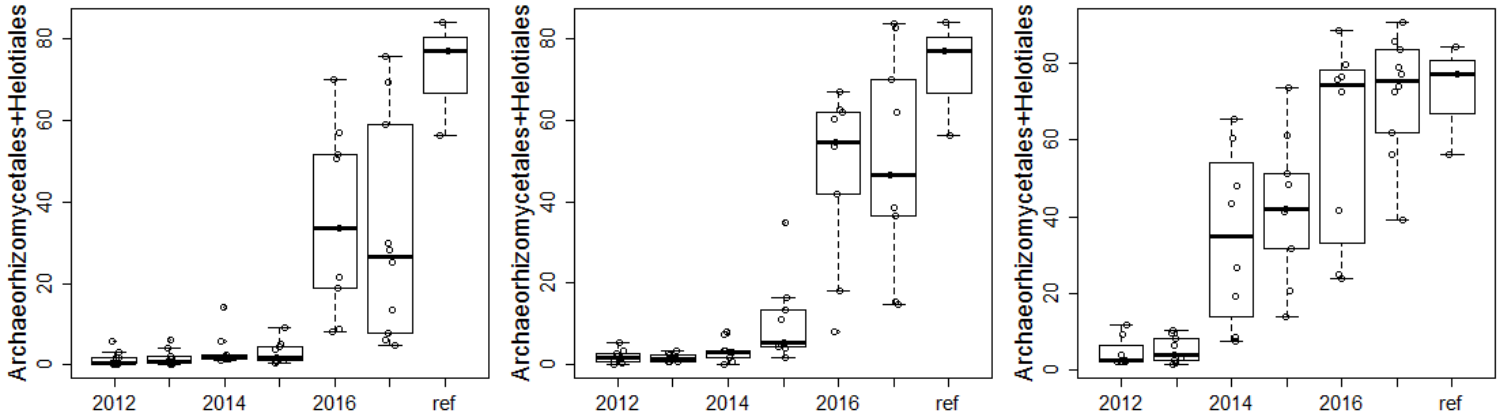


Figure S5.6 Change in the abundance of Archaeorhizomycetales and Helotiales with time for: control (left), hay (middle) and sod treatment (right) compared to the reference.

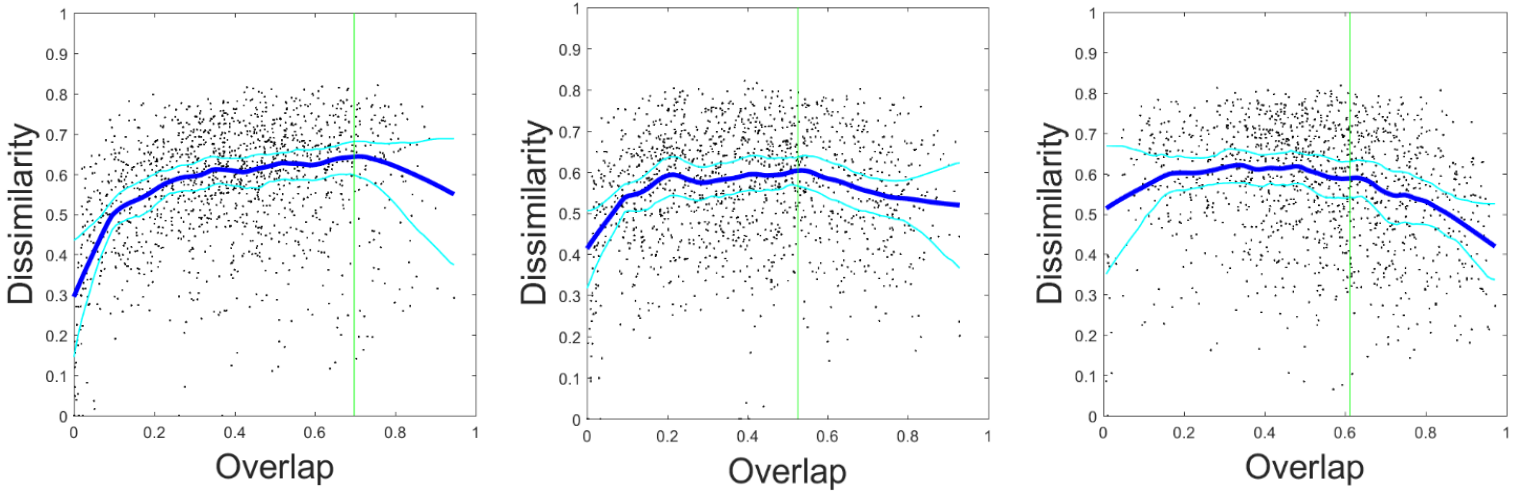


Figure S5.7 DOC analyses of biotic control (left), hay addition (middle) and sod addition treatment (right).The significantly negative relationship between fungal community overlap and dissimilarity observed in the hay and soil additions treatments indicates that, in these treatments, biotic interactions play a significant role in structuring the fungal community composition. Green lines represent the overlap value above which a negative relationship is considered (data points with overlap larger than the median value). Steeper slope after the green line indicates stronger biotic interactions. Dark blue lines represent the smoothed averages and light blue lines 95% confidence intervals.

