# Below-Ground Impacts on Microbial Diversity of Afforestation and Woodland Succession in Great Britain

**Justin G D Byrne**

**Doctor of Philosophy**

**Newcastle University**

**November 2020**

**Main Text word count: 63,166**

## Abstract

Woodland restoration on farmland is gaining support as a method of habitat restoration and carbon sequestration. To understand the consequences of afforestation on soil microbial communities we require comparisons across agricultural, afforested, and established woodlands that differentiate changes in total soil diversity and diversity pertaining to soil functioning. I examine two natural-experiment, afforestation chronosequences describing changes to bacterial and fungal diversity in soils and on decomposing leaf litter using metabarcoding to identify taxa. The use and applications of metabarcoding are described in Chapter 1.

Chapters 2 and 3 report the abiotic and microbiological differences between agricultural and woodland soils over an afforestation and succession gradient. I study land use differences across 21 Woodland creation & Ecological Networks (WrEN) project sites – an afforestation chronosequence - with mixed-effect and multivariate models. The next chapters model how leaf litter decomposition (Chapter 4) and how decomposing litter associated microbes (Chapter 5) vary across woodland age categories (young, mature, and ancient) in 27 woodlands in the North East of England, across five tree species' litter, during spring and summer.

In the first experiment I show key soil changes to carbon and nitrogen concentrations, and to C:N ratio across my afforestation gradient. I also demonstrate the rapid differentiation of woodland microbial communities from agricultural ones, with additional changes resulting from soil properties such as pH. In the second experiment I observe notable, but non-significant increases in decomposition in ancient woodland, as well as significant effects of species, sampling season, and woodland age on microbial diversity, beyond that of the changes expected from soil properties alone. In chapter 6 I discuss how these results integrate into the existing literature on the succession of decomposer communities over long timescales and their implications for woodland management. The work highlights the value of both afforestation and conservation of older woods.

For Jennifer, who supports me in all my labours

## Acknowledgements

I would like to acknowledge innumerable friends, family, and colleagues who have provided insight and support over many years of academic study. Particularly those who have assisted me in the field, in the lab, or behind a computer screen: Joshua Varty, Alex Byrne, Dimitrios "Dimitris" Petsopoulos, Johanna Weston, Dr Lisa Malm, Dr Claire Branston, Dr Fredric Windsor, Dr Kirsten Miller, Mathew Pindar, and Eleanor Collinson. Special thanks to Aifionn Evans for invaluable laboratory assistance in measuring litter dry weight and soil texture. Thanks to Dr Peter Leary for training me in bioinformatics and providing scripts to build analyses upon.

I cannot thank Dr James Kitson enough for training me in molecular techniques and providing many hours of advice and assistance. Additionally, I would like to thank Dr Julia Cooper for advising on the design of the experiment and reading drafts, and Dr Kirsten Wolff for training on laboratory techniques. I would like to thank Sain Atkinson, Dr Christine Tansey, Dr Nick Atkinson, Dr Emma Gilmartin, and all other staff at The Woodland Trust who have provided advice on experimental design or suggestions on chapter drafts. Thanks To Professor Lynne Boddy for advice on conducting litter bag experiments. Thanks to Dr Ian Head for providing me with the source material for the bacterial molecular positives.

Many thanks to Dr Frank Ashwood and Dr Kevin Watts, and other researchers from The WrEN project or Forest Research, for collaborating with me on the first two chapters.

Thanks to the Varty family, the Evans family, The Woodland Trust, The Natural History Society of Northumbria, The National Trust, Durham Wildlife Trust, Durham County Council, and Gateshead Council for allowing me to work at their properties.

I would like to commend Newcastle Carers and Newcastle University for their wellbeing support services that have helped maintain me on an even keel over the years. My parents, the Byrne family, and the Bray family have also helped me in so many ways, underpinning my success.

I give my sincere gratitude to The Woodland Trust for funding me, supporting my professional development, and inspiring the questions discussed herein. Additional thanks to the N8 Partnership for funding provided through their pump-priming scheme. Most of all I would like to than the members of The Woodland Trust for donating the money in the first place.

Finally, my sincere thanks to Professor Darren Evans for providing me with the opportunity to investigate such fulfilling research questions and providing excellent professional and interpersonal support on countless occasions.

# Contents

# Chapter 1. Introduction: The Emergence and Maturation of Soil Science

## 1.1 A History of Soil Science

The scientific study of soils begins in the mid-19th Century with advances such as the mineral theory of plant nutrition and soil biology (Brevik, 2005). Two branches of soil science have dominated for much of its history: maximising plant cover of soil and studying soil as a physical material (Berthelin *et al.*, 2006; Warkentin, 2008). Soil biology emerged around the same time as the development of soil science, with publications from Darwin, Müller, and Drummond; prior to which soil organisms had solely been studied from a taxonomic perspective (Berthelin *et al.*, 2006). At all stages, the focus of soil biology has been determined by technological advancement and availability. Around this time, the production of powerful microscopes and then the developments in cell culture techniques lead research into a mini-renaissance (Berthelin *et al.*, 2006). Microscopy, micromorphology, and later, cell culture, accompanied advances in agronomic microbiology including the discovery of mycorrhizae, bacterial nitrification, soil-forming biological processes, and the role of organisms in nutrient and mineral cycling (Berthelin *et al.*, 2006). These powerful methods changed the world, revolutionising medicine with new understandings of pathogens and the discovery of antibiotics. By 1923, overconfidence in these methods resulted in textbook proclamations that no microorganism could be classified before its characteristics in culture were described (Bergey *et al.*, 1923). However, before the halfway mark of the 20th century, confidence in isolate studies was eroding (Winogradsky, 1949). In his summary of his life work, Winogradsky set out his “principles of soil microbial ecology”, laying the foundation of modern soil biology. Here he suggested that it was “time to distinguish physiological and chemical studies from the ecological research on the microbial dynamics in natural environments”. During the 1980’s the failure of culture and micromorphology methods to describe even a fraction of microbial diversity had become apparent (Pace *et al.*, 1986; Berthelin *et al.*, 2006), new tools were needed to probe deeper into the “black box of soil microbial diversity” (Tiedje *et al.*, 1999). The late 20th century was correspondingly dominated by the development and use of new technological profiling tools of soil communities, including microbial biomass, enzyme activity and eventually genetic profiling approaches. After less than a century we find ourselves in a similar situation; developments in genetics, gene sequencing, and computing power have set

the stage for a new era of soil biology. As Winogradsky noted at the last paradigm shift, these new tools do not replace those that come before, but merely allow new questions to be answered that were previously recalcitrant to study. If history repeats itself, genetic tools are set to dominate soil biology for the next half century.

Changes in the last half-century dramatically shifted how we view soils from physical structures to ecological habitats (Warkentin, 2008). In all aspects of soil science, we have witnessed increasingly holistic approaches to their study that acknowledge their physical and ecological complexity. Recently we have seen an increasing need to interact with non-scientific disciplines and interest groups (Warkentin, 1999) as the huge socio-economic tasks involved in meeting the modern challenges of soil erosion, global food security, and climate change have become apparent (McBratney *et al.*, 2014). The challenges facing soil science today include controlling soil degradation (fertility loss, nutrient loss, erosion, desertification, compaction, salinisation, acidification, and toxification) (Schröter *et al.*, 2005; Dominati *et al.*, 2010), managing soils for water filtration and flood mitigation in order to meet global water security challenges (Dominati *et al.*, 2010; McBratney *et al.*, 2014), facilitating soils carbon sequestration (Lal, 2004), minimising greenhouse gas emissions from soils (Dominati *et al.*, 2010), maintaining soil biodiversity (Dominati *et al.*, 2010), and increasing global agricultural productivity to provide food security (Godfray *et al.*, 2010).

Soils, and their organisms, play a crucial role in facilitating the global cycling of nutrients and minerals that enable life on earth (Prentice *et al.*, 2001; Fowler *et al.*, 2013). Though accounting for only 14% of global reactive nitrogen production, microbially mediated terrestrial biological nitrogen fixation is essential to the support of natural and semi-natural systems that are not the target of intensive fertiliser use (van der Heijden *et al.*, 2008). Root associated mycorrhizal and decomposers play an especially important role in phosphorus cycling (Filippelli, 2008). Finally, soil carbon fluxes facilitated by soil microorganisms form an essential part of the carbon cycle (Burns *et al.*, 2013). These cycles, along with other soil processes and properties are a huge natural capital, providing many ecosystem services to humans (Dominati *et al.*, 2010). Future civilizations rely on our effective management of soils and soil organisms today, as Franklin D Roosevelt put it in his *Letter to all State Governors on*

*a Uniform Soil Conservation Law* (1937): "a nation that destroys its soil destroys itself". However, the fundamental ecology supporting soil security needs to be better understood.

### 1.2 Molecular Approaches Modify Ecological Research

Concurrent advances in fields such as molecular ecology, computer science, and remote sensing have changed the capacity of ecologists to generate and analyse ecological data quickly, cheaply, and accurately (Baird & Hajibabaei, 2012; Coissac *et al.*, 2012; Bohan *et al.*, 2017). The continuous progress in these technologies is leading to them becoming increasingly accessible (Baird & Hajibabaei, 2012; Ji *et al.*, 2013a; Bohmann *et al.*, 2014). Ongoing discussions about the challenges and opportunities present in each new method or platform (Brosi *et al.*, 2015; Cowart *et al.*, 2015; Pompanon & Samadi, 2015; Bucklin *et al.*, 2016; Blanckenhorn *et al.*, 2016) are needed but researchers should not hold off from taking advantage of new techniques until an optimal approach is decided. Such an approach will likely remain illusory for some time. Unfortunately, the pressing demands on ecologists to deal with huge challenges at a global scale necessitates the use of imperfect tools to inform decision making (Baird & Hajibabaei, 2012). In reality, the uncertainty and drawbacks of new approaches have analogues in current ecological techniques, such as uncertainty around abundance estimates, sampling effort, inability to confirm species absence, cryptic species, and the limitations of temporally wide sampling periods. But the advantages of these methods are that they increase researchers' capacity to answer questions that require huge data sets (Derocles *et al.*, 2018).

Ecological questions are often complex, requiring large datasets or meta-analysis to separate out effects from noise and confounding variables. Traditional ecological techniques have practical limits relating to time, money, accuracy, and the availability of expertise and equipment (Pawlowski *et al.*, 2014). These limitations have shaped the kind of questions that ecologists are able to answer, and so they have shaped the kinds of questions that ecologists ask. Collecting more data does not always improve our ability to answer a given question, but given access to larger pools of data we can ask different questions that require more data to come to satisfactory statistical conclusions (Derocles *et al.*, 2018). Technological and computational advances continue to allow us to collect and process larger amounts of

biodiversity data at a lower cost than traditional data collection methods (Baird & Hajibabaei, 2012). Notably, advances in DNA sequencing, ecological statistics and bioinformatics are allowing researchers to ask more sophisticated ecological questions and produce the large amounts of data needed to answer them. Researchers without extensive training in computer science or biochemistry are now able to make use of advanced molecular or computational approaches by drawing on the growing literature and training resources on the topics. Progress in automated remote sensing, cloud computing, and machine learning are also areas that will greatly increase the capacity of ecologists to collect and analyse huge datasets (Bohan *et al.*, 2017). A great many traditional ecological questions are currently best answered using traditional approaches; but with time, automated data collection approaches will supersede traditional ones for even simple ecological surveys. One area of research revolutionised by molecular methods is soil biology (Handelsman, 2004). Now, new problems in microbial ecology can be approached with high taxonomic resolution.

### 1.3 Soil Microbe Multifunctionality

One such problem is disentangling the relationship between microbial diversity and ecosystem functioning (Antwis *et al.*, 2017). It is clear that soil microorganisms play an important role in driving many ecosystem services, such as in nutrient cycling, plant productivity and pest control, or in engineering an environment conducive to their support, either through bioremediation, or improving soil quality (Filip, 2002; Arias *et al.*, 2005; Hooper *et al.*, 2005; Barrios, 2007). However, linking soil microbial diversity to the services they provide is not trivial (Allison & Martiny, 2008), due to high diversity and potential functional redundancy or multifunctionality. Some suggest that microbial diversity or biomass may indicate good soil health, which assumes a relationship between diversity and functioning (Filip, 2002; Arias *et al.*, 2005).

These relationships are not always detected or may be inconsistent, such as in an experiment of functional processes in grassland soils treated with chloroform fumigation, which detected specific changes in processes such as nitrification, but no general relationship between diversity and functioning (Griffiths *et al.*, 2000). In Wagg *et al.* (2014) ecosystem functions were measured across fractionated soil samples of decreasing diversity, showing positive

relationships between litter decomposition and microbial diversity and negative relationships between diversity and both $N_2O$ emission and phosphorus leaching. Similarly, potential nitrification activity in soil was found to reduce when soil microbial diversity was reduced using dilution methods (Philippot *et al.*, 2013). Mycorrhizal diversity has been positively associated with plant diversity and productivity (both root and shoot biomass, and phosphorus uptake) at low levels of diversity generated by soil inoculation (van der Heijden *et al.*, 1998). Measures of community structure have revealed the responses of microorganisms to changes such as woodland development (C. A. Creamer *et al.*, 2016; Mackay *et al.*, 2016), showing that microbial biomass and diversity measures correlate positively with increased carbon sequestration. A variety of functional responses, both positive and negative have been observed in response to decreasing species richness, with few clear, consistent trends. Overall, linking microbial diversity to some soil processes, like uptake of nitrogen by fungal mycorrhizae to transfer to plants (Van Der Heijden *et al.*, 2006; van der Heijden *et al.*, 2015; Field *et al.*, 2020), have been more tractable than others, such as decomposition. The role of microbes in decomposition has been harder to describe and quantify unambiguously, partly due to the difficulty of isolating individual litter-microbe interactions t (Hättenschwiler *et al.*, 2005; Hooper *et al.*, 2005). This may be because the biological contribution of diversity to processes such as decomposition saturates at low species richness (Gessner *et al.*, 2010). However, authors have noted that species composition and interactions between species may be just as important as richness, if not more so (Hooper *et al.*, 2005; Nielsen *et al.*, 2011; van der Wal *et al.*, 2013).

Experimental work detailing the relationships between soil organisms and associated processes is a growing field. However, substantial challenges exist in studying soil *in situ* that include balancing experimental control and reducing artificiality, but especially due to the extremely high diversity of soils microbes that are especially difficult to characterise (Fitter *et al.,* 2005). The majority of microbial species remain undescribed and even the species concept starts to break down within some microbial taxa (Handelsman, 2004; Sharma *et al.*, 2015; Tiedje *et al.*, 1999). Identification through culture and microscopy is insufficient to describe most soil microbes at high taxonomic resolution (Giller, 1996), and community profiling tools grew increasingly popular at the turn of century to account for this. Temperature or Denaturing Gradient Gel Electrophoresis (TGGE/DGGE), phospholipid fatty acid (PLFA)

analysis, and terminal restriction fragment length polymorphism (T-RFLP) all provided methods of assessing community diversity profiles based upon microbial genetic sequences (Handelsman, 2004) with limitations. For example, profiling methods based on non-genetic biological molecules, such as PLFA analysis, provide limited taxonomic resolution risking generalising the role of lower-level taxa to higher level taxonomic groups (Frostegård *et al.*, 2011). Community profiling tools cannot determine which taxa present in one sample are present in another at high taxonomic resolution. As such it is difficult to attribute changes in ecosystem functioning to changes in community composition. Next-generation genetic tools, with their potential for superior taxonomic resolution, improve in this area.

**1.4 What is DNA Metabarcoding**

Advances in DNA based methods of species identification have enabled the detection of species based upon their unique sequence at one of many generally recognised genetic loci, known as a DNA barcode (Hebert *et al.*, 2003). By parallel sequencing bulk samples of DNA, either pooled from multiple individuals or of environmental DNA (eDNA) shed into the environment, DNA metabarcoding enables the taxonomic characterisation of large numbers of individuals or whole communities (Taberlet *et al.* 2012; Yu *et al.* 2012; Ruppert *et al.* 2019). The field is establishing practical, verifiable methods for investigating biodiversity to be used alongside traditional biodiversity surveys (Richardson *et al.*, 2015; Zimmermann *et al.*, 2015; Hanfling *et al.*, 2016; Hawkins *et al.*, 2015; Nichols *et al.*, 2016; Pawlowski *et al.*, 2016; Zaiko *et al.*, 2016). Recent advances in DNA sequencing technology have invigorated research in the area as the price and data output of next-generation sequencing has improved (Shokralla *et al.*, 2012; Reuter *et al.*, 2015). Metabarcoding methods have already been tested in a range of environments and incorporated into study designs in various areas including aquatic systems (Bringloe *et al.*, 2016), soils (Floyd *et al.*, 2002), deep-sea sediments (Guardiola *et al.*, 2015), probiotic drinks (Reva *et al.*, 2015), and ship ballasts (Zaiko *et al.*, 2016).

Numerous genomic loci have been suggested that maximise either the taxonomic resolution or the taxonomic breadth of identification. Frequently the loci chosen for a project is based upon a trade-off between these two requirements depending on the objectives of the work. Online databases, such as the International Barcode of Life (iBOL), have been created to make

identified DNA accessions available to researchers. These have often been derived from samples of tissue of a known taxonomic origin. Additionally, an array of bioinformatics software has been developed to minimise errors in species identification (Bik *et al.*, 2012; Coissac *et al.*, 2012; Boyer *et al.*, 2016). Sequenced barcodes are compared to sequences found in these databases. Samples that cannot be matched to a verified gene sequence may be classified to a Molecular Operational Taxonomic Unit (MOTU or OTU) that may still allow work to distinguish between taxonomic units across samples.

Aspects of metabarcoding remain contentions. Current methods often use PCR amplification to increase the concentration of amplicons in a sample. Variable mismatches in the binding region of the genetic locus mean that certain taxa may be disproportionately represented in the final sequenced data (Piñol *et al.*, 2015). Taxa may also vary in the number of copies of the barcoding region present in their genome, and so become proportionally over-represented in the final dataset. Detecting species present at very low biomasses may also be difficult, as they may be filtered out at the quality control stage of the bioinformatics when low frequency sequences are removed either as read errors or to streamline analysis. In preparation for bioinformatic analysis reads are discarded that may result from the chemical processes of the sequencing machine or from read errors. In principle, distinguishing erroneous sequences from truly rare sequences presents a challenge. Pauvert *et al.* (2019) compares various bioinformatic thresholds, such as species delineation thresholds and the treatment of sequences occurring at low frequencies, showing large effects on the final biodiversity of samples Ecologically, rare sequences may impact their community to a greater degree than their abundance might suggest. Without biological information to base decision making on, the removal of rare reads is a bioinformatic trade off that prevents analyses from becoming overly cumbersome and computationally intensive. PCR free methods for conducting DNA sequencing are also being investigated (Taberlet *et al.*, 2012), with the potential for a shift to occur towards technologies like Oxford Nanopore's *MinION* in the future (Oikonomopoulos *et al.*, 2016). Recent work to improve the cost effectiveness and efficiency of metabarcoding work has highlighted methods of "nested-tagging" that allow for multiple samples to be pooled and processed simultaneously while retaining information relating to their sampling (Binladen *et al.*, 2007; Kitson *et al.*, 2018). The use of "nested

tagging" methods will only increase our ability to produce large, multi-sample assessments of biodiversity.

### 1.5 DNA Metabarcoding Adds Vital Identity Information to Soil Community Profiles

The power and flexibility of metabarcoding has facilitated the use of environmental DNA (eDNA) in biodiversity surveys (Baird & Hajibabaei, 2012; Ruppert *et al.* 2019). Surveys of eDNA can be quicker, cheaper, and of equal or better quality than traditional monitoring methods (Baird & Hajibabaei, 2012; Ji *et al.*, 2013a; Liu *et al.* 2020). The rapidity of sampling and the ease of storing eDNA mean that huge amounts of sampling can be conducted in short periods of time and analysed at leisure. Long sampling windows are problematic, eDNA quickly captures a community snapshot that can overcome may resolve this. The adaptability of the method has seen it employed in practical settings to detect invasive species (Brown *et al.*, 2016; Zaiko *et al.*, 2016; Madden *et al.* 2019) and pathogens (Sapp *et al.*, 2016; Pauvert *et al.*, 2020). Metabarcoding is also being applied to survey community biodiversity (Beng *et al.*, 2016; Yoon *et al.*, 2016; Derocles *et al.* 2018), and used to compare the diversity of species across treatments (Foulon *et al.*, 2016). Now, studies are investigating the effects of species interactions at the community level (Bringloe *et al.*, 2016; Kitson *et al.* 2018; Ruppert et al. 2019). The weight of current work indicates that quantitative assessments of biodiversity remain difficult (Elbrecht and Leese 2015; Blanckenhorn et al. 2016; Deagle et al. 2019; Lamb et al. 2019). However, many researchers have attempted to validate methods for incorporating metabarcoding into biodiversity monitoring (Hanfling *et al.*, 2016; Ji *et al.*, 2013b; Zaiko *et al.*, 2015).

By identifying sequences to a high taxonomic resolution and preserving identity data it may be possible to relate the presence of an organism in a system to its natural history and function. Such approaches will need to account for the capacity for horizontal gene transfer in bacteria and would need to justify assumptions of linking diversity to function. Although these tools are in their infancy, several have been developed to infer functioning of taxa regardless of their accessioning into whole-genome databases (Langille et al., 2013; Wemheuer et al., 2018; Douglas et al., 2020). Due to the complexity of microbial genetic, including horizontal gene transfer, further empirical validation will be needed to determine

the utility of this approach. Experimental data is improved by agreement on standard barcode loci, at which large numbers of accessions can be compiled and accessed by researchers. The reproducibility of barcoding methods and the permanent accession of a study's results makes comparisons across studies simpler, however laboratory and bioinformatic approach does influence results (Pauvert *et al.*, 2019). Early studies may even be revisited in the future and reanalysed once sequence databases are more complete.

The complex, difficult to characterise diversity of the soil microbiome is well suited to examination with next-generation molecular approaches as they can characterise the diversity of soils better than previous methods and generate the huge amount of data needed to sample soils completely (Handelsman, 2004). Increasing the breadth of taxa that can be characterised in studies allows soil communities to be explored in greater detail. However, the presence of species on their own is not enough information to understand an ecological community. The retained taxa identity data across multiple samples means that community composition can be directly compared. Relationship information, in the form of ecological networks, is required to understand how belowground communities function and respond to change and this has led to some metagenomic studies representing outputs in the form of network graphs (Toju *et al.*, 2014; Banerjee *et al.*, 2016; Shi *et al.*, 2020). These are sometimes generated using an algorithm to predict interactions based on patterns of co-occurrence (Barberán *et al.*, 2012), or they may be observed or inferred directly as more common in bi-partite network approaches of defined interactions (Bennett *et al.*, 2013). Networks require large amounts of data to construct and are improved when compartments are defined to a high taxonomic resolution. Traditional methods of constructing networks have been labour intensive, prone to sampling biases, and vulnerable to missing cryptic species (Evans *et al.*, 2016), the effort involved in observing enough interactions to infer species relationships with confidence is substantial. Metabarcoding methods go some way in dealing with these problems and enable the construction of highly resolved interaction networks, so long as the information captured relates to species interactions. Nevertheless, metabarcoding approaches and the ability to infer interactions from co-occurrence from them does not negate the necessity of careful experimental design. Targeted experiments of compartments of soil diversity can better characterise the relationships of biodiversity to specific ecological functions.