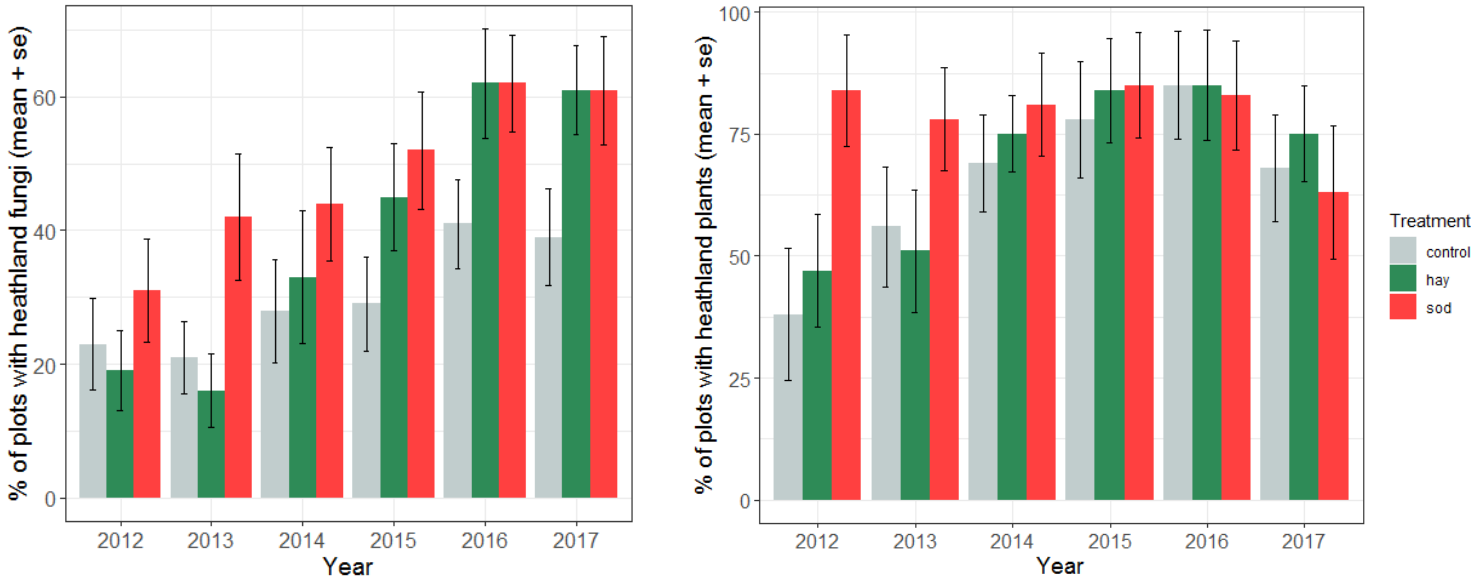


Figure S5.8 Percentage of plots that contained heathland fungal (left) and heathland plant (right) taxa for the three different biotic treatments over time, averaged over each of these heathland taxa (± SE). It can be seen that initially most fungal taxa were detectable at a similar frequency between treatments but stayed behind in the control for subsequent years. For heathland plants, the mean frequency differed initially but was similar and high across all treatments in later years.

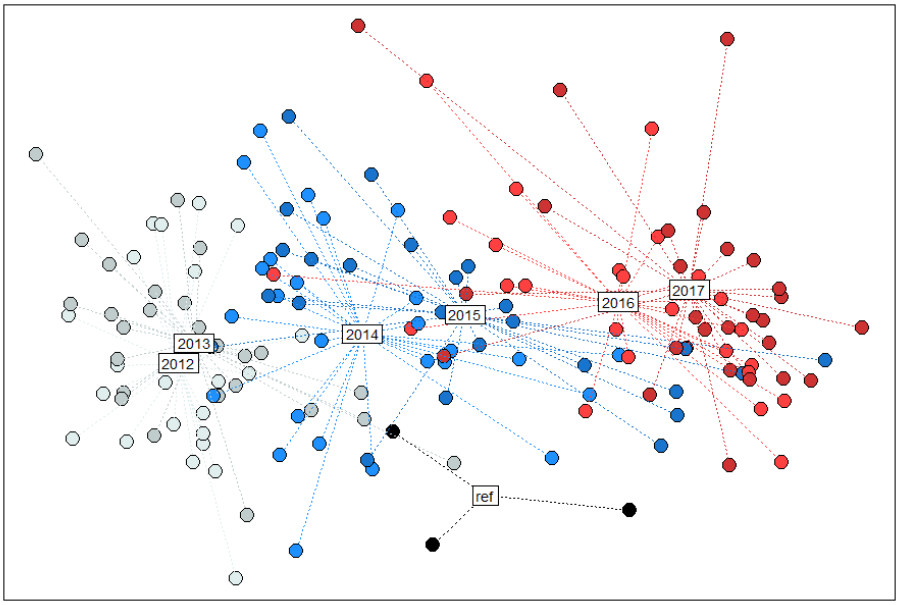


Figure S5.9 Change in fungal community composition over the course of six years (from 2012 to 2017) compared to the reference heathland communities. First and third dimensions are shown. First and second dimensions are presented in the main text (FigureS1)