## 1.6 Future Directions

### *1.6.1 Ecological networks:*

Networks enable researchers to understand complex processes occurring across an ecological community by mapping species interactions. Network metrics such as connectance, the density of links between species, relate to properties of the whole system, providing community level descriptors and introductions to these metrics already serve as excellent resources for newcomers to network analysis (Proulx *et al.*, 2005; Tylianakis *et al.*, 2010; Farine & Whitehead, 2015; Guimarães, 2020). A network, or graph, represents individual species or groups of species in a node, sometimes called a vertex. The relationships between species are indicated by a link, which may be referred to as an edge. Links may represent a trophic interaction, pollination relationship, or any type of interaction. These links may indicate the presence or absence of an interaction alone, or an interaction strength, frequency, direction, or other information. This extra information can dramatically change the importance of certain species within a network and completely alter the network's metrics (Scotti *et al.*, 2007; Jordán *et al.*, 2008). Computationally, this information may be stored in an interaction matrix (Suweis *et al.*, 2013) with information relating to each relationship's strength, direction, or even more complex descriptors. adding weighting and direction to network links has meaningful effects on network analysis, with measures such as node centrality varying greatly between weighted and unweighted networks (Scotti *et al.*, 2007) and weight may also affect the degree to which network measures correlate with functionally important network indices (Jordán *et al.*, 2008). Including link direction in a network changes its structure and can alter the conclusions of the network's analysis (Bascompte *et al.*, 2006). Network analysis methods provide a number of descriptors that refer to the relative importance and place of an individual node within a network, such as the number of links that connect to a certain node i.e. its degree. Otherwise, descriptors may indicate properties possessed by a network as a whole, such as a measure of how much a network is structured into separate clusters, i.e. its modularity (Tylianakis *et al.*, 2010; Deng *et al.*, 2012; Thompson *et al.*, 2012). Scientists focusing on concepts of stability and robustness in ecology have found that these notions can be related to the structural properties of networks (Montoya *et al.*, 2006; Thébault & Fontaine, 2010; Tylianakis *et al.*, 2010). For example, the distribution of species' degrees within a network has been shown to affect if, and how many, secondary extinctions occur following

the loss of a species from a network i.e. its robustness to extinction (Tylianakis *et al.*, 2010). Multipart processes might be broken up into constituent bipartite networks and then assembled into networks of networks (Pocock *et al.*, 2012; Pilosof *et al.*, 2017) to determine vulnerable links.

Merging metabarcoding and the construction of ecological networks would seem to provide clear advantages in both fields, with metabarcoding studies benefiting from the analytical strengths of network analysis in describing communities and networks studies benefiting from the rapid, broad scale assessment of community composition (Bohan *et al.*, 2017). The ease and speed of eDNA collection can bring down the cost and effort involved in constructing networks, enabling increased focus on replicating experiments in time and space. This is important as even when numerous species are able to support an ecosystem function, if they are never present together at the same time of year our understanding of the system may be incorrect. Pollinator systems have shown remarkable annual variation in interactions, though overall network structure remains relatively constant (Dupont *et al.*, 2009; Kemp *et al.*, 2016). The ease of collecting eDNA samples in short sampling periods solves these issues.

### *1.6.2 Bipartite networks of species subsects focus on key actors*

Efforts to relate biodiversity to ecosystem function are closely related to the topic of redundancy in the biological composition of ecosystems (Walker, 1992). Functionally redundant species are those that perform similar roles within their communities, such that the loss of one of them will not greatly undermine an ecological process (Rosenfeld, 2002). Classic papers have used species level measures (e.g. richness) to determine redundancy (Tilman *et al.*, 2001), but in cases where a process occurs between interacting partners (e.g. pollinators and plants), a community level, network approach may provide additional insights (Kaiser-Bunbury *et al.*, 2017). This could be taken further by examining functional redundancy in terms of functionally redundant interactions, rather than redundant species. As the science of network rewiring and network structure fluctuations over time, this approach would allow ecologists to assess the impact of changes in interactions but not species or vice versa.

Networks may be constructed selectively for a specific compartment of ecological diversity, such as from a root tip, or they may be constructed non-selectively across compartments using environmental sampling of eDNA, such as from a water or soil sample. Studies of soil networks have generally involved generic networks (Zhou *et al.*, 2010; R. E. Creamer *et al.*, 2016; Morriën *et al.*, 2017), where links are inferred between species based upon co-occurrence (Toju *et al.*, 2017). As many types of relationships are represented in these networks, it can be very difficult to identify what results are important to specific ecosystem functions. Selective, bipartite networks can be assembled between functional species on either side of an ecosystem process, e.g. between plants and their associated mycorrhizae (Bennett *et al.*, 2013; Toju *et al.*, 2014). Using this approach it becomes possible to relate community level metrics to rates of ecosystem processes occurring at the same scale. De Vries *et al. (2013)* have demonstrated links between nutrient cycling services and food web structure in agricultural soils, using food webs assembled with traditional identification methods and PLFA analysis. Reviews have already called to merge metabarcoding and network analysis (Evans *et al.*, 2016; Bohan *et al.*, 2017; Toju *et al.*, 2017; Derocles *et al.*, 2018). I argue that selective, bipartite networks provide important information about the functional diversity of species and species interactions. So far, each of these tools has been linked with the others independently, functions have been associated with generic networks (R. E. Creamer *et al.*, 2016), metabarcoding has been used to produce selective, bipartite networks (Toju *et al.*, 2014), and metabarcoding data has been linked to function (Strickland *et al.*, 2009). Metabarcoding offers a clear path to quick, highly replicated, highly resolved bi-partite network construction. I suggest that researchers should consider using bipartite networks, constructed through DNA metabarcoding to study the biological provision of ecosystem functions.

To apply these methods to soil systems, I have identified an ecosystem process involving a subset of the community - the decomposition of leaf litter. Here a bipartite network could be assembled between the leaf litters of different species of trees and microbial OTUs identified by metabarcoding. To achieve this, leaf litter could be decomposed in single-species mesh bags. Critically this selectively isolates leaf litter and fungi from larger decomposers. At the end of the study, decomposition rate and fungal diversity could be determined from the remaining material. Bipartite networks of tree species and fungi might be analysed for

generality, specialisation, or modularity allowing multiple decomposition networks (separated in time or space) to be compared. Individual nodes might be identified as keystone species, or of particular importance or vulnerability.

Loosely targeted co-occurrence networks generated from non-selective environmental sampling remain extremely useful for answering ecological questions at the community level. However, when it is crucial to relate specific taxa to a particular process, selective networks become essential. In some cases, many taxonomic groups or families will take part in a process. In the above example, decomposition rates will be influenced by bacterial and fungal microbial decomposers, but also larger taxa such as insects. To assess the contribution of each separately, it is important to isolate them in the experimental design. This might be done by decomposing leaf litter in mesh bags with progressively larger mesh sizes. Selective design does not preclude the eventual construction of a larger network of bipartite networks (Pocock *et al.*, 2012).

This approach is adaptable to many systems but is particularly relevant to belowground communities. Below ground ecosystems have presented particular challenges that have made their description difficult; such as unculturable organisms (Handelsman, 2004), extraordinary diversity (Giller, 1996), and spatial structure variability (Ettema, 2002). The task soil ecologists now face in understanding ecosystem functions within soils faces extreme logistic difficulties that are not present in above ground systems. In these systems, bipartite networks assembled from metabarcoding data provide large amounts of data, highly resolved taxonomic descriptions, and a method of capturing complex, community level data.

### 1.7 Understanding Woodland Soils

This project is funded largely by The Woodland Trust from charitable contributions. It focuses on the differences in soil diversity and functioning between agricultural, afforested, and mature woodland sites. In the following chapters I will set out why these land use transitions are of importance to UK conservation policy and planning, and why they are important to the ecology of the UK. In each chapter I will describe how little is known of how woodland soils

change as they age, and few good comparisons of the microbial diversity of farmland, recently afforest woodland, and ancient woodland exist. Without a strong understanding of the consequences, land management decisions are affecting the age structure of our woodlands; older woodlands are being regularly grubbed out (Rackham, 2008) and UK woodland planting is set to continue at pace in the coming decades (DEFRA, 2018). Existing policy frameworks recognise the need to mitigate for the potential diversity loss of microbes in ancient woodland soils without understanding what that diversity is or why it matters (HS2 Ltd, 2020). At the end of each chapter, and throughout the thesis, I have endeavoured to summarise the information for forestry practitioners, not just academic readers. I have devoted a significant portion of my discussion chapter to this as well. This reflects the invaluable contribution The Woodland Trust and other practitioners have made to this project.

**1.8 References**

Allison, S.D. & Martiny, J.B.H. (2008) Colloquium paper: resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences of the United States of America*. 105 Suppl 111512–11519.

Antwis, R.E., Griffiths, S.M., Harrison, X.A., Aranega-Bou, P., Arce, A., Bettridge, A.S., Brailsford, F.L., de Menezes, A., Devaynes, A., Forbes, K.M., Fry, E.L., Goodhead, I., Haskell, E., Heys, C., James, C., Johnston, S.R., Lewis, G.R., Lewis, Z., Macey, M.C., *et al.* (2017) Fifty important research questions in microbial ecology. *FEMS microbiology ecology*. 93 (5).

Arias, M.E., Gonzalez-Perez, J.A., Gonzalez-Vila, F.J. & Ball, A.S. (2005) Soil health -- a new challenge for microbiologists and chemists. *International microbiology: the official journal of the Spanish Society for Microbiology*. 8 (1), 13–21.

Baird, D.J. & Hajibabaei, M. (2012) Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular ecology*. 21 (8), 2039–2044.

Banerjee, S., Kirkby, C.A., Schmutter, D., Bissett, A., Kirkegaard, J.A. & Richardson, A.E. (2016) Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. *Soil biology & biochemistry*. 97188–198.

Barberán, A., Bates, S.T., Casamayor, E.O. & Fierer, N. (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME journal*. 6 (2), 343–351.

Barrios, E. (2007) Soil biota, ecosystem services and land productivity. *Ecological economics: the journal of the International Society for Ecological Economics*. 64 (2), 269–285.

Bascompte, J., Jordano, P. & Olesen, J.M. (2006) Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science*. 312 (5772), 431–433.

Beng, K.C., Tomlinson, K.W., Shen, X.H., Surget-Groba, Y., Hughes, A.C., Corlett, R.T. & Slik, J.W.F. (2016) The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Scientific reports*. 613.

Bennett, A.E., Daniell, T.J., Öpik, M., Davison, J., Moora, M., Zobel, M., Selosse, M.-A. & Evans, D. (2013) Arbuscular mycorrhizal fungal networks vary throughout the growing season and between successional stages. *PloS one*. 8 (12), e83241.

Bergey, D.H., Harrison, F.C., Breed, R.S., Hammer, B.W. & Huntoon, F.M. (1923) *Bergey's manual of determinative bacteriology*. Baltimore: Williams & Wilkins Co.

Berthelin, J., Babel, U. & Toutain, F. (2006) 'History of Soil Biology', in Benno P. Warkentin (ed.) *Footprints in the Soil: People and Ideas in Soil History*. [Online]. Elsevier. pp. 279–306.

Bik, H.M., Porazinska, D.L., Creer, S., Caporaso, J.G., Knight, R. & Thomas, W.K. (2012) Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in ecology & evolution*. 27 (4), 233–243.

Binladen, J., Gilbert, M.T.P., Bollback, J.P., Panitz, F., Bendixen, C., Nielsen, R. & Willerslev, E. (2007) The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PloS one*. 2 (2), e197.

Blanckenhorn, W.U., Rohner, P.T., Bernasconi, M.V., Haugstetter, J. & Buser, A. (2016) Is qualitative and quantitative metabarcoding of dung fauna biodiversity feasible? *Environmental toxicology and chemistry / SETAC*. 35 (8), 1970–1977.

Bohan, D.A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A.J. & Woodward, G. (2017) Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks. *Trends in ecology & evolution*.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W. & de Bruyn, M. (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in ecology & evolution*. 29 (6), 358–367.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P. & Coissac, E. (2016) obitools: a unix-inspired software package for DNA metabarcoding. *Molecular ecology resources*. 16 (1), 176–182.

Brevik, E.C. (2005) 'A brief history of soil science', in W. Verheye (ed.) *Global Sustainable Development, Theme 1.05: Land Use and Cover,*. [Online]. Oxford, UK: EOLSS Publishers. p. http://www.eolss.net.

Bringloe, T.T., Adamowicz, S.J., Harvey, V.F.I., Jackson, J.K. & Cottenie, K. (2016) Detecting signatures of competition from observational data: a novel approach combining DNA barcoding, diversity partitioning, and checkerboards at small spatial scales. *Freshwater Biology*. 61646–657.

Brosi, B., Bell, K., Burgess, K. & Chu, H.-I. (2015) 'DNA metabarcoding of pollen: progress, technical limitations, and future directions', in *GENOME*. [Online]. 2015 Canadian Science Publishing, NRC Research Press 65 Auriga Dr, Suite 203, Ottawa, On K2E 7W6, Canada. pp. 201–201.

Brown, E.A., Chain, F.J.J., Zhan, A., MacIsaac, H.J. & Cristescu, M.E. (2016) Early detection of aquatic invaders using metabarcoding reveals a high number of non-indigenous species in Canadian ports. *Diversity & distributions*. 22 (10), 1045–1059.

Bucklin, A., Lindeque, P.K., Rodriguez-Ezpeleta, N., Albaina, A. & Lehtiniemi, M. (2016) Metabarcoding of marine zooplankton: prospects, progress and pitfalls. *Journal of plankton research*. 38 (3), 393–400.

Burns, R.G., DeForest, J.L., Marxsen, J., Sinsabaugh, R.L., Stromberger, M.E., Wallenstein, M.D., Weintraub, M.N. & Zoppini, A. (2013) Soil enzymes in a changing environment: Current knowledge and future directions. *Soil biology & biochemistry*. 58216–234.

Coissac, E., Riaz, T. & Puillandre, N. (2012) Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular ecology*. 21 (8), 1834–1847.

Cowart, D.A., Pinheiro, M., Mouchel, O., Maguer, M., Grall, J., Mine, J. & Arnaud-Haond, S. (2015) Metabarcoding Is Powerful yet Still Blind: A Comparative Analysis of Morphological and Molecular Surveys of Seagrass Communities. *PloS one*. 10 (2), 26.

Creamer, C.A., Filley, T.R., Boutton, T.W. & Rowe, H.I. (2016) Grassland to woodland transitions: Dynamic response of microbial community structure and carbon use patterns. *Journal of Geophysical Research: Biogeosciences*. 121 (6), 2016JG003347.

Creamer, R.E., Hannula, S.E., Van Leeuwen, J.P., Stone, D., Rutgers, M., Schmelz, R.M., Ruiter, P.C. de, Hendriksen, N.B., Bolger, T., Bouffaud, M.L., Buee, M., Carvalho, F., Costa, D., Dirilgen, T., Francisco, R., Griffiths, B.S., Griffiths, R., Martin, F., Silva, P.M. da, *et al.* (2016) Ecological network analysis reveals the inter-connection between soil biodiversity and ecosystem function as affected by land use across Europe. *Applied soil ecology: a section of Agriculture, Ecosystems & Environment*. 97112–124.

Deagle, B.E., Thomas, A.C., McInnes, J.C., Clarke, L.J., Vesterinen, E.J., Clare, E.L., Kartzinel, T.R. & Eveson, J.P. (2019) Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? Molecular ecology. 28 (2), 391–406.

DEFRA (2018) *A Green Future: Our 25 Year Plan to Improve the Environment*.

Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F. & Zhou, J. (2012) Molecular ecological network analyses. *BMC bioinformatics*. 13113.

Derocles, S.A.P., Bohan, D.A., Dumbrell, A.J., Kitson, J.J.N., Massol, F., Pauvert, C., Plantegenest, M., Vacher, C. & Evans, D.M. (2018) 'Chapter One - Biomonitoring for the 21st Century: Integrating Next-Generation Sequencing Into Ecological Network Analysis', in David A. Bohan, Alex J. Dumbrell, Guy Woodward, & Michelle Jackson (eds.) *Advances in Ecological Research*. [Online]. Academic Press. pp. 1–62.

Dominati, E., Patterson, M. & Mackay, A. (2010) A framework for classifying and quantifying the natural capital and ecosystem services of soils. *Ecological economics: the journal of the International Society for Ecological Economics*. 69 (9), 1858–1868.

Douglas, G.M., Maffei, V.J., Zaneveld, J.R., Yurgel, S.N., Brown, J.R., Taylor, C.M., Huttenhower, C. & Langille, M.G.I. (2020) PICRUSt2 for prediction of metagenome functions. Nature biotechnology. 38 (6), 685–688.

Dupont, Y.L., Padrón, B., Olesen, J.M. & Petanidou, T. (2009) Spatio-temporal variation in the structure of pollination networks. *Oikos* . 118 (8), 1261–1269.

Elbrecht, V. & Leese, F. (2015) Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass--Sequence Relationships with an Innovative Metabarcoding Protocol. *PloS one*. 10 (7), e0130324.

Ettema, C. (2002) Spatial soil ecology. *Trends in ecology & evolution*. 17 (4), 177–183.

Evans, D.M., Kitson, J., Lunt, D.H. & Straw, N.A. (2016) Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Ecology*.

Farine, D.R. & Whitehead, H. (2015) Constructing, conducting and interpreting animal social network analysis. *The Journal of animal ecology*. 84 (5), 1144–1163.

Field, K.J., Daniell, T., Johnson, D. & Helgason, T. (2020) Mycorrhizas for a changing world: Sustainability, conservation, and society. PLANTS, PEOPLE, PLANET. 2 (2), 98–103.

Filippelli, G.M. (2008) The Global Phosphorus Cycle: Past, Present, and Future. *Elements* . 4(2) (89-95).

Filip, Z. (2002) International approach to assessing soil quality by ecologically-related biological parameters. *Agriculture, ecosystems & environment*. 88 (2), 169–174.

Fitter, A.H., Gilligan, C.A., Hollingworth, K., Kleczkowski, A., Twyman, R.M., Pitchford, J.W. & The members of the NERC Soil Biodiversity Programme (2005) Biodiversity and ecosystem function in soil. *Functional ecology*. 19 (3), 369–377.

Floyd, R., Abebe, E., Papert, A. & Blaxter, M. (2002) Molecular barcodes for soil nematode identification. *Molecular ecology*. 11 (4), 839–850.

Foulon, J., Appelini, C., Durand, A., Valot, B., Girardclos, O., Blaudez, D. & Chalot, M. (2016) Environmental metabarcoding reveals contrasting microbial communities at two poplar phytomanagement sites. *The Science of the total environment*. 5711230–1240.

Fowler, D., Coyle, M., Skiba, U., Sutton, M.A., Cape, J.N., Reis, S., Sheppard, L.J., Jenkins, A., Grizzetti, B., Galloway, J.N., Vitousek, P., Leach, A., Bouwman, A.F., Butterbach-Bahl, K., Dentener, F., Stevenson, D., Amann, M. & Voss, M. (2013) The global nitrogen cycle in the twenty-first century. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 368 (1621), 20130164.

Frostegård, Å., Tunlid, A. & Bååth, E. (2011) Use and misuse of PLFA measurements in soils. *Soil biology & biochemistry*. 43 (8), 1621–1625.

Gessner, M.O., Swan, C.M., Dang, C.K., McKie, B.G., Bardgett, R.D., Wall, D.H. & Hättenschwiler, S. (2010) Diversity meets decomposition. *Trends in ecology & evolution*. 25 (6), 372–380.

Giller, P.S. (1996) The diversity of soil communities, the 'poor man's tropical rainforest'? *Biodiversity and conservation*. 5 (2), 135–168.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M. & Toulmin, C. (2010) Food security: the challenge of feeding 9 billion people. *Science*. 327 (5967), 812–818.

Griffiths, B.S., Ritz, K., Bardgett, R.D., Cook, R., Christensen, S., Ekelund, F., Sørensen, S.J., Bååth, E., Bloem, J., De Ruiter, P.C., Dolfing, J. & Nicolardot, B. (2000) Ecosystem response of pasture soil communities to fumigation-induced microbial diversity reductions: an examination of the biodiversity–ecosystem function relationship. *Oikos* . 90 (2), 279–294.

Guardiola, M., Uriz, M.J., Taberlet, P., Coissac, E., Wangensteen, O.S. & Turon, X. (2015) Deep-Sea, Deep-Sequencing: Metabarcoding Extracellular DNA from Sediments of Marine Canyons. *PloS one*. 10 (10), e0139633.

Guimarães, P.R. (2020) The Structure of Ecological Networks Across Levels of Organization. *Annual review of ecology, evolution, and systematics*. 51 (1), 433–460.

Handelsman, J. (2004) Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and molecular biology reviews: MMBR*. 68 (4), 669–685.

Hanfling, B., Handley, L.L., Read, D.S., Hahn, C., Li, J.L., Nichols, P., Blackman, R.C., Oliver, A. & Winfield, I.J. (2016) Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular ecology*. 25 (13), 3101–3119.

Hättenschwiler, S., Tiunov, A.V. & Scheu, S. (2005) Biodiversity and Litter Decomposition in Terrestrial Ecosystems. *Annual review of ecology, evolution, and systematics*. 36 (1), 191–218.

Hawkins, J., de Vere, N., Griffith, A., Ford, C.R., Allainguillaume, J., Hegarty, M.J., Baillie, L. & Adams-Groom, B. (2015) Using DNA Metabarcoding to Identify the Floral Composition of Honey: A New Tool for Investigating Honey Bee Foraging Preferences. *PloS one*. 10 (8), 20.

Hebert, P.D.N., Cywinska, A., Ball, S.L. & deWaard, J.R. (2003) Biological identifications through DNA barcodes. *Proceedings. Biological sciences / The Royal Society*. 270 (1512), 313–321.

van der Heijden, M.G.A., Bardgett, R.D. & van Straalen, N.M. (2008) The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology letters*. 11 (3), 296–310.

van der Heijden, M.G.A., Klironomos, J.N., Ursic, M., Moutoglis, P., Streitwolf-Engel, R., Boller, T., Wiemken, A. & Sanders, I.R. (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature*. 396 (6706), 69–72.

Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A.J., Vandermeer, J. & Wardle, D.A. (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological monographs*. 75 (1), 3–35.

HS2 Ltd (2020) *High Speed Two Phase 2a: Ancient Woodland Strategy*.

Ji, Y., Ashton, L., Pedley, S.M., Edwards, D.P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P.M., Woodcock, P., Edwards, F.A., Larsen, T.H., Hsu, W.W., Benedick, S., Hamer, K.C., Wilcove, D.S., Bruce, C., Wang, X., Levi, T., Lott, M., *et al.* (2013a) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology letters*. 16 (10), 1245–1257.

Ji, Y., Ashton, L., Pedley, S.M., Edwards, D.P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P.M., Woodcock, P., Edwards, F.A., Larsen, T.H., Hsu, W.W., Benedick, S., Hamer, K.C., Wemheuer, F., Taylor, J.A., Daniel, R., Johnston, E., Meinicke, P., Thomas, T. & Wemheuer, B. (2018) Tax4Fun2: a R-based tool for the rapid prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene marker gene sequences. bioRxiv p.490037.

Wilcove, D.S., Bruce, C., Wang, X., Levi, T., Lott, M., *et al.* (2013b) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology letters*. 16 (10), 1245–1257.

Jordán, F., Okey, T.A., Bauer, B. & Libralato, S. (2008) Identifying important species: Linking structure and function in ecological networks. *Ecological modelling*. 216 (1), 75–80.

Kaiser-Bunbury, C.N., Mougal, J., Whittington, A.E., Valentin, T., Gabriel, R., Olesen, J.M. & Blüthgen, N. (2017) Ecosystem restoration strengthens pollination network resilience and function. *Nature*.

Kemp, J.E., Evans, D.M., Augustyn, W.J. & Ellis, A.G. (2016) Invariant antagonistic network structure despite high spatial and temporal turnover of species and their interactions in a biodiversity hotspot. *Ecography*.

Kitson, J.J.N., Hahn, C., Sands, R.J., Straw, N.A., Evans, D.M. & Lunt, D.H. (2018) Detecting host-parasitoid interactions in an invasive Lepidopteran using nested tagging DNA metabarcoding. *Molecular ecology*.

Lal, R. (2004) Soil carbon sequestration impacts on global climate change and food security. *Science*. 304 (5677), 1623–1627.

Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G. & Taylor, M.I. (2019) How quantitative is metabarcoding: A meta-analytical approach. Molecular ecology. 28 (2), 420–430.

Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R., Beiko, R.G. & Huttenhower, C. (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature biotechnology*. 31 (9), 814–821.

Liu, M., Clarke, L.J., Baker, S.C., Jordan, G.J. & Burridge, C.P. (2020) A practical guide to DNA metabarcoding for entomological ecologists. Ecological entomology. 45 (3), 373–385.

Mackay, J.E., Cunningham, S.C. & Cavagnaro, T.R. (2016) Riparian reforestation: are there changes in soil carbon and soil microbial communities? *The Science of the total environment*. 566-567960–967.

Madden, M.J.L., Young, R.G., Brown, J.W. & Miller, S.E. (2019) Using DNA barcoding to improve invasive pest identification at US ports-of-entry. PLoS.

McBratney, A., Field, D.J. & Koch, A. (2014) The dimensions of soil security. *Geoderma*. 213203–213.

Montoya, J.M., Pimm, S.L. & Solé, R.V. (2006) Ecological networks and their fragility. *Nature*. 442 (7100), 259–264.

Morriën, E., Hannula, S.E., Snoek, L.B., Helmsing, N.R., Zweers, H., de Hollander, M., Soto, R.L., Bouffaud, M.-L., Buée, M., Dimmers, W., Duyts, H., Geisen, S., Girlanda, M., Griffiths, R.I., Jørgensen, H.-B., Jensen, J., Plassart, P., Redecker, D., Schmelz, R.M., *et al.* (2017) Soil networks become more connected and take up more carbon as nature restoration progresses. *Nature communications*. 814349.

Nichols, R.V., Akesson, M. & Kjellander, P. (2016) Diet Assessment Based on Rumen Contents: A Comparison between DNA Metabarcoding and Macroscopy. *PloS one*. 11 (6), 16.

Nielsen, U.N., Ayres, E., Wall, D.H. & Bardgett, R.D. (2011) Soil biodiversity and carbon cycling: a review and synthesis of studies examining diversity–function relationships. *European journal of soil science*. 62 (1), 105–116.

Oikonomopoulos, S., Wang, Y.C., Djambazian, H., Badescu, D. & Ragoussis, J. (2016) Benchmarking of the Oxford Nanopore MinION sequencing for quantitative and qualitative assessment of cDNA populations. *Scientific reports*. 631602.

Pace, N.R., Stahl, D.A., Lane, D.J. & Olsen, G.J. (1986) 'The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences', in *Advances in Microbial Ecology*. Advances in Microbial Ecology. [Online]. Springer, Boston, MA. pp. 1–55.

Pauvert, C., Buée, M., Laval, V., Edel-Hermann, V., Fauchery, L., Gautier, A., Lesur, I., Vallance, J. & Vacher, C. (2019) Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal ecology*. 4123–33.

Pauvert, C., Vallance, J., Delière, L., Bohan, D.A., Buée, M. & Vacher, C. (2020) Inferred microbial networks lack replicability: consequences for next-generation biomonitoring. bioRxiv p.642199.

Pawlowski, J., Esling, P., Lejzerowicz, F., Cedhagen, T. & Wilding, T.A. (2014) Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. *Molecular ecology resources*. 14 (6), 1129–1140.

Pawlowski, J., Lejzerowicz, F., Apotheloz-Perret-Gentil, L., Visco, J. & Esling, P. (2016) Protist metabarcoding and environmental biomonitoring: Time for change. *European journal of protistology*. 5512–25.

Philippot, L., Spor, A., Hénault, C., Bru, D., Bizouard, F., Jones, C.M., Sarr, A. & Maron, P.-A. (2013) Loss in microbial diversity affects nitrogen cycling in soil. *The ISME journal*. 7 (8), 1609–1619.

Pilosof, S., Porter, M.A., Pascual, M. & Kéfi, S. (2017) The multilayer nature of ecological networks. *Nature ecology & evolution*. 1 (4), 101.

Piñol, J., Mir, G., Gomez-Polo, P. & Agustí, N. (2015) Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular ecology resources*. 15 (4), 819–830.

Pocock, M.J.O., Evans, D.M. & Memmott, J. (2012) The robustness and restoration of a network of ecological networks. *Science*. 335 (6071), 973–977.

Pompanon, F. & Samadi, S. (2015) Next generation sequencing for characterizing biodiversity: promises and challenges. *Genetica*. 143 (2), 133–138.

Prentice, I.C., Farquhar, G.D., Fasham, M.J.R., Goulden, M.L., Heimann, M., Jaramillo, V.J., Kheshgi, H.S., LeQuéré, C., Scholes, R.J. & Wallace, D.W.R. (2001) '*The carbon cycle and atmospheric carbon dioxide*', in [Online]. Cambridge University Press.

Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005) Network thinking in ecology and evolution. *Trends in ecology & evolution*. 20 (6), 345–353.

Rackham, O. (2008) Ancient woodlands: modern threats. *The New phytologist*. 180 (3), 571–586.

Reuter, J.A., Spacek, D.V. & Snyder, M.P. (2015) High-throughput sequencing technologies. *Molecular cell*. 58 (4), 586–597.

Reva, O.N., Zaets, I.E., Ovcharenko, L.P., Kukharenko, O.E., Shpylova, S.P., Podolich, O.V., de Vera, J.P. & Kozyrovska, N.O. (2015) Metabarcoding of the kombucha microbial community grown in different microenvironments. *AMB Express*. 58.

Richardson, R.T., Lin, C.H., Quijia, J.O., Riusech, N.S., Goodell, K. & Johnson, R.M. (2015) Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach. *Applications in plant sciences*. 3 (11), 9.

Roosevelt, F.D. (1937) *Letter to all state governors on a uniform soil conservation law*. [Online]

Rosenfeld, J.S. (2002) Functional redundancy in ecology and conservation. *Oikos* . 98 (1), 156–162.

Ruppert, K.M., Kline, R.J. & Rahman, M.S. (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation. 17e00547.

Sapp, M., Lewis, E., Moss, S., Barrett, B., Kirk, S., Elphinstone, J. & Denman, S. (2016) Metabarcoding of Bacteria Associated with the Acute Oak Decline Syndrome in England. *Forests, Trees and Livelihoods*. 7 (5), 95.

Schröter, D., Cramer, W., Leemans, R., Prentice, I.C., Araújo, M.B., Arnell, N.W., Bondeau, A., Bugmann, H., Carter, T.R., Gracia, C.A., de la Vega-Leinert, A.C., Erhard, M., Ewert, F., Glendining, M., House, J.I., Kankaanpää, S., Klein, R.J.T., Lavorel, S., Lindner, M., *et al.* (2005) Ecosystem service supply and vulnerability to global change in Europe. *Science*. 310 (5752), 1333–1337.

Scotti, M., Podani, J. & Jordán, F. (2007) Weighting, scale dependence and indirect effects in ecological networks: A comparative study. *Ecological Complexity*. 4 (3), 148–159.

Sharma, R., Polkade, A.V. & Shouche, Y.S. (2015) 'Species concept' in microbial taxonomy and systematics. *Current science*. 108 (10), 1804–1814.

Shi, Y., Delgado-Baquerizo, M., Li, Y., Yang, Y., Zhu, Y.-G., Peñuelas, J. & Chu, H. (2020) Abundance of kinless hubs within soil microbial networks are associated with high functional potential in agricultural ecosystems. *Environment international*. 142105869.

Shokralla, S., Spall, J.L., Gibson, J.F. & Hajibabaei, M. (2012) Next-generation sequencing technologies for environmental DNA research. *Molecular ecology*. 21 (8), 1794–1805.

Strickland, M.S., Lauber, C., Fierer, N. & Bradford, M.A. (2009) Testing the functional significance of microbial community composition. *Ecology*. 90 (2), 441–451.

Suweis, S., Simini, F., Banavar, J.R. & Maritan, A. (2013) Emergence of structural and dynamical properties of ecological mutualistic networks. *Nature*. 500 (7463), 449–452.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. & Willerslev, E. (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular ecology*. 21 (8), 2045–2050.

Thébault, E. & Fontaine, C. (2010) Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*. 329 (5993), 853–856.

Thompson, R.M., Brose, U., Dunne, J.A., Hall, R.O., Jr, Hladyz, S., Kitching, R.L., Martinez, N.D., Rantala, H., Romanuk, T.N., Stouffer, D.B. & Tylianakis, J.M. (2012) Food webs: reconciling the structure and function of biodiversity. *Trends in ecology & evolution*. 27 (12), 689–697.

Tiedje, J.M., Asuming-Brempong, S., Nüsslein, K., Marsh, T.L. & Flynn, S.J. (1999) Opening the black box of soil microbial diversity. *Applied soil ecology: a section of Agriculture, Ecosystems & Environment*. 13 (2), 109–122.

Tilman, D., Reich, P.B., Knops, J., Wedin, D., Mielke, T. & Lehman, C. (2001) Diversity and productivity in a long-term grassland experiment. *Science*. 294 (5543), 843–845.

Toju, H., Guimarães, P.R., Olesen, J.M. & Thompson, J.N. (2014) Assembly of complex plant-fungus networks. *Nature communications*. 55273.

Toju, H., Yamamichi, M., Guimarães, P.R., Jr, Olesen, J.M., Mougi, A., Yoshida, T. & Thompson, J.N. (2017) Species-rich networks and eco-evolutionary synthesis at the metacommunity level. *Nature Ecology & Evolution*. 10024.

Tylianakis, J.M., Laliberte, E., Nielsen, A. & Bascompte, J. (2010) Conservation of species interaction networks. *Biological conservation*. 143 (10), 2270–2279.

de Vries, F.T., Thébault, E., Liiri, M., Birkhofer, K., Tsiafouli, M.A., Bjørnlund, L., Bracht Jørgensen, H., Brady, M.V., Christensen, S., de Ruiter, P.C., d'Hertefeldt, T., Frouz, J., Hedlund, K., Hemerik, L., Hol, W.H.G., Hotes, S., Mortimer, S.R., Setälä, H., Sgardelis, S.P., *et al.* (2013) Soil food web properties explain ecosystem services across European land use systems. *Proceedings of the National Academy of Sciences of the United States of America*. 110 (35), 14296–14301.

Wagg, C., Bender, S.F., Widmer, F. & van der Heijden, M.G.A. (2014) Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (14), 5266–5270.

van der Heijden, M.G.A., Martin, F.M., Selosse, M.-A. & Sanders, I.R. (2015) Mycorrhizal ecology and evolution: the past, the present, and the future. The New phytologist. 205 (4), 1406–1423.

Van Der Heijden, M.G.A., Streitwolf-Engel, R., Riedl, R., Siegrist, S., Neudecker, A., Ineichen, K., Boller, T., Wiemken, A. & Sanders, I.R. (2006) The mycorrhizal contribution to plant productivity, plant nutrition and soil structure in experimental grassland. The New phytologist. 172 (4), 739–752.

van der Wal, A., Geydan, T.D., Kuyper, T.W. & de Boer, W. (2013) A thready affair: linking fungal diversity and community dynamics to terrestrial decomposition processes. *FEMS microbiology reviews*. 37 (4), 477–494.

Walker, B.H. (1992) Biodiversity and Ecological Redundancy. *Conservation biology: the journal of the Society for Conservation Biology*. 6 (1), 18–23.

Warkentin, B.P. (2008) 'Soil Structure: A History from Tilth to Habitat', in *Advances in Agronomy*. [Online]. Academic Press. pp. 239–272.

Warkentin, B.P. (1999) The return of the 'other' soil scientists. *Canadian journal of soil science*. 79 (1), 1–4.

Winogradsky, S. (1949) *Microbiologie du sol ; problemes et methodes*. Paris: Masson.

Yoon, T.H., Kang, H.E., Kang, C.K., Lee, S.H., Ahn, D.H., Park, H. & Kim, H.W. (2016) Development of a cost-effective metabarcoding strategy for analysis of the marine phytoplankton community. *PeerJ*. 420.

Yu, D.W., Ji, Y.Q., Emerson, B.C., Wang, X.Y., Ye, C.X., Yang, C.Y. & Ding, Z.L. (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in ecology and evolution / British Ecological Society*. 3 (4), 613–623.

Zaiko, A., Samuiloviene, A., Ardura, A. & Garcia-Vazquez, E. (2015) Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. *Marine pollution bulletin*. 100 (1), 53–59.

Zaiko, A., Schimanski, K., Pochon, X., Hopkins, G.A., Goldstien, S., Floerl, O. & Wood, S.A. (2016) Metabarcoding improves detection of eukaryotes from early biofouling communities: implications for pest monitoring and pathway management. *Biofouling*. 32 (6), 671–684.

Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q. & Zhi, X. (2010) Functional molecular ecological networks. *mBio*. 1 (4), 10.

Zimmermann, J., Glockner, G., Jahn, R., Enke, N. & Gemeinholzer, B. (2015) Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Molecular ecology resources*. 15 (3), 526–542.

# Chapter 2. Soil Abiotic Responses to Woodland Creation: A Comparative Study Across a Long-Perspective Natural Experiment

## 2.1 Abstract

Woodland restoration on agricultural land has become an extremely popular strategy for habitat restoration and carbon sequestration, both internationally and within the UK. Understanding the impacts of afforestation projects on a range of ecological processes is of great interest but can only be done in the context of the pedological changes that occur in parallel. Much effort has been spent in monitoring the success of these schemes, charting the changes in soil properties in recent plantings over the last half-century. However, few studies have combined observations from agricultural soils, young woodlands, and with those of mature and ancient semi-natural woodland. Here, I use a large-scale woodland restoration experiment, set up to monitor agricultural and woodland ecosystems in a homogeneous landscape, to examine differences in soil properties. By comparing 210 soil cores collected from 21 sites, I found higher carbon concentration, nitrogen concentration, and C:N ratios in afforested and woodland sites when compared to arable counterparts, while considering spatial autocorrelation. Pasture soils differ from woodland sites less than arable ones, though exhibit very high ammonium levels and lower nitrate concentrations than arable farms. Bulk density is lower significantly in recently reforested sites (<50 years), indicating a legacy of the soil disturbance caused by planting persists for half a century, but disappears in older woodlands. Soil pH differed little across the experimental sites, making them well suited to biotic comparisons, in particular microbial community composition.

## 2.2 Introduction

### *2.2.1 Woodland in the UK*

The UK is in its greatest period of afforestation since the Middle Ages. Broady, woodland cover in Anglo-Saxon and Norse Britain likely declined due to agricultural expansion, but was not dramatically reshaped (Rackham, 1986). Following the Norman conquest and throughout the early Middle Ages woodland cover rapidly declined from an estimated 15% 1086, drawn from the extensive documentation of land cover in the Domesday Book, to around 10% by 1350 – averaging a destruction of around seven hectares a day (Rackham, 1986). The black death,

which arrived in Britain in 1348, curtailed woodland decline and much of the woodland found in 14th century Britain (c. 80% of woods) survived the next three centuries (Cahalan *et al.*, 2011). After this time, woodland cover became more closely tied to the varying demands for charcoal, bark, and timber following the industrial revolution (Rackham, 1986). The boom-and-bust cycles of these industries caused periods of woodland development and neglect, with agricultural booms working antagonistically over that same period and leading to further reduction of woodland. By the early 20th century UK woodland covered only around 5.1% of the country (c. 600,000 ha), its lowest recorded extent (Cahalan *et al.*, 2011). Following the end of the Second World War, successive governments have guided the silviculture industry to meet perceived national needs. Immediately following the wars, the UK Government policy was changed to encourage management for productivity, ushering in a burst of non-native, conifer woodland creation (Cahalan *et al.*, 2011). The desire for quality wood products frequently encouraged the conversion of less productive ancient and mature semi-natural broadleaf woodland to largely non-native conifer plantations (Cahalan *et al.*, 2011), with the conifer forest area increasing from around 382,400 ha in 1947 (41% of the forest area of Great Britain) to around 917,300 ha in 1965 (67% of forest area). During this time, coppice and coppice with standards (uncoppiced trees) declined by 114,800ha (81%) and has never recovered (Mason, 2007).

By 1982 Sitka Spruce (*Picea sitchensis*) had spread to cover over 500,000ha of Britain, becoming the commonest British tree by a considerable margin (Mason, 2007). Sitka has remained Britain's most populous tree ever since (Forestry Commission, 2019). Overall, conifer cover in Great Britain had reached 1.32 million ha (67% forest area) (Mason, 2007). The mid-1980s introduced policy and tax that curbed the conversion of broadleaf woodland to conifer plantations and slowed the rate of woodland growth (Mason, 2007). The decreasing rate of woodland growth was accompanied by a shift towards the planting of broadleaf woodlands and increasing recognition of the social and biodiversity benefits of woods (Cahalan *et al.*, 2011). Between 1982 and 2000, conifer cover increased by only 58,500 ha in Great Britain, while over 320,000 ha of broadleaf woodland was planted. Despite changing attitudes in the forestry industry, Britain's oldest woodlands - ancient semi-natural woodlands (ASNW) - still face threats and have been steadily declining in national cover. Even in their relative state of expansion, woodlands cover only 13% of the UK ('Woodland Statistics - Forest

Research', 2019) compared to around 43% cover over the European continent as a whole (The Scottish Government, 2019).

Changes in UK politics may lead to dramatic changes to the British countryside, propelling forestry policy once again into the national focus (Tatchell-Evans, 2016). The current UK 25-year action plan (DEFRA, 2018) makes clear that increasing woodland cover is a priority. It emphasizes the protection of ancient woodland and tree planting, with particular focus on increasing the proportion of broadleaf woodland and the commercial hardwood timber capacity. Both Scotland and Wales's devolved governments' forestry strategies also place woodland creation and native woodland expansion and protection as a central aspect of future management, with targets to reach 21% woodland cover by 2032 and to plant 2000ha of woodland per year between 2020 and 2030, respectively (Welsh Government, 2018; The Scottish Government, 2019). Furthermore, calls for increased woodland planting (Chazdon & Brancalion, 2019) are drawing on a growing body of evidence supporting the wider socio-economic benefits of broadleaf woodlands. This includes services to health (Maas *et al.*, 2009); carbon sequestration (Cannell, 2003); and a variety of biodiversity-related benefits, such as wildlife corridors and habitat diversity that can increase the variety and size of natural populations; all produced by heterogeneous, highly-connected landscapes that include woodlands of all types (Burton *et al.*, 2018). Combined, woodland regeneration and reforestation are becoming an increasingly popular strategy (Chazdon, 2008) for sustainable stewarding of biodiversity, economic gains, climate mitigation, and human well-being, (Chazdon & Brancalion, 2019).

In the UK, widespread abandonment of historic forest management approaches, including coppicing, pollarding, and hardwood timber production in broadleaf woodlands across the UK in the 20th century has led to a noticeable "age-gap" in the structure of British woods and hedges (Cahalan *et al.*, 2011; Rackham, 2012). When current veteran trees die or regenerate to young growth, many locations will be missing the "next-generation" of veteran trees to provide the unique ecological benefits supported by them (Cahalan *et al.*, 2011), such as the niches they create for insects, fungi, bats, and birds. The changing structure of UK woodlands will likely have far-reaching socio-economic impacts on the UK and will exert a noticeable

influence over the biodiversity of the British Isles. Additionally, ASNWs in the UK are under continuous threat from land development (Rackham, 2008). The importance of ASNW for above-ground organisms is understood (Hermy & Verheyen, 2007), with around 20% classified as Sites of Special Scientific Interest (SSSIs) according to The Woodland Trust (HS2 Select Committee, 2016), compared to only 10% for broadleaf, mixed and yew woodlands generally (Natural England, 2008). However, less is known about the impact of long-term land-use history on below-ground woodland organisms. A strategy is needed to ensure the long-term viability of the UK treescape. Woodland creation will be a large part of the solution. Together, these issues highlight the need to better understand how the ecology of woodland creation and regeneration sites differs from that of mature secondary and ancient woodland sites.

### *2.2.2 Woodland soil research*

Previous work on woodland chronosequences has largely focused on comparisons of the above-ground biodiversity of several age categories of woodlands (Hermy & Verheyen, 2007), demonstrating the conservation value of these environments for protecting woodland specialists that are often dispersal limited (Kimberley *et al.*, 2013), such as the ancient woodland indicator species Dog's Mercury (*Mercurialis perennis*) which disperses via ants and vegetative propagules (Jefferson, 2008). A smaller section of the literature details the soil chemistry changes that occur as woodland environments transition through age structure stages, including changes to pH, carbon stocks, and nutrient pools. Of this, the focus is generally on planted woodlands and documents the first half-century or so of plantation growth. The soil changes studied are most often related to carbon stocks. Additionally, a great deal of this work is conducted in tropical and, to a lesser extent, boreal systems, or on coniferous plantations. Less work examines differences over long timescales, in semi-natural systems, or on a range of soil properties beyond carbon stocks. It is unclear how many of the conclusions gained from studies of coniferous woodlands, or woodlands from a huge geographic, climatic, and geological range, can be applied in temperate, broadleaf woodlands.