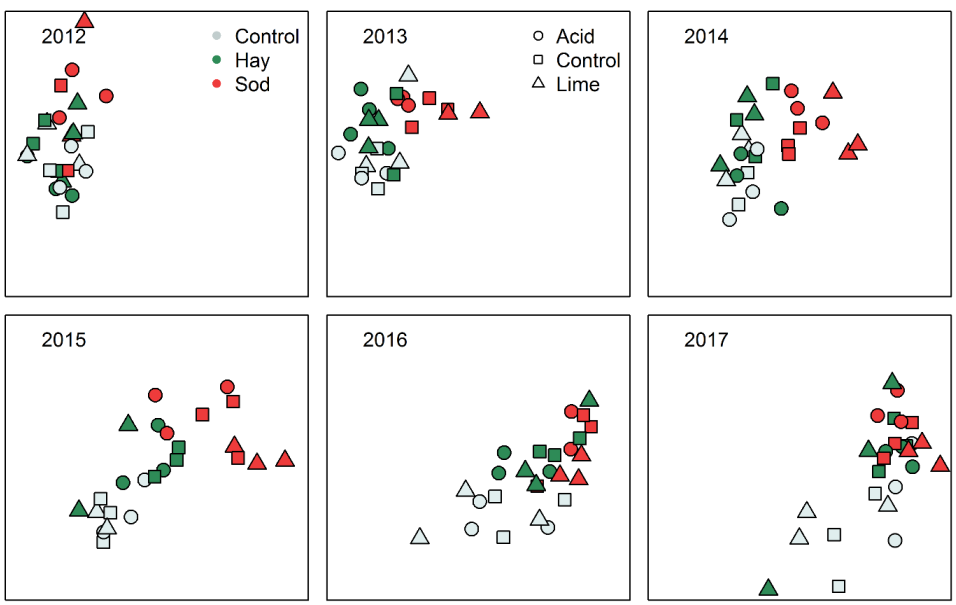


Figure S5.10 Change in fungal community composition over the course of six years (from 2012 to 2017) where different biotic and abiotic treatments are differentiated. First and third dimensions are shown. First and second dimensions are presented in the main text (Figure2a)

Table S5.1 The mean proportion (with standard deviation) of the total diversity in a sample covered by the rarefaction threshold (1,275 reads) for different biotic treatments (control, hay, sod) according to the Chao index.

Control	Hay	Sod
0.975 ± 0.01	0.975 ± 0.009	0.979 ± 0.01

Table S5.2 The results of PERMANOVA analyses using different types of transformations of OTU data.

	Untransformed	**log**	**Hellinger**	**Wisconsin+sqrt**
Model: OTU_full ~ year, strata= plot				
	r^2 = 0.11	r^2 = 0.12	r^2 = 0.12	r^2 = 0.06
year	P = 0.001	P = 0.001	P = 0.001	P = 0.001
*Model: OTU_full ~ abiotic_treat * biotic_treat, strata = year*				
	r^2 = 0.04	r^2 = 0.05	r^2 = 0.05	r^2 = 0.03
abiotic_treat	P = 0.001	P = 0.001	P = 0.001	P = 0.001
	r^2 = 0.05	r^2 = 0.06	r^2 = 0.06	r^2 = 0.04
biotic_treat	P = 0.001	P = 0.001	P = 0.001	P = 0.001
	r^2 = 0.04	r^2 = 0.04	r^2 = 0.04	r^2 = 0.04
interaction	P = 0.001	P = 0.001	P = 0.001	P = 0.001
*Model: OTU_2017 ~ abiotic_treat * biotic_treat*				
	r^2 = 0.11	r^2 = 0.13	r^2 = 0.12	r^2 = 0.11
abiotic_treat	P = 0.02	P = 0.001	P = 0.001	P = 0.001
	r^2 = 0.14	r^2 = 0.15	r^2 = 0.15	r^2 = 0.12
biotic_treat	P = 0.001	P = 0.001	P = 0.001	P = 0.001
	r^2 = 0.14	r^2 = 0.17	r^2 = 0.16	r^2 = 0.17
interaction	P = 0.324	P = 0.02	P = 0.035	P = 0.006

Table S5.3 Proportion of plant cover / fungal reads included in the network analysis out of the total cover / number of fungal reads, averaged for different biotic treatments each year. As can be seen, the taxa included in the network analysis cover the majority of fungal reads and almost all of the total plant cover while rare fungal OTUs and plant species were excluded.

	2013		**2014**		**2015**		**2016**		**2017**	
Treat.	Plants	Fungi	Plants	Fungi	Plants	Fungi	Plants	Fungi	Plants	Fungi
Control	0.95	0.60	0.98	0.74	0.96	0.71	0.96	0.81	0.97	0.76
Hay	0.94	0.72	0.96	0.70	0.96	0.82	0.84	0.87	0.95	0.84
Sod	0.97	0.67	0.99	088	0.98	0.91	0.80	0.90	0.99	0.92

Table S5.4 The list of heathland-related and non-heathland plant species included in the network analysis.

Heathland plant species	Non-heathland plant species
1. *Calluna vulgaris* L.	1. *Gnaphalium uliginosum L.*
2. *Erica tetralix* L.	2. *Cerastium fontanum* Baumg.
3. *Betula pendula* Roth	3. *Trifolium repens* L.
4. *Molinia caerulea* L.	4. *Ranunculus repens* L.
5. *Carex pilulifera* L.	5. *Plantago major* L.
6. *Juncus effusus* L.	6. *Salix cinerea* L.
7. *Juncus bufonius* L.	7. *Plantago lanceolate* L.
8. *Juncus squarrosus* L.	8. *Persicaria maculosa* Gray
9. *Rumex acetosella* L.	9. *Leontodon autumnalis* L.
	10. *Hypochaeris radiacata* L.
	11. *Hypericum perforatum* L.
	12. *Holcus lanatus* L.
	13. *Veronica serpyllifolia* L.
	14. *Sagina procumbrens* L.
	15. *Agrostis stolonifera* L.
	16. *Agrostis capillaris* L.

Table S5.5 Mean OTU richness for each combination of biotic and abiotic treatments throughout the years (standard deviation is indicated).