This concern is reinforced *in silico* as simulations indicate that the recovery trajectories of biotic and abiotic site characteristics should vary across biomes (Krause *et al.*, 2016). Indeed, much of the literature has collected information on soils as the context for their main

experiment or have combined data from many studies and summarised them to make a more complete review. There are challenges to drawing conclusions from these studies that have not been systematically designed to answer the question of how soils age over periods of afforestation and woodland growth. In this chapter, I present the findings broadly, with the caveat that, for all general trends, exceptions exist that demonstrate the primacy of local geography over broad patterns. The majority of the literature makes use of space-for-time experimental designs, which has the limitation of not being able to control for differences in the starting soil characteristics of sites that are being compared, as well as differences in the climate experienced by woodlands at different growth stages (although see (Zhou *et al.*, 2006) for an exception).

A general understanding of carbon and nutrient cycling in soil-vegetation systems would lead one to expect certain important inputs, outputs and cycling processes to importantly vary between land-use types. In the British lowland landscapes that are the focus of this work, land-uses sit on a spectrum of disturbance. Arable sites experience the highest disturbance; followed by pasture sites, where there tends to be less vegetation management; woodlands experience the least disturbance of these land-use types. Disturbance may decrease with woodland age, although the difference between mature and ASNW disturbance levels may depend more on management than age. The regular removal of crops from arable land and its tilling encourages the loss of carbon and nitrogen from the system (Weil & Brady, 2017). The soil carbon and nitrogen pools are ultimately fed from atmospheric carbon and nitrogen pools via biological processes, though they are supplemented by additional input from precipitation. They are depleted when leaching occurs or when vegetation is removed from the system. In less disturbed environments biological processes such as vegetation cycling of nutrients and earthworm activity retain nutrients in the system, preventing their loss to leaching. Generally, plants move minerals and nutrients up the soil profile into living materials and then back to the upper horizons of the soil. However, the factors that govern the rates of these processes exist in positive and negative feedback loops that are difficult to disentangle (Trudgill, 1988). For example, soil pH will depend on the parent material present, but the capacity of the soil to act as a pH buffer will be influenced by processes like organic input, decomposition, chelation, leaching and weathering of parent material that feedback into each other. Over time acid rain and deposition processes will act to acidify soils everywhere, but on

forested sites, the higher turnover of organic material and its decomposition will tend to acidify the soils over time. Despite this, site characteristics and climate may end up being the dominant factors in setting site pH at a given location. Management such as the liming of soils is common on arable soils to prevent acidification and leaching. The bulk density of soils also decreases in soils with more organic matter, which provide heterogeneity to soil structure (Weil & Brady, 2017) and correlate with the presence of soil organisms which may aerate the soil through their activity. Differences between arable, pasture and woodland soils may be predicted based upon these known soil processes. However, changes that occur as woodlands age are less clear and examples from the literature are needed to illustrate how these processes work in reality.

The literature indicates that soil carbon generally increases rapidly after afforestation for the first 30 - 60 years (C. A. Creamer *et al.*, 2016; Mackay *et al.*, 2016; Kurganova *et al.*, 2018; Kalinina *et al.*, 2019; Vindušková *et al.*, 2019). Afterwards, the accumulation of carbon generally slows or levels off in woodlands aged 50 - 150 years (Jangid *et al.*, 2011; Susyan *et al.*, 2011; C. A. Creamer *et al.*, 2016; Kurganova *et al.*, 2018; Kalinina *et al.*, 2019; Vindušková *et al.*, 2019). However, trends for broadleaf and conifer plantings may differ, Ražauskaitė *et al.* (2020) concluded that carbon stocks of conifer afforested soils were lower after 43 years than in younger plantations or in undisturbed forest, perhaps due to litter-soil dynamics. Generally, similar trends are observed for soil nitrogen, although these are more variable. Some studies indicate decreasing soil nitrogen in the first 120 years of woodland growth (Saiz *et al.*, 2006; Trap *et al.*, 2013; Deng & Shangguan, 2017; Wu *et al.*, 2020), but evidence of increasing total nitrogen is more common (e.g. (Jangid *et al.*, 2011; C. A. Creamer *et al.*, 2016; Mackay *et al.*, 2016; Kurganova *et al.*, 2018). Less of this work examines how pH changes in broadleaf woodlands over this establishment period; that which I have found indicates that it differs little between compared woodlands of different ages (Jangid *et al.*, 2011) although in coniferous woodlands soil pH usually acidifies. The literature on ancient semi-natural woodland is universally hampered by lack of knowledge of long-term land-use history. Ancient woodland is often defined using the oldest available maps for the landscape. In Europe generally, studies can often trace back land use for 200-250 years but have limited evidence of land use prior to that, whereas ancient woodlands in England and Wales are required to exist on maps dated to c.1600, and to c. 1750 in Scotland. As I do not know how long these

soils have been covered by woodland, I cannot easily study what changes may be occurring in soils that are over several hundred years old. When studies contrast established, secondary woodlands known to be younger than 250 years old, to ancient woodlands with comparable tree cover, differences in soil characteristics vary greatly in direction and strength with, for example, Nitsch et al (2018) finding increases in soil SOC and N, Susyan et al. (2018) finding increased SOC but Fitchner et al (2014) finding decreased total C and total N. This trend continues through other literature, although Nitsch et al. add that most work considers only the upper layers of the soil (as I will in this work). Often measures show lower soil carbon and nitrogen in ASNW, although no change or a higher carbon and nitrogen is also observed. Changes to, or stability of, pH occurs is highly variable between studies and often not sampled systematically of analysed directly, some have found decreasing pH older woodland and forest (Kurganova et al. 2018), others increasing (Nitsch et al. 2018), and yet others conflicting result (Susyan et al. 2011). Carbon to nitrogen ratio very often was higher in recently afforested and mature woodlands (Mackay *et al.*, 2016; Deng & Shangguan, 2017) and may be comparable between mature secondary and ancient woodlands (Nitsch *et al.*, 2018) or higher in ancient woods (Fichtner *et al.*, 2014). The input of ammonium from animal waste may be higher in natural and pasture landscapes than in arable landscapes. However, in warm, aerated soils ammonium quickly converts to nitrates in days to weeks (Weil & Brady, 2017) and is very rapidly lost to plant uptake and denitrification. One might expect nitrates to be high in systems with a lot of added fertiliser, whether that fertiliser consists of mostly ammonium or nitrate when added and regardless of its origin. Without more information about site history, the timings of fertiliser additions, or the timings of crop removal it is difficult to predict which land-uses will have the greatest soil ammonium and nitrate. Unfortunately, this data is unavailable.

The Woodland creation and Ecological Networks (WrEN) project was established in 2013 to study the effect of habitat fragmentation on biodiversity in UK woodlands on a large-scale natural experiment. Working at this scale benefits from a natural experiment approach, as investigations that manipulate variables at a landscape scale are challenging in environments that have a long history of management. Natural experiments remove the obstacles of time, resources, and work that are needed to design and affect landscape-level experiments that can be prohibitive to researchers. The project selected 106 woodlands, between 0.5 and 32ha in size, 10 and 160 years old, and isolated from other woodlands by between 7 and 1573 m

(Watts *et al.*, 2016). Multiple ecological surveys have been conducted within WrEN project sites to study how site and landscape-scale differences affect biodiversity, including work demonstrating how a hierarchy of spatially influenced decisions made by birds affect their habitat choice (Whytock, Fuentes-Montemayor, Watts, Macgregor, *et al.*, 2018), how local, not landscape, quality affects Diptera habitat choice (Fuller *et al.*, 2018), and how bird community composition is influenced by habitat continuity (Whytock, Fuentes-Montemayor, Watts, Barbosa De Andrade, *et al.*, 2018). Recognising the importance of the below-ground components in these ecosystems, the WrEN project recently presented results from a focused study of earthworm diversity and soil factors from a subset of 21 WrEN sites in the English midlands. The study documented how soil characteristics and earthworm biodiversity varies across sites at different stages of woodland restoration, identifying improved earthworm diversity and soil carbon stocks in older woodlands (Ashwood *et al.*, 2019). *Ashwood et al.* demonstrated that large changes to organic carbon and earthworm diversity occur over a disturbance gradient from arable farmland to pasture farmland, and then to woodlands of increasing age. Soil organic carbon stocks significantly differed between arable and ancient semi-natural woodlands, with intermediate carbon stocks in pasture sites and younger woodlands. They also identified significantly higher C:N ratios in ASNW sites and younger woodlands, when compared to all agricultural sites pooled together. Ashwood *et al* completed an important analysis of the site scale data for soil physical and chemical properties. I build on this data by analysing the data at the soil core scale, incorporating mixed effect GLMMs to avoid issues relating to pseudo-replication. This allows us to examine the differences between land-use types without losing data. I reach additional conclusions about the data and analyse soil characteristics not explored in the initial paper with significant conclusions. To my knowledge, the work of *Ashwood et al.* is the first study to systematically sample and analyse soils from replicated pasture and arable farmland; and young, mature and ancient woodlands simultaneously; though see (Zhu *et al.*, 2010; Susyan *et al.*, 2011) for studies of scrub and woodlands with a broad range of ages and comparison with arable farms, with less replication. However, the soil microbial diversity of the sites is yet to be studied.

### *2.2.3 Hypotheses*

Here, I aim to assess the short-term and long-term effects of woodland restoration on arable and pasture farmland soils in the UK midlands. I will determine which soil properties

consistently change across land-use types with a view to using these changes as potential explanatory variables of soil microbial diversity change in the same samples. Ultimately this will improve understanding of how the abiotic and biotic components of soils vary across a restoration gradient and identify potential interactions between these components that may confound or reinforce relationships between diversity and woodland age. The youngest woodlands in this study (50-60 years old) are older than many of the planted woodlands examined in the literature. Their soils may have already experienced their most rapid period of change and so I expect young woodland soils to be quite differentiated from their agricultural counterparts. Woodland soils older than 60 years may continue to change in the same direction as that of younger woodlands, or changes may slow or reverse direction. I predict that soil pH would be lower in young woodlands than in agricultural soils and decrease or remain level as they age. Soil carbon and nitrogen, and the carbon to nitrogen ratio should increase over the same period, along with the C:N ratio as organic inputs become more lignified. I hypothesised that ammonium concentrations may be higher in pasture soils due to the large input from livestock, whereas nitrate concentrations may be higher in arable soils resulting from fertiliser application. However, due to the short lifetimes of soil ammonium and nitrate, these concentrations may not differ significantly between land-use types. I predict that compaction will be higher in agricultural sites than woodland sites, leading to higher bulk densities.

## 2.3 Methods

### *2.3.1 Site selection and sampling*

For this study, a subsample of sites from the Woodland creation and Ecological Networks (WrEN) project was chosen. The site selection and sampling process are described in detail in previous work (Ashwood *et al.*, 2019). Briefly, Ashwood *et al.* (2019) selected 21 sites in the National Forest area of the UK midlands for a detailed study of the relationship between earthworm biodiversity and woodland creation. Selected sites (Table 2.1) consisted of pastoral and arable farmland, young and mature secondary woodlands, and Ancient Semi-Natural Woodlands (ASNW).

Table 2.1. Numbers and description of sites of each land use type included in this study, located in the English Midlands

| Land Use | Category | Land use continuity | n |
|---|---|---|---|
| Pasture | Agriculture | Unknown | 3 |

| | | | |
|---|---|---|---|
| Arable | Agriculture | Unknown | 4 |
| Young woodland | Woodland | 50 - 60 years | 4 |
| Mature Woodland | Woodland | 110 - 116 years | 3 |
| ASNW | Woodland | Over 400 years | 7 |

The 21 sites were situated in the agrarian lowlands of the midlands and were selected to form a chronosequence with minimal variation due to climate and geography (Watts *et al.*, 2016). Ashwood *et al.* (2019) implemented a systematic selection process; first using historical maps to identify candidate woodlands with a history of agricultural use in the different age categories, then ensuring that woodlands were spatially separated by at least 3km. Candidate sites were screened to ensure that they had similar soil type - Worcester, Denchworth or Ragdale surface-water gleys. Only unmanaged, broadleaf woodlands of between 2 and 5 ha were selected, to be of comparable size.

In October 2016, the group sampled 10 soil cores from a 20 x 20m sampling area at the centre of each site, selected to reduce the influence of edge effects. Soil cores were taken at least 1m from the nearest tree and at least 5m separated each core from the others. The soil core was taken for the 0-20cm soil layer using a Dutch auger. Three cylindrical core rings ($100cm^3$) were taken from the mineral soil horizon to determine soil bulk density. Equipment was cleaned between visiting sites by scrubbing and washing off soil material using a brush and tap water, followed by a thorough application of Propeller isopropanol-based disinfectant spray, in line with biocontrol protocols.

### *2.3.2 Soil characterisation*

For Ashwood *et al.* (2019), soil samples were analysed by Forest Research's soil laboratory services at Alice Holt Lodge, Farnham, UK. Soil pH was measured in a 1:2.5 soil to water ratio suspension. They determined the concentration of available soil nitrate ($NO_3^-$) and ammonium ($NH_4^+$) using a colorimetric analysis of 1M KCL soil extractions. Soil C and N concentrations were measured using a C:N Elemental Analyser (Carlo Erba [THERMO], FLASH EA 1112 Series). After Ashwood's analysis, I determined the predominant soil texture of each site (and

extracted fungal and bacterial DNA, see Chapter 3). I discovered that there was an insufficient spread of soil texture classes across land-use types, making it impossible to model the effects of soil texture and land use in the same GLMM. However, I estimated the clay, sand, and silt percentage of each site's predominant soil texture. Percentage clay, silt and sand estimated to be the approximate geometric centroid of each texture class, identified by overlaying a 1% increment soil texture triangle (Soil Survey Staff, 2014) on to a plot indicating the centroids of each UK soil texture class using the Soil Survey of England and Wales class delineations (Moeys, 2018) following the method of Levi (Levi, 2017). I analysed this separate to the higher resolution data for soil properties.

### *2.3.3 Statistical analysis*

All the analysis and the plotting of figures was conducted in the R software environment (v3.6.0, R Core Team, 2020). The data and a script reproducing the analysis are provided in the appendix.

### *2.3.4 Spatial autocorrelation*

Soil biology and chemistry are extremely likely to be altered by biotic and abiotic processes occurring at a range of spatial scales (Weil & Brady, 2017). To test for spatial autocorrelation, Distance-based Moran's eigenvector maps (dbMEM) were computed for spatial point datasets in the package *adespatial* (Dray *et al.*, 2020), producing what approximates to a PCA of the truncated distance diagonal matrix of the data (Borcard & Legendre, 2002; Dray *et al.*, 2006). The dbMEM components that are useful for explaining variation in the response variables can be determined through forward selection. These components can then be included as fixed effects in future models, with each univariate MEM incorporating information from both spatial axes. Forward selection is known to have inflated type I error (Blanchet *et al.*, 2008). Furthermore, this analysis was conducted on site averages, a simplification of the available data. In light of this, p and r-squared values were not directly used to determine the significance or explanatory power of the spatial data, but instead as a conservative way of determining which, if any, MEMs to include in future models.

I calculated the dbMEM for the spatial distribution of the WrEN sites. Adjacent sites that had previously been allocated the coordinates of the woodland sampling location were separated, assigning the agricultural sites the coordinates of an adjacent field. For the purpose of the analysis the only requirement was that sites did not share coordinates. When there were multiple adjacent sites to choose from the closest was selected. The selection choice is unlikely to meaningfully impact the analysis as the small distance between adjacent sites and the large distance between separate locations differs by an order of magnitude. Forward selection indicated that the first MEM significantly improved model fit for soil pH, nitrate, total nitrogen, and organic carbon. Including additional MEMs beyond the first never significantly improved model fit. In the case of soil ammonium, C:N ratio, or bulk density, no MEMs improved model fit. Despite not improving model fit in all cases, all future models included the first MEM as a fixed effect to maintain model structure consistency across response variables. Soil texture components (estimated percentage of clay, silt, and sand) were not identified as spatially autocorrelated using forward selection. Site information, including the MEM data is listed in Table S2.e.

### *2.3.5 Multivariate analyses of soil properties*

I summarised the correlation between soil characteristics with a Principal Component Analysis (PCA); in which the response variables were rescaled to have a mean of 0 and equal variance. I conducted the analysis using the "stats" package function *prcomp* in R. I attempted to test whether land-use types differed in their multivariate distribution (centroid location and dispersion) with a PERMANOVA and post-hoc pairwise PERAMNOVAs between each pair of land use types. To determine whether significant differences between land-use groups were due to centroid location and not merely differences in dispersion, I tested for homogeneity of dispersion with a permutation test.

### *2.3.6 General linear mixed-effect models of soil properties*

I tested for an effect of land use type on pH, soil organic carbon, total nitrogen, C:N ratio, nitrate, and ammonium. Organic carbon was often the sole component of a core's total carbon, with inorganic carbon values often very low. Because of this, organic carbon results

should be taken as highly suggestive of those for total carbon and modelling of total carbon has been omitted to avoid analysing the same data twice.

General linear mixed-effect models (GLMM) were used to test for differences in soil properties between land-use types. Models of varying structure were compared and selected to minimise the model Akaike Information Criterion (AIC) and maximise model parsimony. Model AIC describes the relative error around predictions, the information loss, and helps ensuring models are neither over nor under fitted. When choosing model error families, I prioritised maintaining similar model structure across all response variables along with reflecting the trends in the raw data. Models were fit using the maximum likelihood (ML) method. I constructed models of the response of individual soil properties to the fixed effects of land-use type. Including spatial data significantly improved the fit of fewer than half of the models, affecting total organic carbon, total nitrogen, and nitrates, but to keep the model structure consistent it has been included in all response variable models. In all cases, mixed-effect model AIC improved or remained equivalent (within 2 AIC) when spatial information, in the form of the MEM, was incorporated as a fixed factor. The final model structure explained the response variable using site land use and the first spatial MEM as fixed effects, with sampling site included as a random effect. The significance of land-use on each response variable was assessed using the *anova(response_model)* function in *R,* which acquired a GLMM method from the lme package.

To determine which land-use categories meaningfully differed from one another, I produced predictions of each response variable effect from the GLMMs for each land use type while controlling for the effect of spatial autocorrelation (i.e. making predictions with the value of the MEM set at 0). The 95% confidence intervals of these values were calculated, and non-overlapping confidence intervals would be taken to be indicative of meaningful differences between land-use types.

### *2.3.7 Land use as a predictor of % clay, % silt, and % sand*

As forward selection did not indicate spatial autocorrelation of soil texture components, ANOVAs were sufficient to compare clay, silt, and sand percentages across sites of different land-use.

## 2.4 Results

### *2.4.1 Response variable correlation*

Correlation of soil characteristics for soil cores was summarised in a PCA (Table 2.2). The first three principal components were retained due to their explanatory power (Figure 2.1) and explained approximately 78% of the data's variance. The largest component (PC1) indicated a positive correlation between total nitrogen, organic carbon, and ammonium. PC2 indicated a negative correlation of pH with the C:N ratio of the soil core. PC3 was mostly composed of an effect of nitrate variation. Sites with different land-use histories generally did not cluster together across these principal components. PERMANOVA testing indicated that there were significant differences in land use grouping multivariate distribution (Supplementary Table S2.c.1), but there were also significant differences in dispersion between some groups (Supplementary Table S2.c.2). Significant results of a PERMANOVA test cannot differentiate between significant differences in dispersion (which are not of interest in testing this hypothesis) and significant differences in the group centroid (the test of our hypothesis). Due to this ambiguity care must be taken in interpreting these results, but I have included them in as supplementary tables.

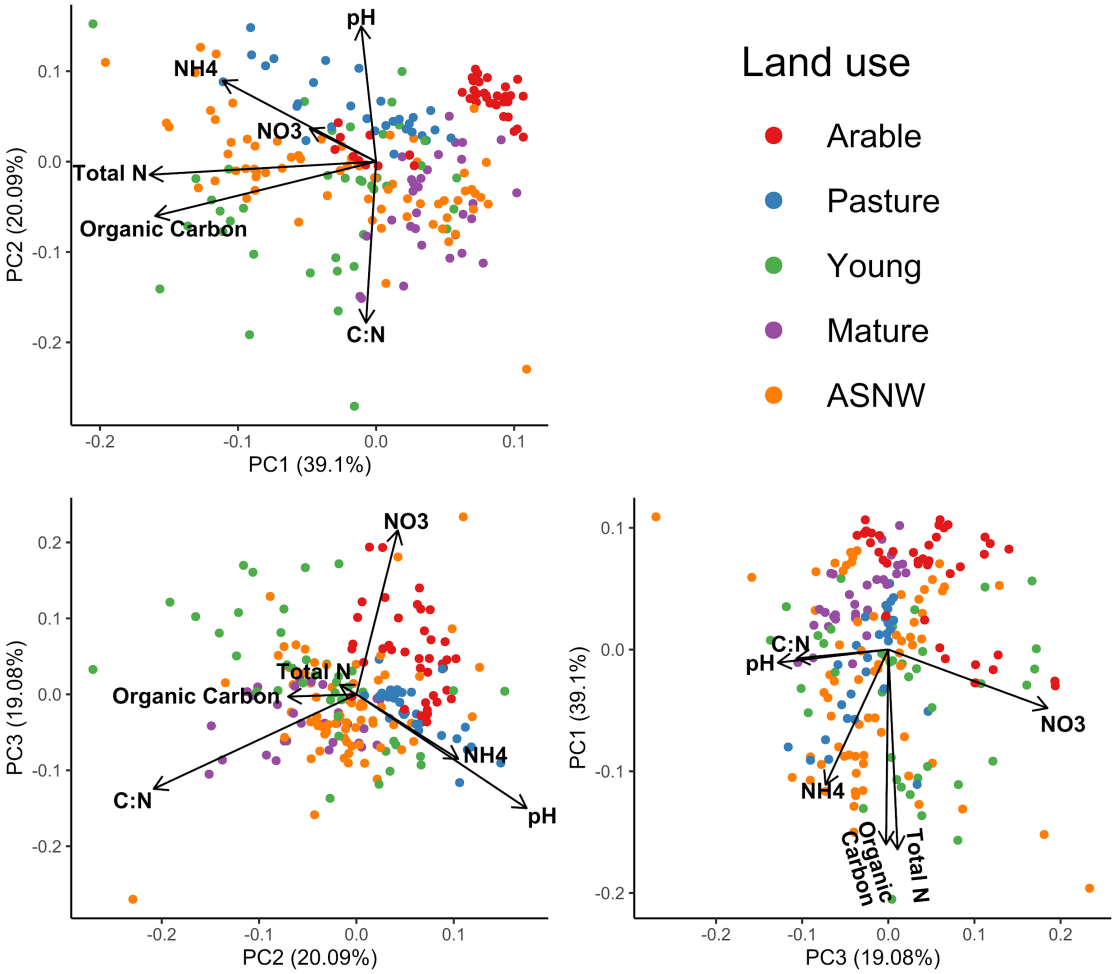


Figure 2.1. Principal Component (PC) scores for each soil core, with land use indicated by point colour. The first three PCs are plotted reciprocally against one another. PC loadings for each variable are indicated by black arrows with the proportion of variance explained by each PC indicated on the axis.

Arable sites showed the greatest degree of clustering, generally being associated with positive values of PC1 and PC3 (i.e. exhibiting low concentrations of total nitrogen, organic carbon, and ammonium, and high concentrations of nitrates as well as low C:N ratios. This reflects the interrupted cycling of carbon and nitrogen caused by the removal of crops, and potentially the addition of nitrate fertilisers).

Table 2.2. Principal component loadings for all soil properties

| component | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 |
|---|---|---|---|---|---|---|
| pH | -0.042 | 0.577 | -0.494 | 0.601 | -0.243 | -0.042 |
| Total N (%) | -0.632 | -0.056 | 0.042 | -0.035 | -0.319 | 0.702 |
| Organic C (%) | -0.617 | -0.231 | -0.001 | 0.010 | -0.278 | -0.699 |
| C:N | -0.028 | -0.687 | -0.410 | 0.470 | 0.349 | 0.126 |
| $NH_4$ mg $kg^{-1}$ | -0.428 | 0.345 | -0.280 | -0.299 | 0.727 | -0.026 |
| $NO_3$ mg $kg^{-1}$ | -0.185 | 0.140 | 0.712 | 0.572 | 0.334 | -0.018 |

Table 2.3: The significance of the effect of land-use type on each soil response variable, generated from an ANOVA on each mixed-effect model. Adjusted p-values provided using the Benjamini & Hochberg adjustment method. Here bolding indicates a p-value < 0.05, and (*) indicates p <0.001.

| Response Variable | d.f. | F-value | P-value | Adjusted p-value |
|---|---|---|---|---|
| pH | (4, 15) | 0.868 | 0.505 | 0.505 |
| Bulk Density | (4, 15) | 4.355 | **0.016** | **0.022** |
| $NO_3^-$ | (4, 15) | 3.112 | **0.047** | 0.055 |
| $NH_4^+$ | (4, 15) | 8.490 | **<0.001*** | **0.003** |
| Organic Carbon | (4, 15) | 6.144 | **0.004** | **0.009** |
| Total Nitrogen | (4, 15) | 4.643 | **0.012** | **0.021** |
| C:N ratio | (4, 15) | 8.414 | **<0.001*** | **0.003** |

### *2.4.2 Mixed effect models of soil core characteristics*

There was a statistically significant effect of land use on all response variables except for pH and nitrate (Table 2.3). However, differences in nitrate concentrations were observed before p-value adjustment for multiple testing, undertaken using the Benjamini & Hochberg method. The predicted effect of land use on each response variable and their 95% confidence intervals was compared following Thomas *et al.* (2017). Post-hoc tests of GLMMs remain contentious methods of contrasting levels within significant fixed effects. Readers who prefer the comfort

of p-values may find the results of these tests in the supplementary material, though the interpretation of the data suggested by such tests is identical to that provided by the confidence intervals of Figure 2.2.

The effects of land-use type on pH were non-significant and all land use types had overlapping confidence intervals (Figure 2.2A). The confidence intervals around the mean effect sizes of land use on bulk density suggest a significant difference between arable and young woodland sites, with arable farms having higher bulk density than young woodlands (Figure 2.2 B) where all other land-use types have soils of intermediate densities. Nitrate and ammonium values differed between agricultural land-use types, with pasture sites having lower nitrate and higher ammonium concentrations; pasture ammonium concentrations were also higher than those found in all woodland land-use types (Figure 2.2, C and D). Although, p-value adjustment indicates that observed differences in nitrate concentration may result from Type I error. Total organic carbon and total nitrogen were both higher in young woodland and ASNW than in arable farmland. However, for pasture and mature woodland organic carbon and total nitrogen levels were not different from other land use types (Figure 2.2, E and F). Broadly carbon and nitrogen concentrations were higher in less disturbed systems with woodlands generally having higher concentrations than pasture, and pasture higher concentrations than arable farmland. Carbon to nitrogen ratio was significantly higher for woodland sites than agricultural sites, but not meaningfully different within those groups (Figure 2.2, G). Although working at the scale of individual soil cores lent itself to analysing core concentration data, rather than landscape-scale stocks data, this information is calculated and summarised in Table 2.4.

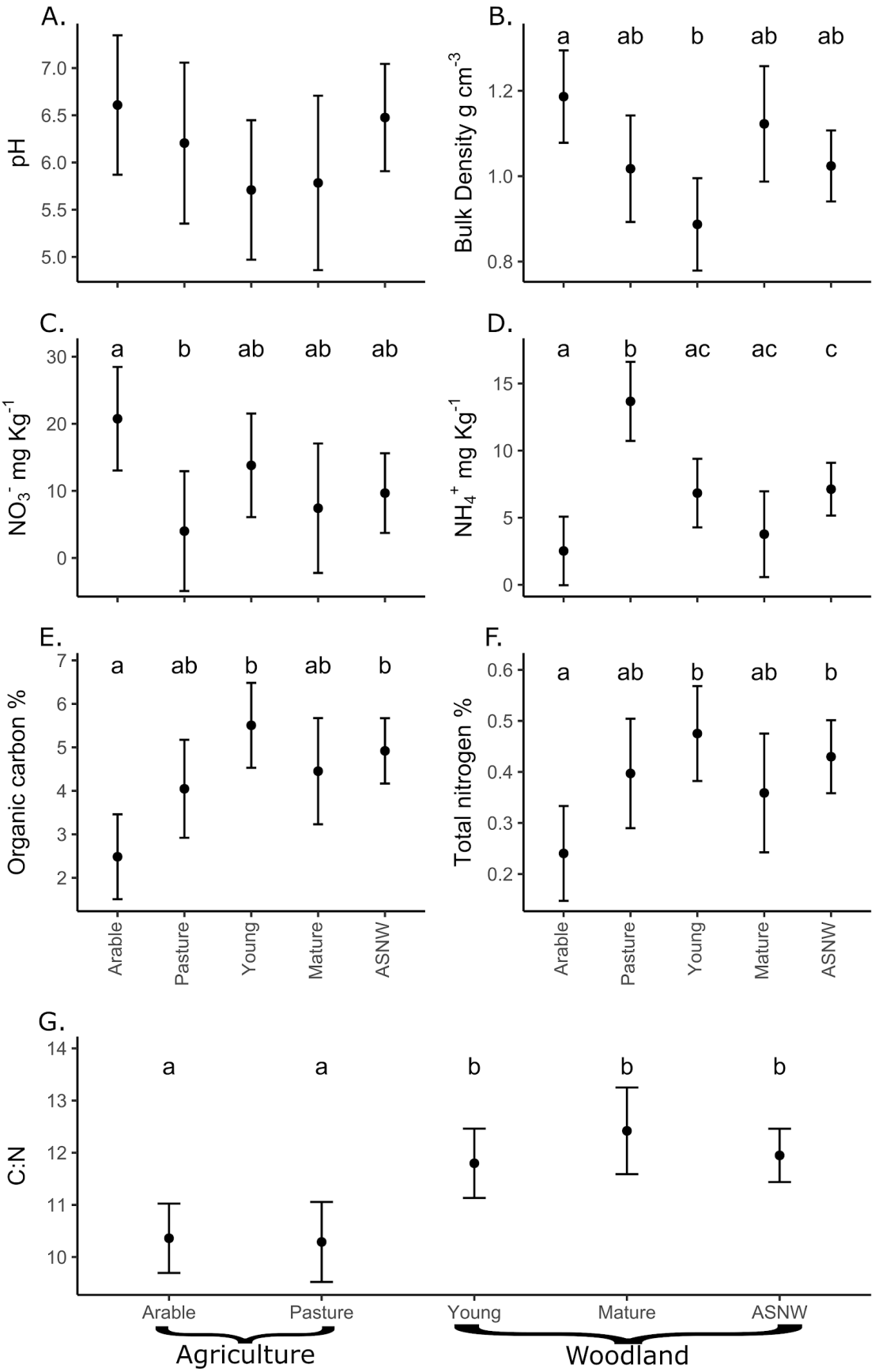


Figure 2.2. Mean predicted effect of land use on soil characteristics, error bars represent their 95% confidence intervals. Lettering indicates groupings with non-overlapping confidence intervals, the size and position of individual plots was chosen merely to ease interpretation

Table 2.4. Carbon and nitrogen stocks for the 0-20 cm soil layer estimated from the product of the average bulk density of each site at 0-20cm and the percentage of each nutrient converted to tonnes per hectare.

| Land Use | Total C (t $ha^{-1}$) | Organic C (t $ha^{-1}$) | Total N (t $ha^{-1}$) | Bulk Density (g $cm^{-3}$) |
|---|---|---|---|---|
| Arable | 58.533 | 58.205 | 5.637 | 1.182 |
| Pasture | 82.964 | 82.474 | 8.084 | 1.016 |
| Young woodland | 96.014 | 94.398 | 8.161 | 0.883 |
| Mature woodland | 86.345 | 86.207 | 6.796 | 1.165 |
| ASNW | 105.718 | 103.283 | 9.023 | 1.011 |

### *2.4.3 Soil textural components did not differ between land uses*

I found no significant difference between sites with different land uses for estimated percentages of clay ($F_{(6,16)} = 1.49$, $p = 0.25$), silt ($F_{(6,16)} = 2.25$, $p = 0.11$), or sand ($F_{(6,16)} = 1.09$, $p = 0.40$). The texture classes of each site are displayed in Figure 2.3.

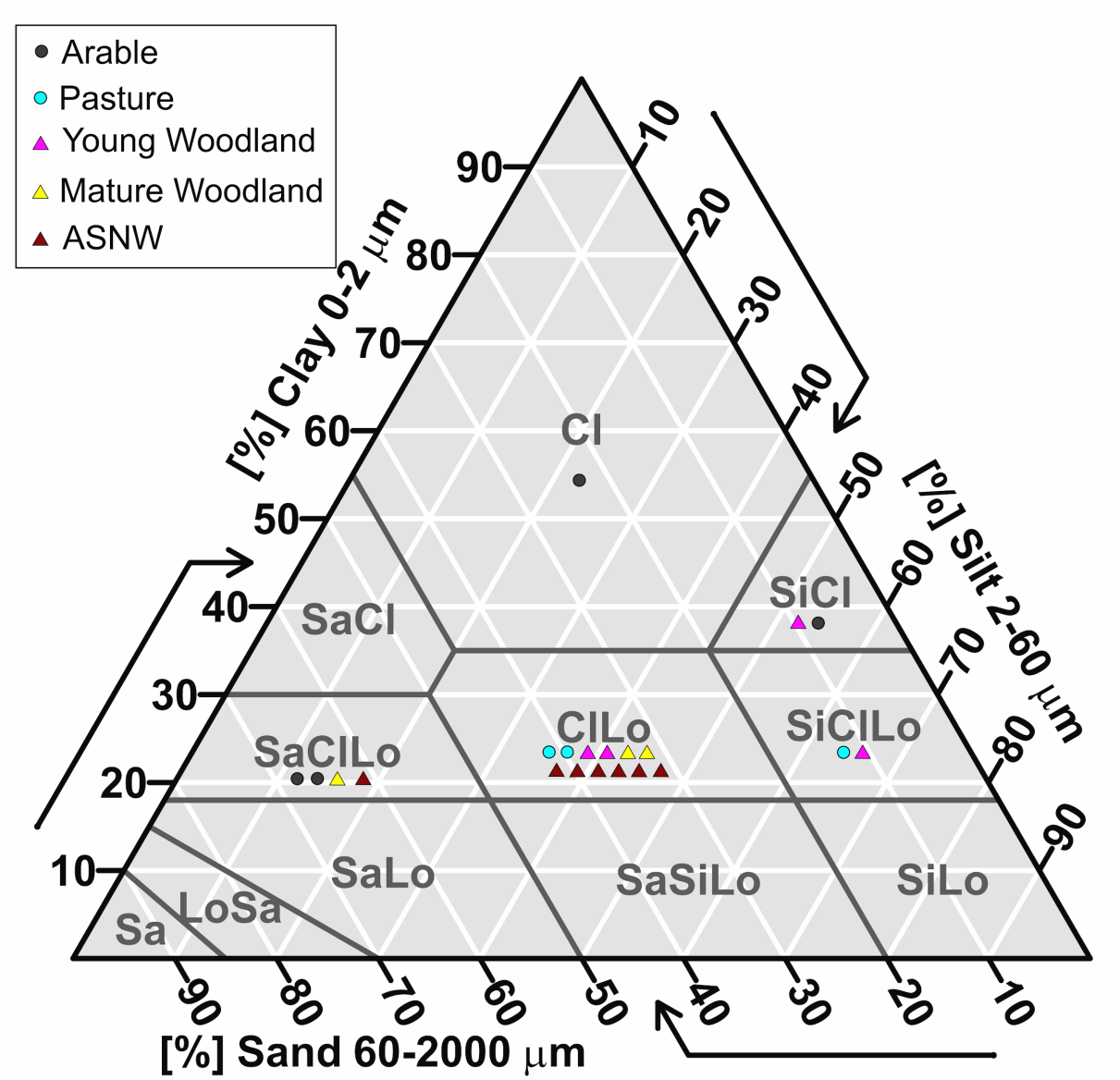


Figure 2.3. Predominant soil texture class for each site. Agricultural sites are indicated with circles (arable - grey, pasture - cyan) and woodland with triangles (young - magenta, mature - yellow, ASNW - dark red). All points in a texture class are estimated as having the same value: its geometric centroid. Points are arranged within the class to aid visibility.

## 2.5 Discussion

### *2.5.1 Soil changes across a wooded agricultural landscape*

Woodland creation in the agrarian landscape of the UK midlands has produced meaningful differences in the soil characteristics of sites with different land-use histories. As I hypothesised, soil organic carbon and soil nitrogen of sites generally decreased across a gradient of low to high disturbance, inverse to woodland age, with significantly lower values for arable sites compared to young or ancient woodlands. The C:N ratio of soils was significantly higher following woodland creation, with woodland sites exhibiting a higher C:N ratio than both arable and pasture agricultural sites. Although, C:N ratio was higher in older woodlands there were no significant differences between the C:N ratios of woodland types. Surprisingly, pH did not vary across land use categories and bulk density only significantly differed between arable sites and young woodlands, suggesting that compaction was highest at arable sites but that there was great variation in compaction and density in pastures and older wooded sites. Young, wooded sites had the lowest bulk densities, perhaps indicating that the disruption of the soils in preparation for and during planting of trees may have persistent effects on soil density for half a century or longer. Ammonium concentrations were significantly higher on pasture sites where animal waste is regularly input than all other land use types, as hypothesised based upon nitrogen inputs (Cameron et al., 2013), and higher in ancient woodlands than arable farms. Nitrates were higher in arable soils than pasture, although not significantly so after accounting for multiple testing. Neither were significantly different from woodlands of any age category, although woodlands, especially older woodland generally exhibited lower nitrate levels, more similar to pasture sites. The three components of soil texture (estimated clay, silt, and sand percentages) did not differ significantly between land uses, reinforcing the assertion that these sites are generally comparable. Despite the relative homogeneity of the landscape, sufficient heterogeneity existed to successfully demonstrate differences in soil characteristics between land use types while accounting for the spatial autocorrelation of sites.

My results indicate that broadleaf woodlands in this landscape are broadly similar in their soil characteristics regardless of age structure. Although mature woodlands did not significantly differ from arable farmland for any measure, except for their higher C:N ratio, woodland creation and ASNW soils also had significantly higher concentrations of organic carbon and

total nitrogen, consistent with previous work that indicates that soil carbon and nitrogen increases with afforestation (C. A. Creamer *et al.*, 2016; Jangid *et al.*, 2011). Pasture farmland only differed from the other land-use types when examining the C:N ratio, ammonium concentration and nitrate concentration. Differences in C:N ratio will relate both to the increased input by weight of organic material, with high C:N ratio, of all types into woodlands when compared to farmland, but also the quality of this input in terms of its C:N ratio may be higher for lignified, woody material than for grass or cropland vegetation (Dickinson, 2012). These patterns of nutrient concentration broadly matched those of nutrient stocks at the site level, differing only in that the low bulk density of young woodland soils resulted in young woodlands having lower carbon and nitrogen stocks than ASNW (Table 2.4) despite having higher concentrations of carbon and nitrogen per gram of soil.

Differences in carbon, nitrogen, and C:N ratio are likely to be related to the history of the sites. At all sites, the immobilisation of decomposing matter promotes the accumulation of both carbon and nitrogen in the soils. These organic inputs to the soil generally contain much more carbon than nitrogen, but highly lignified litter from trees tends to have higher C:N ratios than ground flora or grass (Dickinson, 2012). The limited range of organisms capable of degrading lignin makes it hight recalcitrant (Rahman et al. 2013) and keeps lignified carbon in the soil for a long time. This is reflected in the increased ratio of carbon to nitrogen in woodland soils. As most vegetative material is harvested from arable sites, the organic input from plants is low and of low C:N ratio. On pasture grazing, and the eventual removal of grazers, also removed organic matter from the system although some is returned in the form of animal waste with a comparatively low C:N ratio. The soil microbiome also influences and is influenced by soil C:N ratios due to the differing requirements of soil bacteria and fungi and different capacity to transform carbon and nitrogen through various nutrient cycling processes (Weil & Brady, 2017).

Seasonal effects may also be important in explaining the differences between arable, pasture and woodland sites. Previous work has shown that the magnitude of differences between woodland and grassland C:N ratios may exhibit seasonality (C. A. Creamer *et al.*, 2016). Creamer *et al.* observed a significant difference between grassland and woodland C:N ratios

in May, but not October, in southern Texas woodland-grassland transitions. Without knowing more about the seasonality of the C:N ratio for the sites included in my study it is unclear whether the C:N ratio difference between land use types would change over the course of the year.

The nitrification process is facilitated by temperature-sensitive reactions. As the sampling occurred in October 2016, when the temperature was relatively cool (averaging 11°C) the normally rapid conversion of ammonium to nitrate may have slowed, leading to the build-up of ammonium on the pasture sites from animal waste. Higher concentrations of nitrate on arable sites, which is normally removed by crop uptake, may be excess fertiliser that was not uptaken. However, without further details of the management of these farms it is difficult to know this with confidence. Additionally, the differences in nitrate concentrations between land-use types was not significant after p-value adjustment for multiple testing.

I did not observe any meaningful differences in pH across land-use types. Some previous work has documented the gradual acidification of soils in response to woodland ageing (Zhu *et al.*, 2010), but the phenomenon is not ubiquitous (Wilson *et al.*, 1997; Fichtner *et al.*, 2014). Soil pH, alongside soil organic carbon and soil texture, is well understood to be a key driver of microbial community assembly (Fierer & Jackson, 2006; Griffiths *et al.*, 2011; R. E. Creamer *et al.*, 2016; Uroz *et al.*, 2016; Lladó *et al.*, 2017). As these sites are broadly similar in pH it may be easier to differentiate the effects of site age on microbial community assembly from those of pH change.

Taken together, my results indicate that changes to soil characteristics following woodland creation can occur very rapidly, most likely during the first 50 years of establishment. Woodlands over 50 years old (such as the young woodlands included in this study) were already quite distinct from arable farmland and were similar to ASNW. No other work has considered such a breadth of land-use types in a single study. Because of this, one cannot assume that the combined finding of previous research, showing differences in soil characteristics observed across sites, is a product of their land use, rather than their highly

varied geography and methods of analysis. This work removes differences in methods of analysis that are known to inflate the error surrounding measures of soil chemistry between studies (Sun *et al.*, 2015); e.g. collection methods, soil storage times, soil handling, and geographic separation. As such, it is the first time that agricultural sites, recently afforested woodland, and ancient woodland sites will be compared using identical methods in a single study. Previous work has characterised the changes to soils following reforestation within this short period in space for time studies in a variety of woodland systems, highlighting the accumulation of soil carbon and nitrogen (Zhu *et al.*, 2010; Larkin *et al.*, 2014; Mackay *et al.*, 2016). Prior to this study, the degree to which British soils would mimic the response to afforestation exhibited in different biomes with very different vegetation cover was unknown. Differences in soil response to land-use change across biomes is likely to be non-trivial - recent modelling approaches have indicated that the response trajectory of important soil properties, such as soil carbon storage, to afforestation will markedly differ between vegetation types and climates (Krause *et al.*, 2016). Nonetheless, rapid alteration to soil properties over the first 50 years of afforestation is consistent with previous studies.

As I do not know the type of farming conducted on these woodland sites prior to woodland creation it is unclear whether the changes in soil characteristics following woodland creation would be more dramatic in previous pasture sites than in previous arable sites. Woodland soil properties were very similar across all woodland ages. However, the variations in soil properties across these sites do not show a directional trend with increasing woodland age. Nonlinear processes may dominate as the woodland ages and canopy structure changes. Many studies have documented the accumulation of carbon and nitrogen in soils during woodland creation (Zhu *et al.*, 2010; Jangid *et al.*, 2011; Larkin *et al.*, 2014; C. A. Creamer *et al.*, 2016; Mackay *et al.*, 2016). When there are good temporal comparisons this often shows a flattening curve, with the rate of carbon and nitrogen accumulation slowing in older woodland creation sites after maturation. Work comparing older woodlands sometimes indicates declines in soil carbon and nitrogen (both stock and concentrations) when comparing mature secondary and ancient woodland (Zhu *et al.*, 2010; Wilson *et al.*, 1997; Fichtner *et al.*, 2014), but not in all cases e.g.(Wilson *et al.*, 1997; Hoogmoed *et al.*, 2012). Carbon and nitrogen stocks may be lower in ASNW than mature woodlands because of changes in woodland productivity over time (Berger *et al.*, 2004). Soil nutrient stocks are a balance of

flows into and out of soils, and if the above-ground pool of carbon exhibits lower productivity and smaller turnover in older woodlands, inputs into the system decline and may reduce the size of the pools of soil carbon and nitrogen. Interpreting these space-for-time studies for woodland age must be done with caution; site attributes and specific history are often more important than broadscale effects of succession and nutrient accumulation (Walker *et al.*, 2010). Studies have often focused on comparisons between agricultural and ex-agricultural woodlands (particularly silvicultural plantations) or on differences between mature and ancient woodlands. Fewer studies have examined woodland differences across long chronosequences, so nonlinear soil changes over time have been documented (Zhu *et al.*, 2010), but in few ecosystems. This study adds British woodlands as a new ecosystem to long-term afforestation studies.