Biotic	**control**			**hay**			**sod**		
Abiotic	**acid**	**control**	**lime**	**acid**	**control**	**lime**	**acid**	**control**	**lime**
2012	61.7 ± 11	76 ± 23	78 ± 27	74.7 ± 12	76.7 ± 1	108.7 ± 30	90 ± 30	132 ± 8	97.5 ± 70
2013	90.3 ± 12	86.7 ± 13	93 ± 10	88.3 ± 12	76 ± 23	97.7 ± 18	100.8 ± 24	90.7 ± 19	84.5 ± 5
2014	67.7 ± 12	79 ± 11	98.7 ± 38	71.7 ± 45	90 ± 6	87 ± 18	63.3 ± 18	72.7 ± 8	50.7 ± 11
2015	82.3 ± 6	137 ± 26	111 ± 10	93.7 ± 42	80.3 ± 24	107 ± 28	57.3 ± 23	37.7 ± 11	36.3 ± 11
2016	58.7 ± 18	62.7 ± 11	97 ± 34	65.7 ± 16	57 ± 6	47.7 ± 30	44.5 ± 1	31.3 ± 16	56.7 ± 6
2017	29.7 ± 15	66.7 ± 11	54.7 ± 23	61.3 ± 7	58.7 ± 25	61.7 ± 28	48.7 ± 21	43 ± 9	34.3 ± 6

Table S5.6 The percentage of plant and fungal taxa (out of the total number included in the network analyses) that were found in at least one plot exposed to one of the three biotic treatments (control, hay addition, sod addition) each year. Most fungal and plant taxa included in the network analysis were present in each of the treatments in 2013, suggesting that there was no absolute dispersal limitation. See Figure S5.8 for the mean frequency of heathland-associated plant and fungal taxa in each biotic treatment.

	2013		**2014**		**2015**		**2016**		**2017**	
Treat.	Plants	Fungi	Plants	Fungi	Plants	Fungi	Plants	Fungi	Plants	Fungi
Control	92	80	96	80	84	89	96	84	77	84
Hay	92	71	92	75	88	87	84	80	88	69
Sod	84	75	88	83	84	80	80	71	68	63

Table S5.7 Soil pH (in NaCl) and percentage of soil organic matter (SOM) measured at each plot from 2012 to 2016.

Abiotic treatment	Biotic treatment	Replicate #	Year	pH (NaCl)	SOM [%]
Acidification	Control	1	2012	4.9	1.1
Acidification	Control	2	2012	4.9	1.3
Acidification	Control	3	2012	5.0	2.6
Acidification	Hay	1	2012	4.8	1.7
Acidification	Hay	2	2012	4.6	1.9
Acidification	Hay	3	2012	5.2	2
Acidification	Sod	1	2012	4.7	2
Acidification	Sod	2	2012	4.9	5.1
Acidification	Sod	3	2012	4.8	3.9
Control	Control	1	2012	4.8	1.8
Control	Control	2	2012	5.3	2.3
Control	Control	3	2012	5.2	1.5
Control	Hay	1	2012	5.0	2.2
Control	Hay	2	2012	5.1	2.1
Control	Hay	3	2012	4.9	2.6
Control	Sod	1	2012	5.1	2.4
Control	Sod	2	2012	5.3	1.1
Control	Sod	3	2012	5.1	2.1
Liming	Control	1	2012	5.2	1.8
Liming	Control	2	2012	5.9	NA
Liming	Control	3	2012	6.3	2
Liming	Hay	1	2012	5.9	1.3
Liming	Hay	2	2012	5.2	1.7
Liming	Hay	3	2012	6.1	1.9
Liming	Sod	1	2012	5.2	2.8
Liming	Sod	2	2012	6.4	2.2
Liming	Sod	3	2012	5.3	NA
Acidification	Control	1	2013	4.8	1.7
Acidification	Control	2	2013	5.0	2.4
Acidification	Control	3	2013	5.0	2.4
Acidification	Hay	1	2013	5.0	2.3
Acidification	Hay	2	2013	5.0	2
Acidification	Hay	3	2013	5.0	2.5
Acidification	Sod	1	2013	4.6	2.1
Acidification	Sod	2	2013	4.8	1.5
Acidification	Sod	3	2013	4.7	1.8
Control	Control	1	2013	5.0	2.1

Control	Control	2	2013	5.1	1.3
Control	Control	3	2013	5.3	2.2
Control	Hay	1	2013	4.9	1.8
Control	Hay	2	2013	5.1	1.7
Control	Hay	3	2013	4.8	3.7
Control	Sod	1	2013	4.8	2.4
Control	Sod	2	2013	5.1	1.3
Control	Sod	3	2013	4.8	2
Liming	Control	1	2013	5.9	2.1
Liming	Control	2	2013	5.6	1.4
Liming	Control	3	2013	5.4	2.8
Liming	Hay	1	2013	5.8	2.5
Liming	Hay	2	2013	5.2	2.2
Liming	Hay	3	2013	5.4	2.4
Liming	Sod	1	2013	5.7	5.2
Liming	Sod	2	2013	5.7	1.4
Liming	Sod	3	2013	5.1	2.8
Acidification	Control	1	2014	4.5	1.7
Acidification	Control	2	2014	4.9	1.5
Acidification	Control	3	2014	4.7	2.7
Acidification	Hay	1	2014	4.5	1.7
Acidification	Hay	2	2014	4.1	1.7
Acidification	Hay	3	2014	4.6	2.8
Acidification	Sod	1	2014	4.4	3
Acidification	Sod	2	2014	4.8	1.9
Acidification	Sod	3	2014	3.8	3.8
Control	Control	1	2014	4.9	2.9
Control	Control	2	2014	5.1	1.3
Control	Control	3	2014	5.4	2
Control	Hay	1	2014	4.8	2.2
Control	Hay	2	2014	5.1	1.1
Control	Hay	3	2014	4.7	2.9
Control	Sod	1	2014	5.1	1.4
Control	Sod	2	2014	5.0	1.7
Control	Sod	3	2014	4.9	2.2
Liming	Control	1	2014	5.4	1.6
Liming	Control	2	2014	5.4	1.5
Liming	Control	3	2014	5.5	3.3
Liming	Hay	1	2014	5.6	1.8
Liming	Hay	2	2014	5.4	1.9
Liming	Hay	3	2014	5.9	4.2