The bulk density of soils only meaningfully differed between arable and young woodlands. The bulk density data was collected at the site level, so it is not possible to determine the bulk density of individual soil cores. It is possible that woodland creation activities, such as tillage may have loosened soils, the legacy of this may lead to lower bulk density in these young woodlands even 50 years after planting. Alternatively, soil organic carbon may be influencing bulk density by improving soil structure through the encouragement of aggregation (Weil & Brady, 2017). Soil carbon strongly, negatively correlates with bulk density at these sites, and this may be one reason why young woodlands have such low bulk density. It is possible to examine soil properties at the landscape scale by scaling up the proportion by weight of the soil nutrients that were examined to kg per hectare and modelling the effect of site land use on these landscape-level soil properties. However, as density cores must be sampled separately from soil cores used to determine soil properties it is not possible to sensibly pair soil property data to soil density data at the sampling level of the soil core. Instead, one would have had to transform the values of soil properties for each core by an average bulk density for the site. The variation would then be partitioned into the model's random site effect, adding little to the model. This information is no doubt of interest at the site, landscape, or biome levels and so I have included the carbon and nitrogen stocks of each land-use type in Table 2.4. It should be noted that these results broadly follow expected patterns for carbon and nutrient stocks accumulation. Furthermore, they reinforce that ancient woodlands remain important carbon sinks both above and below ground. An analysis of the carbon and

nitrogen stocks of these soils can be found in Ashwood *et al.* (2019). As important landscape-scale measures are, changes to local microbial communities may be dictated by processes occurring at the scale of a single soil core, rather than at the landscape scale (Jiao *et al.*, 2018; Chen *et al.*, 2019). As I intend to examine the microbial diversity of these soil cores in the next chapter's analysis of the soils at this small scale is the most appropriate.

Similar to my analysis Ashwood *et al.* (2019) found young woodlands (aged 50-116 years) had higher carbon stocks than arable farmland and differentiated from all farmland in their soil C:N ratios. However, most of their analysis was conducted at the site scale, giving stocks of compounds or elements per hectare, modifying the raw measures by the density of the soil. This approach is less appropriate for analysis of the microbiological of soils, see the next chapter, than analysis of the raw proportional data (e.g. percentages). Additionally, Ashwood *et al.* do not make use of the full potential of mixed-effect models to account for data non0independence. Instead, they account for data non-independence by averaging values within a site which sacrifices statistical power. My analysis preserves statistical power by employing random factors for sites. This allows for additional, more complex analyses to be conducted and for estimates of effects with lower error.

Soil chemistry and structure is a broad and deep field, of which no study can do more than scratch the surface. By focusing on certain aspects of the soil I was limited in my ability to draw conclusions regarding others. I was not able to consider all aspects of soil chemistry. The dataset lacked potentially informative information for essential macronutrients - soil sulphur, phosphorus, and potassium. Although I obtained information regarding the predominant soil texture at each site and the site geology, some combinations of land use and soil texture fully overlap, preventing both variables being included in a statistical model. As land use is often determined by the suitability of soil for cultivation, this overlap is not surprising, though due to the low number of sites for each land use it is likely to be due to chance. The alternative method used, of estimating the percentage of clay, silt, and sand for each site is imprecise, but not without precedent (Levi, 2017). Levi compared the accuracy of geometric centroids and centroids estimated from existing US soil collections in estimating true soil composition. The article highlights the imprecision of geometric centroids at estimating soil composition, but the method remains the only one available to me at this time. These analyses indicate that the similarity of these sites makes them well suited for investigating the effect of site age on soil chemistry, but less well suited for broad-scale descriptions of soil characterisation across

geological and soil texture gradients. Finally, fitting GLMM models to nitrate and ammonium concentrations was challenging, as the data best fitted a gamma error distribution. However, I was unable to produce a GLMM model that reflected the shape of raw data as well as the GLMM model assuming a normal error distribution, the approach used for all models. The scripts and data needed to investigate these results are provided in the supplementary material.

### *2.5.2 Conclusions*

Few studies have effectively compared the differences in soil chemistry experienced across agricultural land and broadleaf woodland sites spanning a large woodland age range. Until now, understanding this process in British woodlands was lacking. The WrEN project has been able to do this for a subset of woodlands in the British midlands. These woodland sites had higher C:N ratios than arable ones, also having generally higher organic carbon and nitrogen concentrations; and differed from pasture in having higher C:N ratio and lower ammonium levels. Broadly woodland soils were more like pasture than arable soils, the former rarely differing from any category of woodland in a meaningful way. Mature woodland often exhibited similar characteristics to arable land (for organic carbon, nitrogen, ammonium, and bulk density) while one or both other categories of woodland were meaningfully differentiated from arable soils. This highlights the non-linear direction of change in ageing woodland soils. This work is consistent with previous studies and well describes successional processes in this region. However, these results may not be descriptive of woodland creation and ageing processes across the country or in coniferous plantations. Future work should aim to ask: how these soil characteristics affect the microbial communities found in these woodlands? And does microbial composition closely match patterns observed in soil properties, or does the age of these woodlands influence the assembly of these communities?

### *2.5.3 Synthesis and applications*

As reported in Ashwood *et al.* (2019) afforestation on agricultural land restores the carbon stocks of woodland creation sites relatively quickly within the first 50 years. My reanalysis using mixed-effect models allows for the separation of young (<50 years since planting) and mature (established around 110 years ago) woodland without sacrificing statistical power.

Additionally, I reframe the focus from the landscape scale (units per hectare) to a fine scale, more appropriate for microbiological systems (units per gram or percentages). This also allows for the analysis of soil properties without transforming values by soil density, as is required for per hectare calculations that deal in volumes and weights. Although our conclusions support one another, the reanalysis is necessary in the context of an exploration of microbial systems in the following chapters. These young woodland creation sites have high soil organic carbon and total nitrogen concentrations, but low bulk density, older woodland typically had higher soil bulk density but slightly lower organic carbon and total nitrogen concentrations. This may have implications for how planting is conducted as soil disturbance during planting is likely the cause of this low bulk density. Despite this, woodlands over 50 years old did not significantly differ in their organic carbon or total nitrogen concentrations, or stocks as Ashwood reports. This suggests that factors other than woodland age are the dominant influences of carbon accumulation in woodlands as they grow.

**2.6 References**

Anon (2019) *Woodland Statistics - Forest Research*. [Online] [online]. Available from: https://www.forestresearch.gov.uk/tools-and-resources/statistics/statistics-by-topic/woodland-statistics/ (Accessed 31 October 2019).

Ashwood, F., Watts, K., Park, K., Fuentes-Montemayor, E., Benham, S. & Vanguelova, E.I. (2019) Woodland restoration on agricultural land: long-term impacts on soil quality. *Restoration Ecology*. 42182.

Berger, U., Hildenbrandt, H. & Grimm, V. (2004) Age-Related Decline in Forest Production: Modelling the Effects of Growth Limitation, Neighbourhood Competition and Self-Thinning. *The Journal of ecology*. 92 (5), 846–853.

Blanchet, F.G., Legendre, P. & Borcard, D. (2008) Forward selection of explanatory variables. *Ecology*. 89 (9), 2623–2632.

Borcard, D. & Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological modelling*. 153 (1), 51–68.

Burton, V., Moseley, D., Brown, C., Metzger, M.J. & Bellamy, P. (2018) Reviewing the evidence base for the effects of woodland expansion on biodiversity and ecosystem services in the United Kingdom. *Forest ecology and management*. 430366–379.

Cahalan, C., Hester, A., Humphrey, J., Kirby, K., Moffat, A. & Valatin, G. (2011) 'Woodlands', in National Ecosystem Assessment (ed.) *The UK National Ecosystem Assessment Technical Report*. [Online]. Cambridge: UNEP-WCMC. pp. 241–294.

Cameron, K.C., Di, H.J. & Moir, J.L. (2013) Nitrogen losses from the soil/plant system: a review. The Annals of applied biology. 162 (2), 145–173.

Cannell, M.G.R. (2003) Carbon sequestration and biomass energy offset: theoretical, potential and achievable capacities globally, in Europe and the UK. *Biomass and Bioenergy*. 24 (2), 97–116.

Chazdon, R. & Brancalion, P. (2019) Restoring forests as a means to many ends. Science 365 (6448) p.24–25.

Chazdon, R.L. (2008) Beyond deforestation: restoring forests and ecosystem services on degraded lands. *Science*. 320 (5882), 1458–1460.

Chen, W., Jiao, S., Li, Q. & Du, N. (2019) Dispersal limitation relative to environmental filtering governs the vertical small-scale assembly of soil microbiomes during restoration Gaowen Yang (ed.). *The Journal of applied ecology*. 355.

Creamer, C.A., Filley, T.R., Boutton, T.W. & Rowe, H.I. (2016) Grassland to woodland transitions: Dynamic response of microbial community structure and carbon use patterns. *Journal of Geophysical Research: Biogeosciences*. 121 (6), 2016JG003347.

Creamer, R.E., Hannula, S.E., Van Leeuwen, J.P., Stone, D., Rutgers, M., Schmelz, R.M., Ruiter, P.C. de, Hendriksen, N.B., Bolger, T., Bouffaud, M.L., Buee, M., Carvalho, F., Costa, D., Dirilgen, T., Francisco, R., Griffiths, B.S., Griffiths, R., Martin, F., Silva, P.M. da, *et al.* (2016) Ecological network analysis reveals the inter-connection between soil biodiversity and ecosystem function as affected by land use across Europe. *Applied soil ecology: a section of Agriculture, Ecosystems & Environment*. 97112–124.

DEFRA (2018) *A Green Future: Our 25 Year Plan to Improve the Environment*.

Deng, L. & Shangguan, Z. (2017) Afforestation Drives Soil Carbon and Nitrogen Changes in China. *Land Degradation & Development*. 28 (1), 151–165.

Dickinson, C.H. (2012) *Biology of Plant Litter Decomposition V2*. Elsevier.

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N. & Wagner, H.H. (2020) *adespatial: Multivariate Multiscale Spatial Analysis*.

Dray, S., Legendre, P. & Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological modelling*. 196 (3), 483–493.

Fichtner, A., von Oheimb, G., Härdtle, W., Wilken, C. & Gutknecht, J.L.M. (2014) Effects of anthropogenic disturbances on soil microbial communities in oak forests persist for more than 100 years. *Soil biology & biochemistry*. 7079–87.

Fierer, N. & Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*. 103 (3), 626–631.

Forestry Commission (2019) *Forestry Statistics 2019*.

Fuller, L., Fuentes-Montemayor, E., Watts, K., Macgregor, N.A., Bitenc, K. & Park, K.J. (2018) Local-scale attributes determine the suitability of woodland creation sites for Diptera Jörg Müller (ed.). *The Journal of applied ecology*. 55 (3), 1173–1184.

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M. & Whiteley, A.S. (2011) The bacterial biogeography of British soils. *Environmental microbiology*. 13 (6), 1642–1654.

Hermy, M. & Verheyen, K. (2007) Legacies of the past in the present-day forest biodiversity: a review of past land-use effects on forest plant species composition and diversity. *Ecological research*. 22 (3), 361–371.

Hoogmoed, M., Cunningham, S.C., Thomson, J.R., Baker, P.J., Beringer, J. & Cavagnaro, T.R. (2012) Does afforestation of pastures increase sequestration of soil carbon in Mediterranean climates? *Agriculture, ecosystems & environment*. 159176–183.

HS2 Select Committee (2016) *Petitions and publications - The Woodland Trust Petitioner*.

Jangid, K., Williams, M.A., Franzluebbers, A.J., Schmidt, T.M., Coleman, D.C. & Whitman, W.B. (2011) Land-use history has a stronger impact on soil microbial community composition than aboveground vegetation and soil properties. *Soil biology & biochemistry*. 43 (10), 2184–2193.

Jefferson, R.G. (2008) Biological Flora of the British Isles: Mercurialis perennis L. *The Journal of ecology*. 96 (2), 386–412.

Jiao, S., Chen, W., Wang, J., Du, N., Li, Q. & Wei, G. (2018) Soil microbiomes with distinct assemblies through vertical soil profiles drive the cycling of multiple nutrients in reforested ecosystems. *Microbiome*. 6 (1), 146.

Kalinina, O., Cherkinsky, A., Chertov, O., Goryachkin, S., Kurganova, I., Lopes de Gerenyu, V., Lyuri, D., Kuzyakov, Y. & Giani, L. (2019) Post-agricultural restoration: Implications for dynamics of soil organic matter pools. *Catena*. 181104096.

Kimberley, A., Blackburn, G.A., Whyatt, J.D., Kirby, K. & Smart, S.M. (2013) Identifying the trait syndromes of conservation indicator species: how distinct are British ancient woodland indicator plants from other woodland species? Martin Hermy (ed.). *Applied Vegetation Science*. 16 (4), 667–675.

Krause, A., Pugh, T.A.M., Bayer, A.D., Lindeskog, M. & Arneth, A. (2016) Impacts of land-use history on the recovery of ecosystems after agricultural abandonment. *Earth System Dynamics*. 7 (3), 745–766.

Kurganova, I.N., Lopes de Gerenyu, V.O., Mostovaya, A.S., Ovsepyan, L.A., Telesnina, V.M., Lichko, V.I. & Baeva, Y.I. (2018) Effect of Reforestation on Microbiological Activity of Postagrogenic Soils in European Russia. *Contemporary Problems of Ecology*. 11 (7), 704–718.

Larkin, D.J., Steffen, J.F., Gentile, R.M. & Zirbel, C.R. (2014) Ecosystem Changes Following Restoration of a Buckthorn-Invaded Woodland. *Restoration Ecology*. 22 (1), 89–97.

Levi, M.R. (2017) Modified Centroid for Estimating Sand, Silt, and Clay from Soil Texture Class. *Soil Science Society of America journal. Soil Science Society of America*. 81 (3), 578–588.

Lladó, S., López-Mondéjar, R. & Baldrian, P. (2017) Forest Soil Bacteria: Diversity, Involvement in Ecosystem Processes, and Response to Global Change. *Microbiology and molecular biology reviews: MMBR*. 81 (2).

Maas, J., Verheij, R.A., de Vries, S., Spreeuwenberg, P., Schellevis, F.G. & Groenewegen, P.P. (2009) Morbidity is related to a green living environment. *Journal of epidemiology and community health*. 63 (12), 967–973.

Mackay, J.E., Cunningham, S.C. & Cavagnaro, T.R. (2016) Riparian reforestation: are there changes in soil carbon and soil microbial communities? *The Science of the total environment*. 566-567960–967.

Mason, W.L. (2007) Changes in the management of British forests between 1945 and 2000 and possible future trends: Changes in British forest management. *The Ibis*. 14941–52.

Moeys, J. (2018) *soiltexture: functions for soil texture plot, classification and transformation. R package version 1.3. 3*.

Natural England (2008) *State of the Natural Environment*.

Nitsch, P., Kaupenjohann, M. & Wulf, M. (2018) Forest continuity, soil depth and tree species are important parameters for SOC stocks in an old forest (Templiner Buchheide, northeast Germany). *Geoderma*. 31065–76.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rackham, O. (1986) *The History of the Countryside*. London: J M Dent & Sons Ltd.

Rackham, O. (2008) Ancient woodlands: modern threats. *The New phytologist*. 180 (3), 571–586.

Rackham, O. (2012) *Woodlands*. London: William Collins.

Ražauskaitė, R., Vanguelova, E., Cornulier, T., Smith, P., Randle, T. & Smith, J. (2020) Measured and modelled soil organic carbon changes in a pine chronosequence on a podzolic soil. *Frontiers of Environmental Science & Engineering in China*. 8204.

Rahman, M.M., Tsukamoto, J., Rahman, M.M., Yoneyama, A. & Mostafa, K.M. (2013) Lignin and its effects on litter decomposition in forest ecosystems. Chemistry and Ecology. 29 (6), 540–553.

Saiz, G., Byrne, K.A., Butterbach-Bahl, K., Kiese, R., Blujdea, V. & Farrell, E.P. (2006) Stand age-related effects on soil respiration in a first rotation Sitka spruce chronosequence in central Ireland. *Global change biology*. 12 (6), 1007–1020.

Soil Survey Staff (2014) *Keys to Soil Taxonomy*. Vol. 12th ed. Washington, DC.: USDA-Natural Resources Conservation Service.

Sun, S.-Q., Cai, H.-Y., Chang, S.X. & Bhatti, J.S. (2015) Sample storage-induced changes in the quantity and quality of soil labile organic carbon. *Scientific reports*. 517496.

Susyan, E.A., Wirth, S., Ananyeva, N.D. & Stolnikova, E.V. (2011) Forest succession on abandoned arable soils in European Russia – Impacts on microbial biomass, fungal-bacterial ratio, and basal CO2 respiration activity. *European journal of soil biology*. 47 (3), 169–174.

Tatchell-Evans, M. (2016) *Implications of Leaving the EU Forestry*.

The Scottish Government (2019) *Scotland's Forestry Strategy 2019–2029*.

Thomas, R.J. and the Guidebook Team. (2017) *Data analysis with R statistical software: A guidebook for scientists*. Caerphilly: Eco-Explore.

Trap, J., Hättenschwiler, S., Gattin, I. & Aubert, M. (2013) Forest ageing: An unexpected driver of beech leaf litter quality variability in European forests with strong consequences on soil processes. *Forest ecology and management*. 302338–345.

Trudgill, S.T. (1988) *Soil and Vegetation Systems*. Oxford: Oxford University Press.

Uroz, S., Buée, M., Deveau, A., Mieszkin, S. & Martin, F. (2016) Ecology of the forest microbiome: Highlights of temperate and boreal ecosystems. *Soil biology & biochemistry*. 103471–488.

Vindušková, O., Pánek, T. & Frouz, J. (2019) Soil C, N and P dynamics along a 13 ka chronosequence of landslides under semi-natural temperate forest. *Quaternary science reviews*. 21318–29.

Walker, L.R., Wardle, D.A., Bardgett, R.D. & Clarkson, B.D. (2010) The use of chronosequences in studies of ecological succession and soil development. *The Journal of ecology*. 98 (4), 725–736.

Watts, K., Fuentes-Montemayor, E., Macgregor, N.A., Peredo-Alvarez, V., Ferryman, M., Bellamy, C., Brown, N. & Park, K.J. (2016) Using historical woodland creation to construct a long-term, large-scale natural experiment: the WrEN project. *Ecology and evolution*. 6 (9), 3012–3025.

Weil, R.R. & Brady, N.C. (2017) *The Nature and Properties of Soils, Global Edition*. Harlow: Pearson Education Ltd.

Welsh Government (2018) *Woodlands for Wales*.

Whytock, R.C., Fuentes-Montemayor, E., Watts, K., Barbosa De Andrade, P., Whytock, R.T., French, P., Macgregor, N.A. & Park, K.J. (2018) Bird-community responses to habitat creation in a long-term, large-scale natural experiment: Birds and Habitat Creation. *Conservation biology: the journal of the Society for Conservation Biology*. 32 (2), 345–354.

Whytock, R.C., Fuentes-Montemayor, E., Watts, K., Macgregor, N.A., Williams, L. & Park, K.J. (2018) Context-dependent colonization of terrestrial habitat 'islands' by a long-distance migrant bird. *Proceedings of the Royal Society B: Biological Sciences*. 285 (1885), 20181490.

Wilson, B.R., Moffat, A.J. & Nortcliff, S. (1997) The Nature of Three Ancient Woodland Soils in Southern England. *Journal of biogeography*. 24 (5), 633–646.

Wu, X., Xu, H., Tuo, D., Wang, C., Fu, B., Lv, Y. & Liu, G. (2020) Land use change and stand age regulate soil respiration by influencing soil substrate supply and microbial community. *Geoderma*. 359113991.

Zhou, G., Liu, S., Li, Z., Zhang, D., Tang, X., Zhou, C., Yan, J. & Mo, J. (2006) Old-growth forests can accumulate carbon in soils. *Science*. 314 (5804), 1417.

Zhu, W., Cai, X., Liu, X., Wang, J., Cheng, S., Zhang, X., Li, D. & Li, M. (2010) Soil microbial population dynamics along a chronosequence of moist evergreen broad-leaved forest succession in southwestern China. *Journal of Mountain Science; Dordrecht*. 7 (4), 327–338.

# Chapter 3. Soil Microbial Community Responses to Woodland Creation: Long-Term Perspectives from a Natural Experiment


## 3.1 Abstract

Woodland creation on agricultural land is increasingly advocated for, but its effect on soil diversity, and the following successional changes, are poorly understood. It is unclear whether woodland creation sites adequately mitigate for the loss of ancient woodland when ancient woods are lost to land-use change, as the biodiversity of woodland soils by successional stage is mostly unknown. Molecular methods allow for the highly replicated characterisation of soil communities and insights into complex ecological questions. Here I use next-generation sequencing to examine the differences in soil bacterial and fungal biodiversity in 210 soil cores taken from 21 sites spanning a gradient of afforestation made up from four agricultural croplands and three pastures, and three recently forested, four mature, and seven ancient woodlands. I demonstrate that coarse metrics of microbial alpha diversity do not change significantly across these land-use categories, but important aspects of microbial community composition change between woodland and arable sites. Pasture site diversity sits intermediately along a spectrum, having similarities to both arable land and woodlands. This study highlights stark differences between microbial diversity in ancient woods and arable fields, as well as showing that community composition continues to change even after centuries of forest cover.


## 3.2 Introduction

### *3.2.1 Soil microbial systems*

Our understanding of agricultural and woodland ecosystems above-ground is much more comprehensive than our understanding of their below-ground compartments because of the inherent challenges of studying soil biota. Improving our knowledge of biological processes in agricultural landscapes is of vital interest in food security and land-sharing based approaches to biodiversity conservation. Less managed systems, such as woodland ecosystems, are also recognised as important habitats and corridors for biodiversity by national governments across Europe who are seeking ways to mitigate climate change and biodiversity loss. In the UK woodland creation has been advocated for as a carbon-sequestering method for ecological

restoration and has been built into the strategic plans of the devolved UK governments (Forest Service, 2006; DEFRA, 2018; Welsh Government, 2018; The Scottish Government, 2019). Woodland creation is sometimes viewed as a method of improving biodiversity by creating more natural environments or may be used to mitigate the loss of biodiversity from grubbing out mature woods or Ancient Semi-Natural Woodland (ASNW). This is based upon the assumption that woodlands are better for biodiversity than the more managed agricultural systems. While evidence of high above-ground diversity in afforested land compared to agriculture exists, in general, the evidence base is concentrated on a few specific cover categories, such as conifer plantations, and select indicator taxa, such as birds (Burton *et al.*, 2018), small mammals (Moro and Gadal 2007), and ground beetles (Ng et al. 2018). Studies have rarely directly compared the biodiversity of afforested land to that of the previous cover class of the land. Below-ground biodiversity differences between land-use types are even less well understood, especially for microbial communities. However, recently analysis of the soils examined in this study and the ecotones of these sites has revealed higher enzyme activity in woodland sites relating to an increased prevalence of mycorrhizal fungi (Błonska *et al.* ,2020).