Liming	Sod	1	2014	5.0	3.6
Liming	Sod	2	2014	5.7	1.8
Liming	Sod	3	2014	5.0	4.1
Acidification	Control	1	2015	4.8	1.4
Acidification	Control	2	2015	4.7	1.3
Acidification	Control	3	2015	4.5	2.6
Acidification	Hay	1	2015	4.7	2.7
Acidification	Hay	2	2015	4.4	1.1
Acidification	Hay	3	2015	4.7	2.5
Acidification	Sod	1	2015	4.6	1.6
Acidification	Sod	2	2015	4.8	2
Acidification	Sod	3	2015	4.5	2.4
Control	Control	1	2015	4.9	1.7
Control	Control	2	2015	5.1	1.6
Control	Control	3	2015	5.1	1.8
Control	Hay	1	2015	4.9	2.1
Control	Hay	2	2015	5.1	1.3
Control	Hay	3	2015	4.8	2.5
Control	Sod	1	2015	5.0	1.5
Control	Sod	2	2015	5.2	2.1
Control	Sod	3	2015	4.9	1.7
Liming	Control	1	2015	5.3	1.6
Liming	Control	2	2015	5.4	2.5
Liming	Control	3	2015	5.6	2.1
Liming	Hay	1	2015	5.8	2.4
Liming	Hay	2	2015	5.1	2
Liming	Hay	3	2015	5.6	2.8
Liming	Sod	1	2015	5.4	2.3
Liming	Sod	2	2015	5.6	1.8
Liming	Sod	3	2015	5.3	3.2
Acidification	Control	1	2016	4.3	1.7
Acidification	Control	2	2016	4.9	0.9
Acidification	Control	3	2016	5.1	3
Acidification	Hay	1	2016	4.8	2.3
Acidification	Hay	2	2016	4.7	1.1
Acidification	Hay	3	2016	4.7	1.9
Acidification	Sod	1	2016	4.7	1.6
Acidification	Sod	2	2016	4.9	1.9
Acidification	Sod	3	2016	4.8	3
Control	Control	1	2016	4.9	2.5
Control	Control	2	2016	5.2	1.3

Control	Control	3	2016	5.2	1.5
Control	Hay	1	2016	4.7	1.7
Control	Hay	2	2016	4.9	1.5
Control	Hay	3	2016	4.9	3.2
Control	Sod	1	2016	5.1	1.7
Control	Sod	2	2016	5.2	1.3
Control	Sod	3	2016	4.9	2.2
Liming	Control	1	2016	5.1	1.7
Liming	Control	2	2016	5.2	1.8
Liming	Control	3	2016	5.1	2.7
Liming	Hay	1	2016	5.6	1.8
Liming	Hay	2	2016	5.1	1.5
Liming	Hay	3	2016	5.1	4.3
Liming	Sod	1	2016	5.7	3.2
Liming	Sod	2	2016	5.7	2.1
Liming	Sod	3	2016	5.3	2

Appendix S5.1 Testing the effect of storage conditions

In order to test for the effect of storage conditions on the fungal community composition, we additionally analysed the soil samples from 4 different *Nardus stricta* L. grasslands (located in the vicinity of the plots) that were not part of this study. These samples were taken each year in the period between 2012 and 2016 and they were handled and stored in the same manner as the samples from our experimental site. The rationale was that because these grasslands were developed and thus relatively stable (unlike the experimental plots) eventual substantial changes in their soil biotic community composition through time would be caused by the storage time/conditions (while natural interannual variations would likely produce random rather than directional variation). However, we detected no significant consistent trend in community composition between the soils collected in different years, especially not when compared to the strong effect observed in the developing heathlands (Figure S5.2). This suggests that the differences in the storage time likely did not have a substantial effect on soil fungal community composition at our experimental site, at least not in such a manner to be confounded with the effect of a succession of fungal communities in a developing system. Another line of evidence that the storage conditions did not significantly alter the soil fungal community composition is that the communities from 2017 (where the DNA was extracted shortly after the sampling) followed the pattern of change observed in the previous years (Figure 5.1) which would likely not occur if the pattern was caused by differing storage times.

Chapter VI

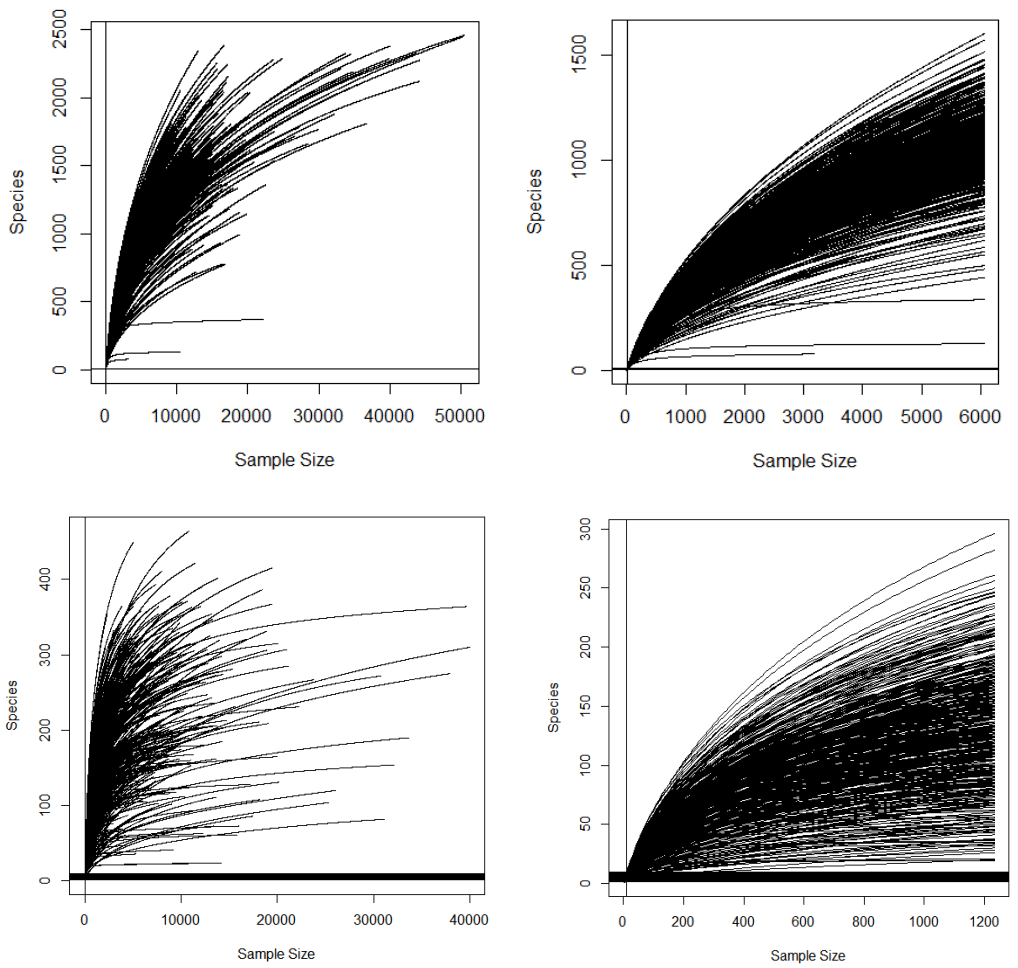


Figure S6.1 Rarefactions curves for bacteria (top) and fungi (bottom). The graphs on the left show the curves for non-rarefied data, while the graphs on the right show the curves after rarefactions at 6046 and 1231 reads for bacteria and fungi, respectively.

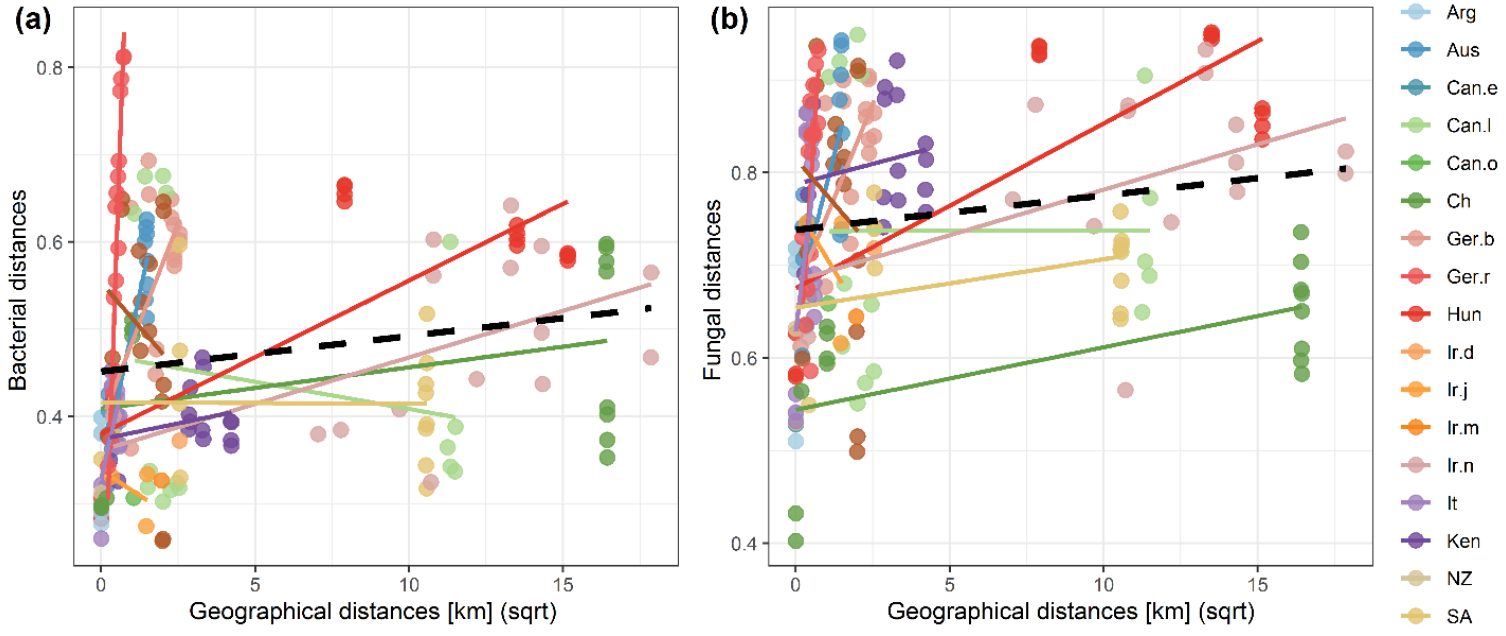


Figure S6.2 The relationship between within-site geographical distances and within-site distances in a) bacterial and b) fungal community composition. Out of the six sites that have relatively high distances between some plots (between cc 10 and 318 km), the relationship was found to be relatively strong (R^2 from 0.3 to 0.57) for only one site in the case of bacteria (Hun) and two sites for fungi (Hun and Ch). In all cases, plant community composition explained a much higher proportion of the variation in microbial communities (between 0.77 and 0.98). This demonstrates that within-site geographical distance is not the main cause for the observed relationships between environmental variables/plant community composition with microbial community composition shown in Figure 6.3.