Differences exist in the reproductive rate, dormancy capacity, dispersal range, speciation rate, genetic transfer, and nutrient acquisitions pathways between organisms of the macro- and micro- scales (Nemergut *et al.*, 2013). Because of this, our understanding of how community assembly and succession occurs in the macro-world may not apply to microbes in the same ecosystems even when organisms operating at macro- and micro- scales interact often and exert great influence on one another (Balser *et al.*, 2006). Recent work is beginning to characterise how soil ecology varies across commonly occurring environmental gradients revealing how chemistry shapes microbial diversity and community composition (Uroz *et al.*, 2016), particularly pH and carbon (Rillig *et al.*, 2019), and how vegetation can exert top-down pressure on microbial communities (Mitchell *et al.*, 2010). Land-use history and human disturbance are clearly important too (Jangid *et al.*, 2011; Susyan *et al.*, 2011; de la Peña *et al.*, 2016; Turley *et al.*, 2020) with evidence that microbial communities take over a century to recover from a history of agricultural use (Fichtner *et al.*, 2014). Traditional microbial perspectives might assume that “everything is everywhere: but the environment selects” (Baas Becking, 1934; O’Malley, 2008), given the potential for microbes for long-range dispersal and the top-down pressures exerted by vegetation. Long-range dispersal capacity may not

necessarily lead to increased short-range dispersal capacity and short-range dispersal limitation in soils may be important to community assembly (Chen *et al.*, 2019). Neutral processes such as dispersal may be less limited in microbial systems (Barberán *et al.*, 2014), and biological and evolutionary response to niche pressure might be realised in comparatively short time scales (Nemergut *et al.*, 2013). Reinforcing this, recent work has demonstrated divergent trends in diversity between macro and micro scales at the national scale (George *et al.*, 2019). This has sparked interest in the relative roles of niche and neutral processes in microbial community assembly and change (Barberán *et al.*, 2014).

Differences between soil biodiversity between land-use types will result from a wide range of processes. In aboveground, woodland systems, abiotic and biotic gradients are known to influence biodiversity (Hermy & Verheyen, 2007); we have extensive accounts in the *Database for the Biological Flora of the British Isles* of the habitat preferences of hundreds of species of British flora regarding their tolerance for shade or soil acidity, fertility, and wetness. In the UK the National Vegetation Classification scheme relies on these patterns of co-occurrence and avoidance that appear to structure the British countryside (Hall *et al.*, 2012). However, it is still unclear how factors that influence above-ground diversity affect microbial communities, where a hyper-heterogeneity of ecological conditions exists even within only a few square metres of space (Ettema, 2002; Dumbrell *et al.*, 2010). Above-ground, transitions in the structure of woodland at the site scale are grouped into successional stages as trees establish and woodland structure develops. These may be groupings by description and demography, e.g. stand initiation, stem exclusion, demographic transition, and multi-aged from Frielich's model (2002), or by stand age (Brunet *et al.*, 2011). Typically, a mixed approach is taken in the literature, often using analogous terms to the following categories: newly created woodland; new wood regeneration; mature or secondary woodland; plantation; and native, old-age, remnant or ancient woodland (Susyan *et al.*, 2011). It is unclear whether the categories that seem coherent above-ground produce concomitant changes to below ground communities, or if the timing of soil microbial shifts matches traditional stages succession. Molecular comparisons of plant-Arbuscular Mycorrhizal Fungi networks in young planted (25 ya) and old-growth woodlands indicated that established woodlands are richer in AMF and have a significantly higher proportion of specialists (Bennett *et al.*, 2013), but the timescales over which these shifts take place is still unknown. The very different biology of microbial life forms

may mean that the timescales of microbial succession in woodland soils are rapid compared to large organisms, responding quickly to local soil chemistry changes and biotic interactions. Soil conditions, and perhaps microbial communities, exhibit considerable variation within medium spatial scales, i.e. metres (Dumbrell *et al.*, 2010; Chen *et al.*, 2019). The influence of these medium-scale factors, e.g. individual tree associations (Chen *et al.*, 2019) or variations in soil chemistry, are important. This may be especially true over short, ecological time scales that might be experienced during habitat restoration. Microbial communities underpin the provisioning of a suite of ecosystem services that are essential to human and environmental health; such as soil structure, fertility, and carbon sequestration via their interactions with biogeochemical cycles and the filtering, recycling or control of contaminants, waste, and pests (Barrios, 2007; Haygarth & Ritz, 2009; Dominati *et al.*, 2010; Wall *et al.*, 2015). These services are provided in numerous habitats, not least in agricultural and woodland soil systems (Sylvain & Wall, 2011; Lladó *et al.*, 2017). The previous generation of microbial assays consists of diverse methods that included single-variable indicators of soil processes; such as soil respiration; community profiling measuring a subset of microbial diversity at low resolution; such as community-level physiological profiles, denaturing gradient gel electrophoresis, or terminal restriction fragment length polymorphism assays; or count-based techniques such as cultures colony counts or spore counts (Winding *et al.*, 2005). Advances in these approaches have produced sophisticated phenotypic characterisations of microbial communities, such as Biolog's microbial assay plates. These have been able to characterise these communities at a generally low taxonomic or functional resolution, relating broad-scale community changes to differences in vegetation type (Zak *et al.*, 2003; Strickland *et al.*, 2009), land-use history (Jangid *et al.*, 2011; Creamer *et al.*, 2016), and soil properties such as nitrogen concentration (Zak *et al.*, 2003) and carbon (Bossio *et al.*, 1998). These methods have been useful and effective and are normally in broad agreement in assessing the diversity or health of soils. However, greater taxonomic precision is needed to link these processes directly to diversity. Without a clear understanding of the relationship between soil diversity, soil properties and soil function, important management decisions are being taken regarding the preservation of ecological diversity in soils without a proper understanding of their consequences.

### *3.2.2 Molecular methods in soil ecology*

Molecular methods, in cheaply producing huge datasets, are increasing our potential to answer complex ecological questions. Despite the rapid turnover of technology and methods in this area, existing techniques are ready to answer pressing theoretical and applied questions that need not wait for an illusory optimum molecular approach. Advances in molecular biology and computing have allowed us to better understand microbial systems over the last century. High-throughput sequencing platforms, such as Illumina or PacBio sequencers, can sample microbial communities with high completeness and identify specific taxonomic units individually and provide a relative measure of biomass or quantity. Studies published using earlier techniques have been limited by technological capability. The earliest methods of quantifying soil biodiversity have included culture-based techniques that exhibit high taxonomic bias (Handelsman, 2004). Later techniques sometimes involved measuring proxies of microbial activity as a measure of diversity. Phospholipid Fatty Acid Extraction (PLFA), enzyme activity, and soil respiration all provided useful measures of very specific types of microbial activity that have specific uses to quantify the rates or elucidate the mechanisms of biologically mediated soil processes. PLFAs may be used to detect the activity of different taxonomic or phenotypic groupings, such as fungi or gram positive or negative bacteria (Creamer *et al.* 2016). Enzymes might be associated with the decomposition of specific compounds such as cellulose (Schneider et al. 2012), or of the cycling of specific nutrients such as phosphorus (Stout et al. 2014). However, many have used these techniques to infer differences in soil microbial diversity, based on the assumption that these measures are consistently related to richness or diversity. Without more accurate technology, this step was necessary but contentious (Frostegård *et al.*, 2011). Genetic measures of community diversity were also produced, such as Temperature or Denaturing Gradient Gel Electrophoresis (TGGE and DGGE) and Terminal Restriction Fragment Length Polymorphism (TRFLP). These methods do not identify taxa but characterise the genetic variability of a sample, one cannot say that taxa A is present in two analysed samples, and multiple taxa may be grouped together in results, masking true diversity. Despite these limitations, these methods are useful at coarsely characterising differences between communities and are still advocated for today as a less complex alternative to next-generation sequencing techniques (De Vrieze *et al.*, 2018). Early sequencing platforms, such as Ion Torrent, did provide the ability to identify taxa to high resolution but were limited in sequencing power (Shokralla *et al.*, 2012). As technology advances, current next-generation techniques will face the same criticism. Critically, next-

generation techniques allow researchers to identify taxonomic groups present across samples, the first true molecular analogue to a traditional ecological survey. Just as with manual survey techniques, issues of taxonomic resolution, sampling bias, and detectability are present with next-generation sequencing, but these are not new problems to ecology (Remsen, 1994; Durso *et al.*, 2011; Iknayan *et al.*, 2014). Soil microbial ecology is ripe ground for re-examination with molecular tools that are allowing higher taxonomic resolution than any previous approach (Balser *et al.*, 2006).

With recent developments in molecular ecology, we can ask more questions about where abiotic, biotic, or neutral processes influence community assembly in woodland soils during afforestation and succession. By comparing the trends in soil chemistry to those of microbial biodiversity across a gradient of habitat succession we can assess the strength of the influence these properties have on community assembly. If microbial diversity differs markedly between similar woodland soils of different ages, or between woodland and pasture that have similar soil characteristics, biological or neutral processes, such as dispersal and competition may be having a strong influence on the assembly of those communities. Whereas if land-use effects can be adequately explained by soil properties without a need to consider woodland age or land use, then we will have strong evidence for the dominance of abiotic control of niche processes in microbial diversity.

### *3.2.3 Hypotheses*

In this chapter, I examine how bacterial and fungal soil biodiversity vary across a chronosequence of woodland restoration, from agricultural sites, through to young mature and ancient semi-natural woodlands. By incorporating soil chemistry and spatial information about the site, analysed in the previous chapter, I investigate the importance of these factors in community assembly. Understanding how these site characteristics, influence the diversity and functioning of belowground communities is important for planning woodland conservation and planting at a landscape scale. Prioritising site acquisition and conservation to maximise beta-diversity at the landscape scale may be informed by knowing more about the initial site characteristics or know how altering them may produce different communities over time. I take advantage of the existing “natural experiment” study system of woodlands

identified by the Woodland creation and Ecological Networks (WrEN) project. The WrEN project aims to study the effects of spatial ecological network properties; such as isolation, age, and size; on woodland creation sites (Watts *et al.*, 2016). This natural experiment approach is the only way to study woodlands with centuries difference in age within useful timescales. This exclude direct experimental approaches where potential confounding variables are controlled. However, in the centuries it would take to conduct this experiment, many land use decisions will have been made without the benefit of any research into the topic and the questions originally posed may no longer be relevant to practitioners. Multiple ecological surveys have been conducted within WrEN project sites to study how site and landscape-scale changes affect biodiversity, including work demonstrating how spatially hierarchical decisions affect bird habitat choice (Whytock, Fuentes-Montemayor, Watts, Macgregor, *et al.*, 2018), how local, but not landscape, quality affects Diptera habitat choice (Fuller *et al.*, 2018), and how bird community composition is influenced by habitat continuity (Whytock, Fuentes-Montemayor, Watts, Barbosa De Andrade, *et al.*, 2018). Recent work compared the earthworm diversity and site-scale soil properties of a selection of these sites to nearby agricultural and ASNW sites; Ashwood *et al.* (2019) demonstrated how site scale pools of soil organic carbon increase in size in reforested land as well as shifting to higher C:N ratios, rapidly shifting in the first 50 years of woodland growth. Young woodlands also had high earthworm diversity. However, the soil microbial diversity of these sites is not well understood. Błonska *et al.* (2020) examined the samples from Ashwood *et al.* for enzyme activity and soil organic matter fractions (excluding those of non-adjacent woodland and farmland. The study focussed on hydrolytic enzymes relating to carbon cycling (β-D-cellobiosidase, β-Glucosidase, β-Xylosidase) and N cycling (N-acetyl- β-Glu-cominidase). They demonstrated that woodland-agricultural ecotones transition sharply with little edge-effect and that enzyme activity associated with C transformation was higher in woodlands. This work suggested that microbial diversity differed between these land-use types, attributing the increased enzyme activity to high mycorrhizal prevalence in woodlands. However, these methods cannot demonstrate differences in biodiversity directly.

I added to the WrEN soil property data for these samples in the previous chapter and reconducted their analysis using a finer scale approach and more complex statistical techniques, more appropriate to match the data to changes in the microbiome. Specifically, I shifted the scale of the analysis from the site scale (per hectare) to the microscopic scale

(percent or per gram) and I preserved statistical power by not averaging non-independent data and instead using a mixed-effect model with random factors. By preserving statistical power I increased the confidence in the results and was able to examine woodland age categories separately that were combined in Ashwood *et al* (2019). By rescaling, my results can be interpreted in a microbiological context and applied to microbiological questions. Aside from these differences, my results generally agreed with trends identified in Ashwood *et al.* (2019). In this chapter, I considerably expand this work by assessing the microbial diversity of the soil cores taken from these sites, implementing next-generation sequencing approaches to discover how these communities change across land-use categories. The WrEN project provides an opportunity to study woodland soils in a well-documented ecosystem, in a way that could be integrated into a crucially-needed whole ecosystem woodland assessment (Baldrian, 2017). I hypothesise that 1) across a land-use gradient of disturbance, from arable farmland, to pasture, young woodland, mature woodland and then to ASNW. I expect to see differences in fungal and bacterial diversity between land-use types, not only are bacteria richer in diversity, but their functional diversity may also lead to them varying in response to site conditions to a greater extent. I may observe higher species richness and biological diversity of bacteria and fungi at woodland sites than agricultural ones, and higher richness and diversity at pasture sites than arable ones if less disturbed sites have higher diversity (Giller *et al.* 1997) and time since afforestation (opportunity for colonisation) positively correlates with diversity. Alternatively, diversity may be highest in pasture or young woodlands, where niche diversity – which is positively related to taxonomic diversity in soil environments (Giller, 1996; Lladó et al., 2018) – is high while a mixture of woodland and grassland environments exist. 2) Differences in alpha diversity by land-use will persist even when the effects of soil properties on diversity are accounted for, i.e. a minimum adequate model of alpha diversity for these samples will include significant effects of land use. 3) Ecological communities, represented in multivariate space, will differ by land use. If soil diversity continues to change following afforestation, then we expect that similarity from agricultural communities to be lower in mature woodlands than young woodlands, and highest in ASNW. In the previous chapter, I demonstrated that none of the soil properties examined for these sites differed significantly between woodlands of different age categories; however, soil properties may still influence variation within land-use categories.

## 3.3 Methods

### *3.3.1 Soil analyses*

Ashwood et al (2019) selected study sites from suitable sites within the WrEN project's available woodlands located in the English midlands and complemented these sites with nearby agricultural and ASNW sites for comparison. In October 2016 they collected ten soil samples from the top 20cm of the soil at random locations within a 400$m^2$ square at the centre of each site. They separated sampling locations by at least 5m and positioned them at least 1m from any tree. Additionally, the project transect-sampled agricultural-woodland ecotones and for all samples they collected the soil organic horizon and a soil sample at 20-40 cm depth. I did not examine these samples in this work or the previous chapter, but one may find an analysis of them in Błonska *et al.* (2020). I am focusing on the soil core samples collected from the centre of the site in the most active upper 20 cm depth of the soil. These samples were analysed by forest research to determine their pH in solution. They also measured samples' total carbon and nitrogen (Thermo/Carlo Erba, FLASH EA 1112 Series Nitrogen and Carbon analyser), nitrates (via 1M KCL extraction), and ammonium (by colourimeter analysis). Ashwood *et al.* (2019)calculated bulk density from three 100 $cm^3$ cylindrical soil cores taken at the sites. I assessed the texture of all the samples, analysed in the previous chapter. The similarity of soils in this landscape meant that there was not sufficient variation between soil types and land-use types to analyse both variables in models, so this data has not been analysed in this chapter. More details are available regarding the WrEN project in Watts *et al.* (2016), the soil sampling in Ashwood *et al.* (2019), and the ecotone sampling in Błonska *et al.* (2020).

### *3.3.2 Biodiversity data generation*

High-throughput, next-generation sequencing platforms produce large amounts of sequence data; but keeping track of hundreds of samples on a single sequencing run is non-trivial. I made use of a nested tagging approach to identify pooled samples (Kitson *et al.*, 2018). I randomly allocated samples into libraries to be processed through the workflow simultaneously. I created libraries using a 96-well plate format, with a tagged primer pair, added at the PCR amplification step, identifying the location of each sample on that plate. I added a second tagged primer combination to each sample using a low cycle number PCR, identifying the library to which it belonged. This second tag also included the adapter that enabled

sequencing on Illumina platforms. Each library contained 24 samples, in addition to an extraction negative, PCR negative, DNA positive and PCR positive. These negatives enable me to identify contamination that may have occurred during sample extraction or the initial PCR. Positive samples demonstrated the success of the PCR steps. Together the negatives and positives can be used to diagnose problems that may have arisen during the workflow. Primers targeted at fungal taxa made use of extracted *Flammulina velutipes* DNA for their DNA positive (showing the success of the first PCR) and *Lentinula edodes* for their second PCR positive. Primers targeting bacteria used DNA extracted from isolates of *Shewanella oneidensis* for their DNA positive, and *Planococcus alkanoclasticus* for their second PCR positives. I choose these positive samples as they are unlikely to occur in British soils. I amplified PCR positives separately from the other libraries but added to libraries before their first bead clean-up.

### *3.3.3 Laboratory protocols*

I extracted DNA from 0.25g of soil from each core in individual 5ml Eppendorf tubes. The extraction protocol is a modified version of the *PowerSoil protocol* (MO BIO, Carlsbad, CA, USA) and is available in full detail in the appendix. I physically disrupted the samples using a combination of 2g of crushed, acid-washed garnet and lysis solutions. These were aqueous solutions, the first contained 147 mM guanidine thiocyanate, 228 mM trisodium phosphate, 26 mM sodium chloride, 67 mM Tris HCl, and 27 mM EDTA at 9.0 pH and I added 2200µL to the garnet and sample before disruption. The second solution contained 90 mM aluminium ammonium sulphate and 1.25% (w/v) SDS, of which I added 800µL prior to disruption. I shook this mixture of sample garnet and lysis solution at 1750 RPM in a Geno/grinder 2010 for four minutes. I centrifuged the sample at 4,000 x g for 1 minute, transferred the supernatant to a 1.5ml Eppendorf tube and then centrifuged this again at 10,000 x g for 1 minute, transferring 500µL of the supernatant to a fresh tube. All following centrifuge steps were at 10,000 x g for 1 minute. I combined the supernatant with 200µL of aqueous 5M solution of ammonium acetate as a protein flocculant then I incubated this over ice for at least 10 minutes. I centrifuged the mixture and transferred the supernatant to a fresh tube. I removed additional amplification inhibitors by adding 200µL of freshly combined inhibitor flocculant solution, consisting of equal volumes of aqueous solutions of 180 mM aluminium ammonium sulphate and 204 mM calcium chloride dihydrate. I centrifuged and transferred the resulting supernatant as with the previous step. I combined the supernatant with 1568µL of 5.5 M

aqueous guanidine HCl solution to bind sample DNA to a silica spin column (*Bio Basic*, EZ-10 DNA Mini Spin Columns), passing the solution through the column over 2 or 3 successive centrifuge steps depending on the volume and discarding the flow-though. I washed remaining inhibitors from the filter with two 80% ethanol washes centrifuged through the column before eluting the purified DNA in 313µL of 1 mM tris (pH 8.0), warmed to 70°C which I pipetted onto the filter, left to elute for 2 minutes then centrifuged into a fresh tube.

I PCR amplified the purified DNA at two loci. For fungal barcoding I selected the ITS1 locus, using the ITS1F and ITS2 forward and reverse primer pair (White *et al.*, 1990; Gardes & Bruns, 1993). For bacterial barcoding, I used the 16S rRNA locus, using the 515F and 806R primer pair (Caporaso *et al.*, 2011). These primer pairs have been shown to successfully metabarcode soil bacteria and fungi on Illumina platforms in many previous studies (e.g. in (Shi *et al.*, 2020). I ordered tagged versions of these primers, with 16 forward and 12 reverse primer tags, following the protocols developed in Kitson *et al.* (2018) and allowing for 192 different potential tag combinations for individual library samples. I reserved four forward and reverse primer pairs at each locus for the two positive and two negative samples to be included in a library. This left 96 primer combinations for identifying sample locations within a library, one for each sample on a 96-well plate. I applied nested tags at the within-library and across-library level in two PCR steps. I added the first PCR tags when I amplified samples within libraries, the second PCR amplification added tags and Illumina adapter sequences to each library. The primer sequences contained heterogeneity spacers to improve sequencing performance (Fadrosh *et al.*, 2014). The PCR of ITS1 involved 35 cycles (95°C for 30s, 53°C for 30s, and 72°C for 60s) in 20µl reactions using MyFi Mix (Bioline), 2 µl of template DNA and primers (0.25 µM final concentration). The 16S PCRs involved 30 cycles (95°C for 15s, 49°C for 30s, and 72°C for 60s) in otherwise identical conditions to the ITS1 PCR reactions. To prevent cross-well contamination I sealed wells with a drop of mineral oil in 96 well 0.35 ml PCR plates and I pipetted reactants pipetted under this droplet. I sealed the plate again with a polypropylene seal. I confirmed pre-library amplification via the imaging of an electrophoresis gel. I did not normalise the failed samples or pool them with their counterparts but instead repeated and replaced them before normalisation. I normalised the concentration of samples with a clean-up using carboxylated paramagnetic beads for solid-phase reversible immobilisation (SPRI) (Hosomichi *et al.*, 2014). I followed the protocol of Jolivet and Foley (2015), which produces 1

mg $ml^{-1}$ SPRI beads stock solutions from *Sera-Mag Magnetic SpeedBeads*, 1N HCl, 5 M NaCl, 10% (v/v) Tween 20, 50% (w/v) PEG 8000, 1 M Tris base, and 0.1 M EDTA. This SPRI bead stock solution can be substituted for more expensive ferric, carboxylated bead alternatives such as AMPure XP or RNAClean XP beads (*Beckman Coulter*). I combined 12µl of normalised PCR product with 13 µl of 10mM Tris and 20µl of diluted SPRI beads at (0.1x strength - 0.1mg/ml) in a 0.8 ratio of beads to sample. This ratio size selects PCR products to bind to the bead solution while leaving unused primers and primer dimers in solution. After 3-5 minutes, I pelleted the DNA bound beads to the side of the well using a magnetic plate. I washed these pellets with two successive 60 seconds duration, 80% ethanol washes then left the beads to air dry for 15 minutes. I eluted the product in 35µl of 10mM Tris. By applying this normalisation to all samples, I reduced the variability of sample concentrations, to reduce the variability in read depth of samples during sequencing. I pooled 10µl of product from each normalised and size selected sample. I completed a second SPRI bead clean-up on 100µl of pooled pre-library with 180µl of 1 mg $ml^{-1}$ SPRI beads, eluting in 35ml Tris to concentrate the pre-library for the second PCR.