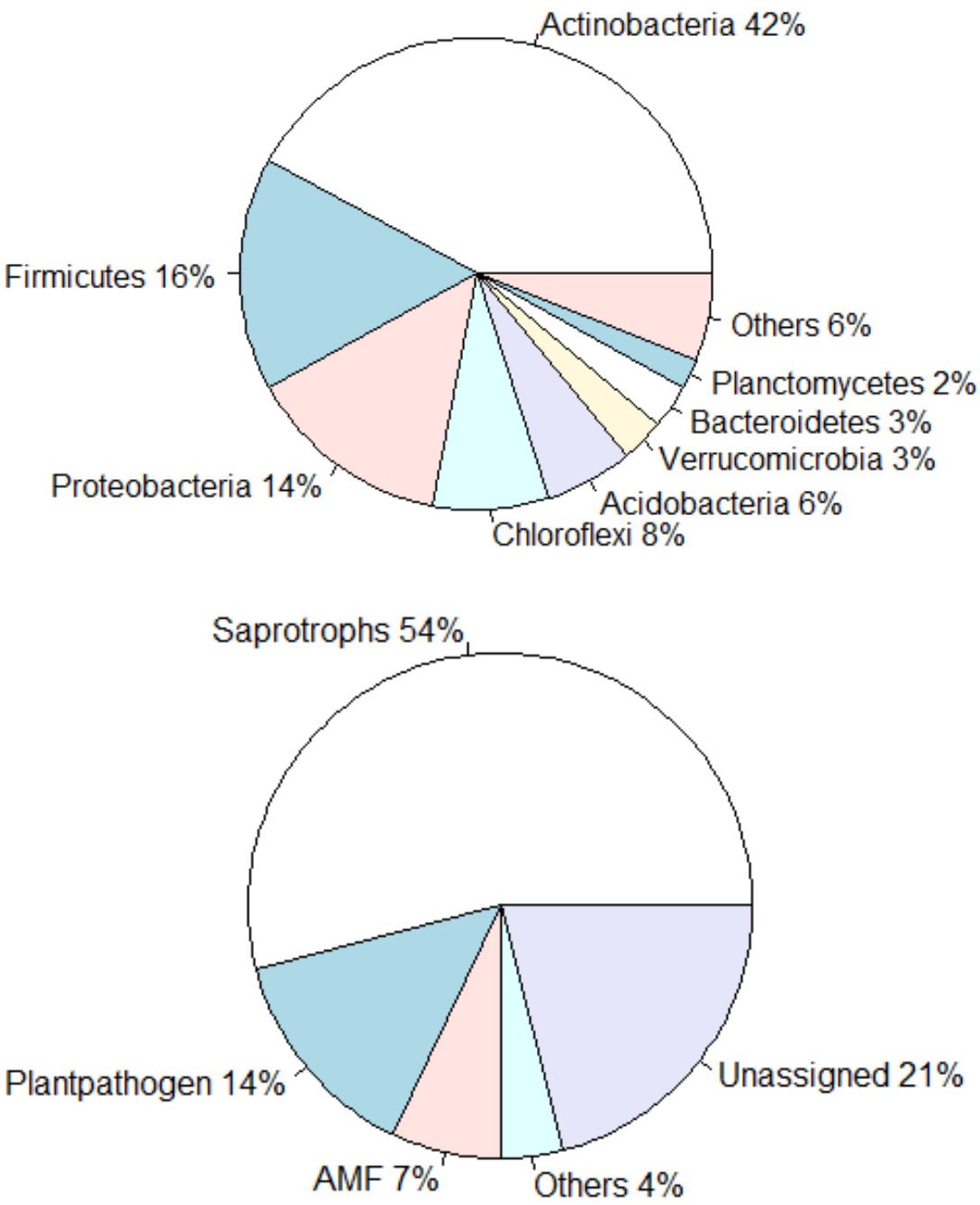


Figure S6.3 The percentage of reads belonging to particular bacterial phyla (top) and fungal functional groups (bottom) in all the samples included in the study.

Table S6.1 Information about 18 sites included in the analyses. The sites shaded in grey contained a productivity gradient.

Site ID	*Site name*	*Country*	*# of plots*	*Year: biomass sampling*	*Year: soil sampling*
Aus	Pitztal	Austria	6	2012	2017
Arg	Córdoba	Argentina	4	2012	2017
Can.e	Elginfield Observatory	Canada	2	2012	2018
Can.l	Lac du Bois	Canada	6	2012	2018
Can.o	Onefour, AB	Canada	2	2012	2017
Ch	Inner Mongolia	China	6	2017	2018
Ger.b	Bayreuth	Germany	6	2012	2017
Ger.r	Rostock	Germany	4	2015	2017
Hun	Soroksár, Fülöpháza, Battonya	Hungary	6	2012	2018
Ir.d	Damavand	Iran	2	2012	2017
Ir.j	Javaherdeh	Iran	3	2012	2017
Ir.m	Masuleh	Iran	2	2012	2017
Ir.n	North Khorasan, Golestan	Iran	6	2012/2017	2017
It	Torricchio Nature Reserve	Italy	6	2012	2017
Ken	Laikipia	Kenya	6	2012	2018
NZ	Waitati South Island	New Zealand	2	2012	2018
SA	Loskopdam	South Africa	6	2012	2017
USA	Fort Keogh	USA	6	2012	2017

Table S6.2 The coefficients and P values of selected environmental predictors in the best multiple regression model explaining bacterial and fungal community composition at high and low productivity levels.

	Bacterial communities				Fungi communities			
	Low productivity $R^2 = 0.81$		High productivity $R^2 = 0.73$		Low productivity $R^2 = 0.48$		High productivity $R^2 = 0.55$	
	Coeff.	P-value	Coeff.	P-value	Coeff.	P-value	Coeff.	P-value
Geo. distance	-	-	-	-	0.01	0.001	0.009	0.001
MAT	0.02	0.003	0.02	0.012	0.02	0.001	0.02	0.001
MAP	0.02	0.02	-	-	-	-	-	-
N deposition	0.04	0.001	0.05	0.001	0.02	0.001	0.03	0.001
Plant biomass	0.02	0.004	-	-	-	-	-	-
pH	0.05	0.001	0.03	0.005	0.03	0.001	0.03	0.001
CEC	-	-	0.04	0.001	-	-	-	-
C:N	-	-	0.03	0.001	0.009	0.01	0.01	0.017
BS	0.06	0.001	0.04	0.002	-	-	-	-
Sand	-	-	0.04	0.007	0.01	0.004	-	-
SOM	-	-	-	-	0.01	0.003	0.01	0.21
Silt	-	-	-0.04	0.005	-	-	-	-
Mg	-	-	-	-	-	-	-	-
Ca	0.03	0.003	-	-	0.02	0.009	-	-
P(avail)	-	-	-	-	-	-	0.01	0.009

Table S6.3 The percentage of reads belonging to different bacterial and fungal phyla in the study.

Bacterial phyla	% of total reads	Fungal phyla	% of total reads
Actinobacteria	41.6730	Ascomycota	62.3605
Firmicutes	16.4703	Basidiomycota	20.6830
Proteobacteria	13.7921	Glomeromycota	7.2779
Chloroflexi	7.8897	Unknown	5.8253
Acidobacteria	6.2957	Mortierellomycota	2.4193
Verrucomicrobia	3.0793	Chytridiomycota	0.4171
Bacteroidetes	2.9213	Olpidiomycota	0.4162
Planctomycetes	2.2374	Rozellomycota	0.2707
Thaumarchaeota	1.6295	Entorrhizomycota	0.1831
Gemmatimonadetes	1.5125	Mucoromycota	0.0627
Tectomicrobia	0.5941	Zoopagomycota	0.0324
Saccharibacteria	0.5819	Blastocladiomycota	0.0140
Armatimonadetes	0.4865	Kickxellomycota	0.0093
Nitrospirae	0.1723	Aphelidiomycota	0.0077
Elusimicrobia	0.1121	Monoblepharomycota	0.0071
Cyanobacteria	0.0676	Entomophthoromycota	0.0067
Latescibacteria	0.0610	Basidiobolomycota	0.0038
Unknown	0.0586	Fungi_phy_Incertae_sedis	0.0021
Parcubacteria	0.0549	Calcarisporiellomycota	0.0011
FBP	0.0522		
Chlorobi	0.0393		
Ignavibacteriae	0.0362		
TM6	0.0292		
Fibrobacteres	0.0256		
Euryarchaeota*	0.0207		
Omnitrophica	0.0165		
Chlamydiae	0.0139		
Candidatus	0.0118		
BRC1	0.0101		
Deinococcus.Thermus	0.0094		
Spirochaetae	0.0083		
FCPU426	0.0083		
Peregrinibacteria	0.0058		
Woesearchaeota*	0.0044		
Microgenomates	0.0041		
BJ.169	0.0027		
WS2	0.0023		

Tenericutes	0.0015	
Hydrogenedentes	0.0011	
GAL15	0.0010	
RBG.1	0.0009	
WWE3	0.0008	
Fusobacteria	0.0007	
CPR2	0.0005	
Gracilibacteria	0.0004	
SBR1093	0.0004	
Bathyarchaeota*	0.0004	
Nitrospinae	0.0003	
Aenigmarchaeota*	0.0003	
SR1	0.0003	
Deferribacteres	0.0002	
Parvarchaeota*	0.0002	
Lentisphaerae	0.0001	
PAUC34f	0.0001	
Aminicenantes	0.0000	
WS1	0.0000	
Synergistetes	0.0000	

* Archaea

Appendix 6.1 – Microbial community composition predictors at different productivity levels

At high productivity, environmental factors (including climate, soil, geographical distance and plant biomass) explained 72% of the variation in bacterial community composition, and the most influential individual predictor was base saturation explaining alone 35% of the variation. Plant communities explained 34% of the variation alone and when added to the best model of environmental factors they increased the variation explained to 77%. Environmental factors explained 82% of variation at low productivity, where the most important individual factor was again base saturation ($R^2 = 0.58$). Plant communities added 5% of variation when included to the selected model (and alone they explained 34% of the variation).

At high productivity, environmental factors explained 55% of the variation in fungal community composition and the most influential individual predictor was geographical distance ($R^2 = 0.24$). Plant communities alone explained 67% of the variation and when added to the model with abiotic factors the total variance explained increased from 55% to 76%. At low productivity, 48% of the variation was explained by the best model with environmental factors, where geographical distance explained the most individual variation ($R^2 = 0.18$). When plant community composition alone explained 78% of the variation and when added to the best mode with environmental variables, the total amount of variation increased from 48% to 78%.

Picture credits:

Chapter I:	Dajana Radujković
Chapter II:	Modified from: Michelle Maria (Creative Commons Attribution 3.0 Unported license)
Chapter III:	Dajana Radujković
Chapter IV:	PLECO microbial lab
Chapter V:	Dajana Radujković
Chapter VI:	Miguel Portillo-Estrada
Chapter VII:	Modified from: Alan Vernon (Creative Commons Attribution 2.0 Generic license)