I added library level tags and Illumina adapter sequences in the second PCR. These were 12 cycle reactions (98°C for 20 s, 72°C for 30 s) in 20µl reactions using MyFi Mix (Bioline), 5 µl of normalised and size selected pre-Library and library primers (final 0.25 µM reaction concentration). I pipetted the reagents under oil and PCR amplified them in the same manner as the previous reactions, except for the use of individually capped PCR strips rather than PCR plates. This produced libraries with two sets of unique molecular tag combinations at the sample and library level. I pooled 60µl of product from three identical replicates of the second PCR for concentration. I concentrated these using 96µl of 1 mg $ml^{-1}$ SPRI beads to remove unused primers, primer-dimers, and increase DNA concentration. I pooled equimolar aliquots of this final set of libraries by locus into two duplicate libraries, each run on a separate Illumina MiSeq v3 (2x 300 bp) flow cell. Libraries were sequenced and libraries demultiplexed by NU-OMICS at the University of Northumbria at Newcastle.

### *3.3.4 Bioinformatics*

I demultiplexed samples within individual libraries on the software metaBEAT (Hahn & Lunt, 2019). I conducted all other analyses in *R* (v3.6.0). I processed the demultiplexed data using *DADA2* (Callahan, McMurdie, *et al.*, 2016), (v1.135), removing remaining primers using *cutadapt* v1.18 (Martin, 2011). DADA2 filtered and trimmed sequences based upon read quality, merged paired-end reads, removed chimeras, inferred Amplicon Sequence Variants (ASVs, see (Callahan *et al.*, 2017)) and assigned taxonomy where possible based upon the existing UNITE and SILVA databases (UNITE v 8.0, SILVA v132). The remaining analysis is adapted from the standard pipeline making use of the *DADA2* and *phyloseq* packages (Callahan, Sankaran, *et al.*, 2016). I set a prevalence filtering threshold to remove the least abundant 5% of ASVs, or to remove all ASVs with fewer than 10 reads, whichever was higher. Normalisation for differences in sequencing depth occurred in two steps. First *DESeq2 (Love et al., 2014)* normalised samples by the number of reads. At each locus, I calculated ASV richness and Shannon diversity in the package *phyloseq* (McMurdie & Holmes, 2013) which borrows functions from *vegan* (Oksanen *et al.*, 2019). Using *DESeq2*, I applied a variance stabilising transformation (VST) to the data to reduce heteroscedasticity and transform the data into an approximately normal distribution. I rlog transformed this data, as is recommended for ordination techniques (Love *et al.*, 2014). I used this transformed data to compare community NMDS ordinations by land use, described below. I produced measures of ASV richness and an NMDS analysis based upon ASV occurrence for each locus, to allow for a comparison that would be less sensitive to taxonomic bias in read abundance.

I analysed the 16S data at the level of the individual sample. I analysed the ITS1 data by pooling sample data by site due to challenges in normalising samples with DESeq2. These difficulties arose from the heterogeneity of sample taxonomic composition that led to many ASVs being absent from most samples. Important gradients present across more common ASVs were obscured by the very large number of rare ASVs present in few samples. This produced incoherent multivariate comparison of sample composition and so analysis of the full ASV dataset was abandoned in favour of analysis of the species data. These difficulties also led to challenges implementing VST and rlog abundance transformations on the data, and so to calculate VST abundance metrics I filtered ASVs to include only those that had been taxonomically assigned to the species level. This subset included 11,144 ASVs of the 28,886

total ASVs that passed the quality filtering described above. Without the noise generated by extremely rare species, multivariate gradients became identifiable.

### *3.3.5 Statistical analysis*

*Hypothesis 1 - Modelling the effect of land-use differences prima facie*

To test for an overall effect of land-use on alpha diversity, I modelled the response of ASV richness and Shannon diversity to changes in land-use for both loci. For the bacterial 16S data, I modelled this for individual samples, using linear mixed effect models from the package *nlme* (Pinheiro *et al.*, 2019) with land-use as a fixed effect and site as a random effect. For fungal ITS1 data aggregated by site, I used linear models generated in the *stats* package from base *R* with land-use as the only explanatory variable.

*Hypothesis 2 - Modelling an effect of land-use after accounting for the effect of soil chemistry*

To test for an effect of land-use after accounting for environmental variables (Hypothesis 2), I fitted global models of the response of ASV richness and Shannon diversity to land-use and soil properties. I included the same soil properties included in the global models as in chapter 2: pH, organic carbon concentration, C:N ratio, ammonium concentration, nitrate concentration, bulk density, and spatial information. I describe the method for calculating a measure of spatial autocorrelation described in the previous chapter. Briefly, I supplied spatial coordinates to the *dbmem* function in the package *adespatial (Dray et al., 2020)*, creating dbMEM eigenfunctions, similar to a PCA of spatial coordinates. I chose eigenfunctions to include in models using forward selection. I excluded total nitrogen as it correlates strongly with organic carbon across these soils. For bacterial 16S data, examined at the level of the individual samples, I fitted global mixed effect models with the site as a random variable. Soil pH appeared to be important in shaping bacterial communities, despite similarities in soil pH across land-use types so I included an interaction between pH and land-use in bacterial models. The complete array of interaction effects saturated model AIC, so I included only the pH and land-use interaction in the global model. For the fungal analysis aggregated to sites, I excluded this interaction due to the low number of replicates and fitted linear models by default as pseudoreplicates at each site had already been aggregated. I refined a minimal

model from these global models using backwards refinement to identify the most parsimonious explanation of the biological data. Linear models were preferred to equivalent mixed models, as were models without interaction terms. I only kept model terms if their removal caused a significant increase in deviance, i.e. significant terms could be removed from a model if a simpler model did not significantly differ in fit as assessed with the *anova* function. Post hoc comparisons would only be undertaken for the main land-use effects, not interaction effects, as this is the study focus and keeps the number of multiple testing corrections low.

*Hypothesis 3 - Modelling differences in community structure*

To test for an effect of land-use on community composition (Hypothesis 3), I compared the distribution of land-use types across community NMDS ordinations based upon occurrence and VST transformed abundance. I produced ordination for individual sample level bacterial communities and site-level fungal communities. I created two-dimensional NMDS plots from occurrence community matrices or VST transformed community read abundance matrices using the function *metaMDS* in *phyloseq* which inherits methods from *vegan*, with ten random starts and without auto-transformation. Permanova testing was unable to discern differences between land-use types at the unaggregated sample level due to differences in dispersion between land-use types, see the supplementary information for the previous chapter. Instead, I produced minimal models explaining the response of community composition along the NMDS axis most associated with community composition differences across land use. The method used to produce minimal models of community change was identical to that use to model changes in alpha diversity.

### *3.3.6 Correction for multiple testing*

I applied a p-value adjustment to all statistical results using the core *R* package *stats* with the Benjamini and Hochberg method (R Core Team, 2019). In each case I corrected for 60 tests, one for each p-value provided (20 tests), and one for each land-use contrast in post-hoc testing (four contrasts with 10 combinations). This is a conservative method. I only conducted post hoc testing on the primary effect of land-use as this was the subject of the study.

## 3.4 Results

### *3.4.1 Sequencing results*

The 16S samples produced 14.8 million reads, 10.5 million remained after filtering, denoising, merging paired ends, and chimera removal. The removal of short reads, positive and negative samples, and rare ASVs led to a final dataset of 9,870,382 reads. DADA2 assigned these to 15,558 bacterial ASVs (9,780,175 reads across all samples - 99.1%), 46 archaea (88,933 reads across 191 of the samples - 0.9% of reads), seven eukaryotic ASVs (194 reads - 0.002% of reads across 2 samples), and 10 unidentified ASVs (1080 reads present across 9 samples - 0.01% of reads). After filtering, denoising, and merging, and chimera removal, the ITS1 samples produced 6.6 million reads. Of these 6,244,609 reads remained following the removal of positives and negatives and the filtering of low-quality ASVs. These were assigned to 28,886 ASVs all of which were fungal. Total read information entering and exiting the DADA2 pipeline can be found in supplementary Table S5.1. Read data for negatives during the DADA2 pipeline can be found in the supplementary materials Table S5.2.

### *3.4.2 Heterogeneity of ASV composition at the sample and site Level*

I produced qualitatively similar ordination plots and alpha diversity statistics from both sample level analysis and site aggregated analysis for the 16S data, although within-site variation was still noticeable, with richness in aggregated samples on average having 3 times higher richness than individual samples. This was very homogeneous compared to the ITS data, where average aggregated richness was nine times higher than the average sample richness, indicating an extremely low overlap in community composition between any of the ten soil samples from the same site. This may be due to the challenges inherent to producing paired-end ITS1 reads, such as the sheer heterogeneity of the locus's length that leads to systematic under sampling of certain fungal taxa (Schmidt et al. 2013; De Filippis et al. 2017). The degree to which it indicates actual ecological heterogeneity is not easy to disentangle.

### *3.4.3 Hypothesis 1 - Land-use effects on alpha diversity*

Hypothesis 1 poses the applied question: does land-use category predict microbial alpha diversity as a product of abiotic succession and concurrent biotic interactions. Mixed effect models with sampling site as a random factor were the most parsimonious model structure

for modelling the response of bacterial ASV richness and Shannon diversity. Richness did not significantly differ across land-use categories for richness or Shannon diversity; for either bacteria analysed at the sample level, or fungi analysed at the site level (Figure 3.1, Table 3.1). Models including site as a random variable had significantly improved fit during model selection, and clustering of sites is apparent in (Figure 3.1 A and B). Although apparent differences in fungal diversity between land-use types may appear discernible (Figure 3.1 C and D), there is high uncertainty around these site scale estimates which is reflected in the non-significant model outputs (Table 3.1).

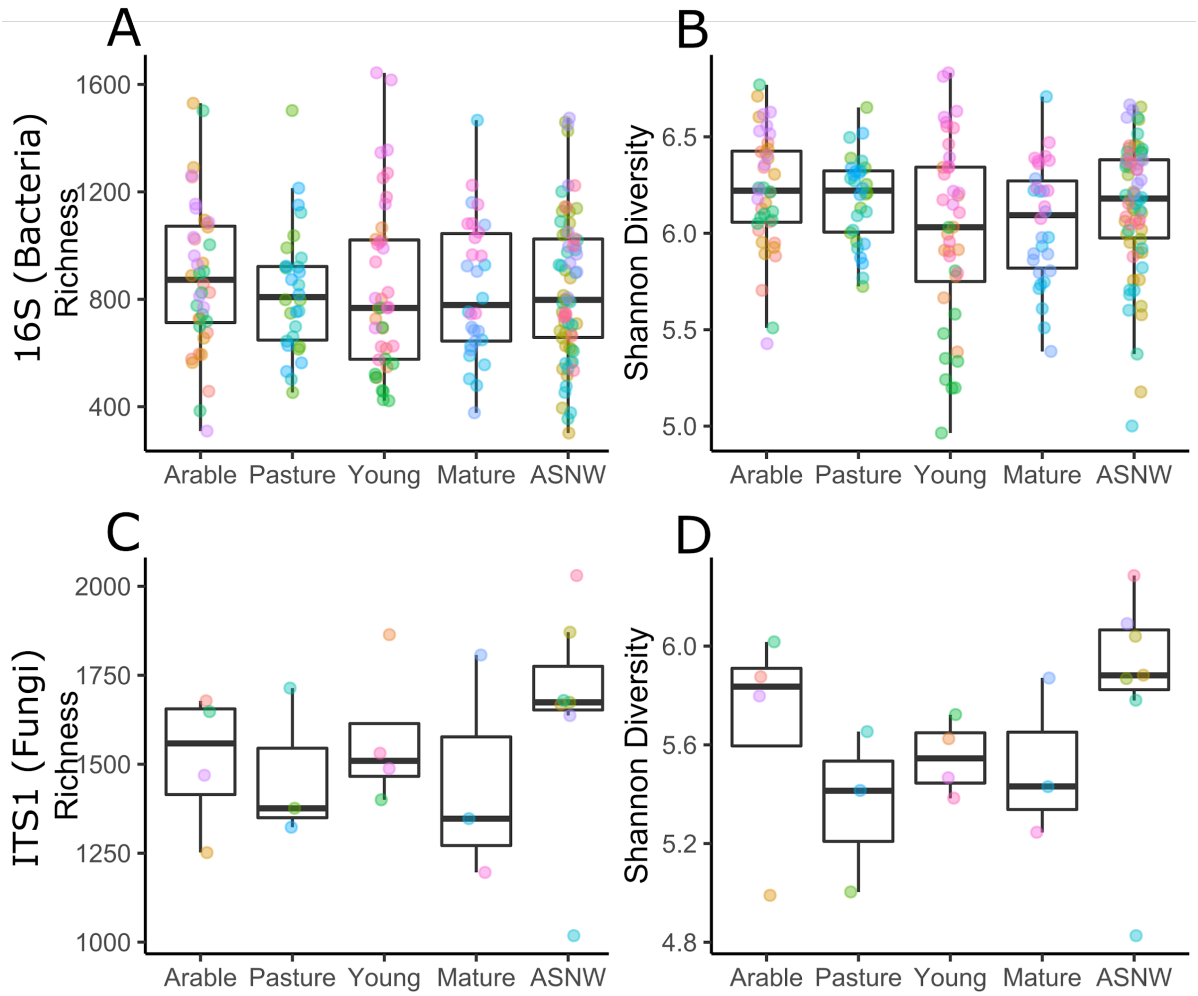


Figure 3.1 Alpha diversity by land-use for WrEN project soils collected in England from two arable and pasture farmland and young, mature and Ancient Semi-Natural Woodland (ASNW). Bacterial samples (A and B) and site-pooled fungal samples (C and D) are coloured by site. Sample richness (A and C) and Shannon diversity (B and D) calculated from ASV abundance normalised by read depth. Sites are anonymised at request.

Table 3.1 Coefficients and significance of effects in alpha diversity models. Adjusted p-values generated the Benjamini & Hochberg adjustment method (correcting for 60 tests). Here bolding indicates a p-value < 0.05.

| Response Variable | Coefficient (95% CIs) | d.f. | F-value | P-value | Adjusted p-value |
|---|---|---|---|---|---|
| ***Bacterial (16S) effect of land-use only (mixed-effect models)*** | | | | | |
| Richness | Figure 3.1 A | (4, 16) | 0.097 | 0.98 | 1.000 |
| Shannon diversity | Figure 3.1 B | (4, 16) | 0.557 | 0.70 | 1.000 |
| ***Fungal (ITS1) effect of land-use only (linear models)*** | | | | | |
| Richness | Figure 3.1 C | (4, 16) | 0.466 | 0.76 | 1.000 |
| Shannon diversity | Figure 3.1 D | (4, 16) | 0.908 | 0.48 | 1.000 |
| ***Minimal Model Bacterial (16S)*** | | | | | |
| *Richness (linear model)* | | | | | |
| Land use | Figure 3.2 A | (4,199) | 0.348 | 0.85 | 1.000 |
| pH | 106 (-7 to 218) | (1, 199) | 49.57 | **<0.001** | **<0.001** |
| Land-use:pH | Figure 3.2 B | (4, 199) | 3.29 | **0.012** | 0.074 |
| *Shannon Diversity (mixed-effect model)* | | | | | |
| Land use | Figure 3.2 C | (4, 16) | 2.33 | 0.10 | 0.429 |
| pH | 0.172 (0.02 to 0.32) | (1, 183) | 71.14 | **<0.001** | **<0.001** |
| Land-use:pH | Figure 3.2 D | (4, 183) | 4.40 | **0.002** | **0.020** |
| ***Minimal Model Fungal (ITS1)*** | | | | | |
| *Richness (linear model)* | | | | | |
| Organic Carbon | 82 (26 to 139) | (1,19) | 8.39 | **0.009** | 0.062 |
| *Shannon Diversity (mixed-effect model)* | | | | | |
| Organic Carbon | 0.09 (-0.0 to 0.19) | (1,19) | 3.16 | 0.091 | 0.422 |

### *3.4.4 Hypothesis 2 - Minimal models of alpha diversity*

Hypothesis 2 asks whether biotic dynamics relating to succession and land-use change alone influence microbial alpha diversity after accounting for the abiotic changes occurring over the same period. I produced similar minimum adequate models explaining bacterial ASV richness and Shannon diversity. For bacterial richness models, fit did not improve by including site as a random effect, so I used a linear model. Bacterial Shannon diversity models including site as a random effect had meaningfully improved fit over those without. This led me to model Shannon diversity response using a mixed-effect model including a random site effect. For both measures of bacterial alpha diversity, the explanatory variables remaining after model refinement were land use, pH, and an interaction between the two (Table 3.1). Models containing land-use had a significantly better fit during refinement, but the terms themselves were not significant (Table 3.1, Figure 3.2 A and C). In agreement with previous work (Griffiths *et al*., 2011), higher pH, more neutral soils had higher richness ($F_{(1,199)} = 49.57$, adjusted $p < 0.001$) and Shannon diversity ($F_{(1,183)} = 71.14$, adjusted $p < 0.001$) than acidic soils. But there was a land-use interaction with pH with both measures ($p < 0.05$), but after p-value adjustment, the interaction only remained significant for Shannon Diversity measures ($F_{(4,183)} = 4.40$, $p_{adj} = 0.020$), not richness measures. I did not conduct post-hoc testing to determine which land-use contrasts significantly differed in their land-use to pH interaction, as it was not the focus of this investigation and post-hoc testing of mixed-effect models remains controversial (Bates, 2006; Bates *et al.*, 2015). This reduced the number of statistical tests for which we needed to apply a p-value correction. However, the land use-pH interaction coefficients are displayed in Figure 3.2 B and D, indicating that the effect of pH was reversed or uncertain in pasture sites.

The minimal model needed to explain fungal ASV richness was a linear model indicating that higher organic carbon concentrations in soils were associated with higher fungal richness (Table 3.1). However, the significance of this result was lost after p-value adjustment ($F_{(1,19)} = 8.39$, unadjusted $p = 0.009$, adjusted $p = 0.062$). Shannon diversity was not sufficiently explained by any terms - the minimal model was a null model. The final term removed in backwards selection was organic carbon concentration, so I have presented that model indicating its non-significance (Table 3.1).

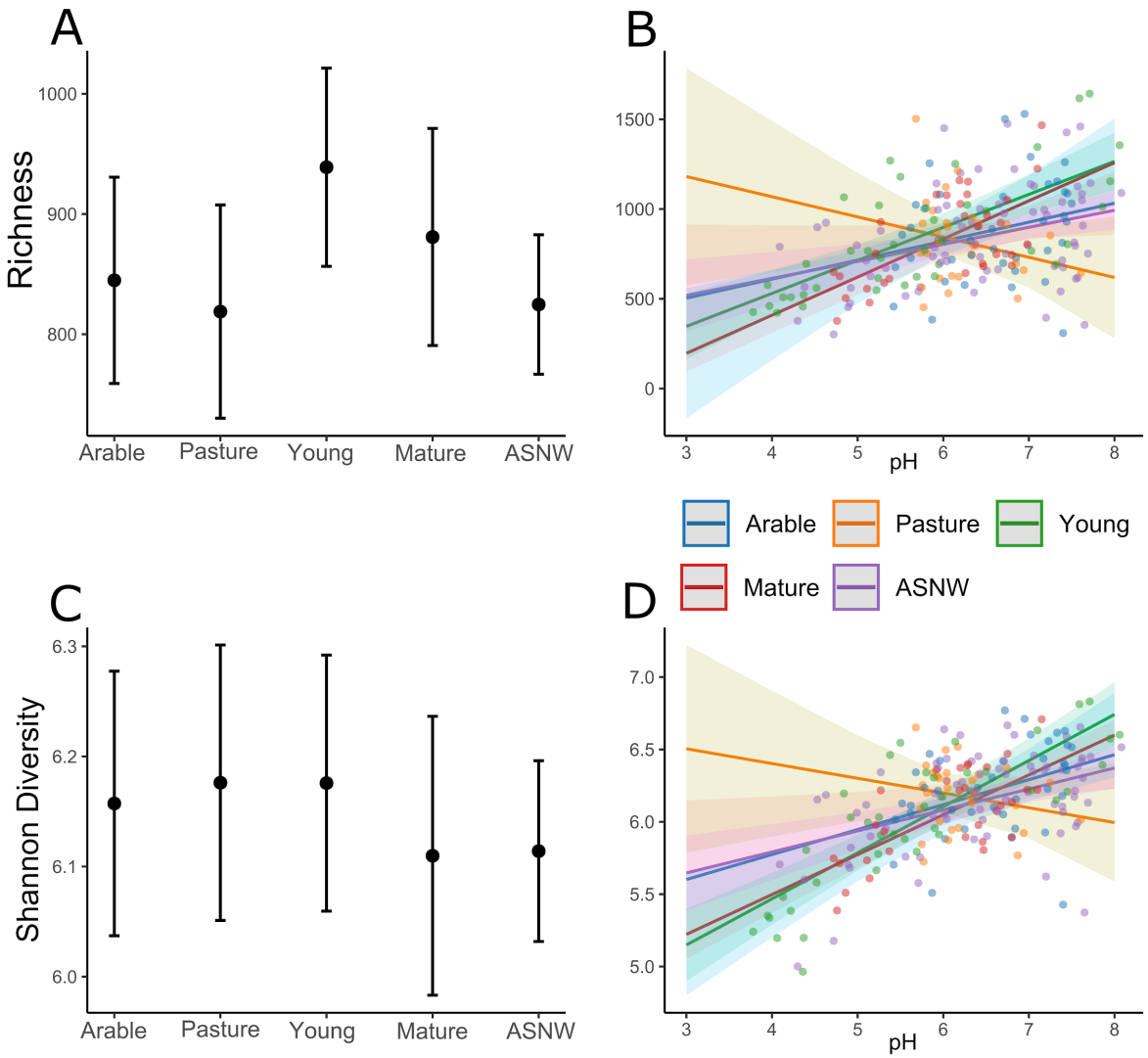


Figure 3.2. Land-use modified effects from minimal models of bacterial alpha diversity. Bacterial richness (A) and Shannon diversity (C) do not significantly vary by land use. But for both measures of diversity, pastures significantly differed from other land-use types in its response to pH (B and D). In both plot types, error bars and shaded envelopes represent 95% confidence intervals around the coefficient.

### *3.4.5 Hypothesis 3 - Differences in beta-diversity by land-use*

Bacterial and fungal ordinations indicated strong differentiation of microbial communities by land-use (Figure 3.3). Differences are apparent at a coarse scale, comparing woodlands to agricultural land, but also between woodland and agricultural types. In the bacterial communities, this spread of community differences by land-use category was best captured across the second NMDS axis in both the occurrence (Figure 3.3 A) and VST transformed abundance (Figure 3.3 B) ordinations. For the fungal data, the first NMDS axis best

demonstrated differentiation by land use, somewhat obfuscated in the fungal ASV occurrence ordination (Figure 3.3 C) but more clearly in the VST transformed species-only abundance ordination (Figure 3.3 D). Although not presented, I observed additional differentiation by land-use type along the second axis of the fungal VST transformed species ordination after accounting for the impacts of pH and C:N ratio.

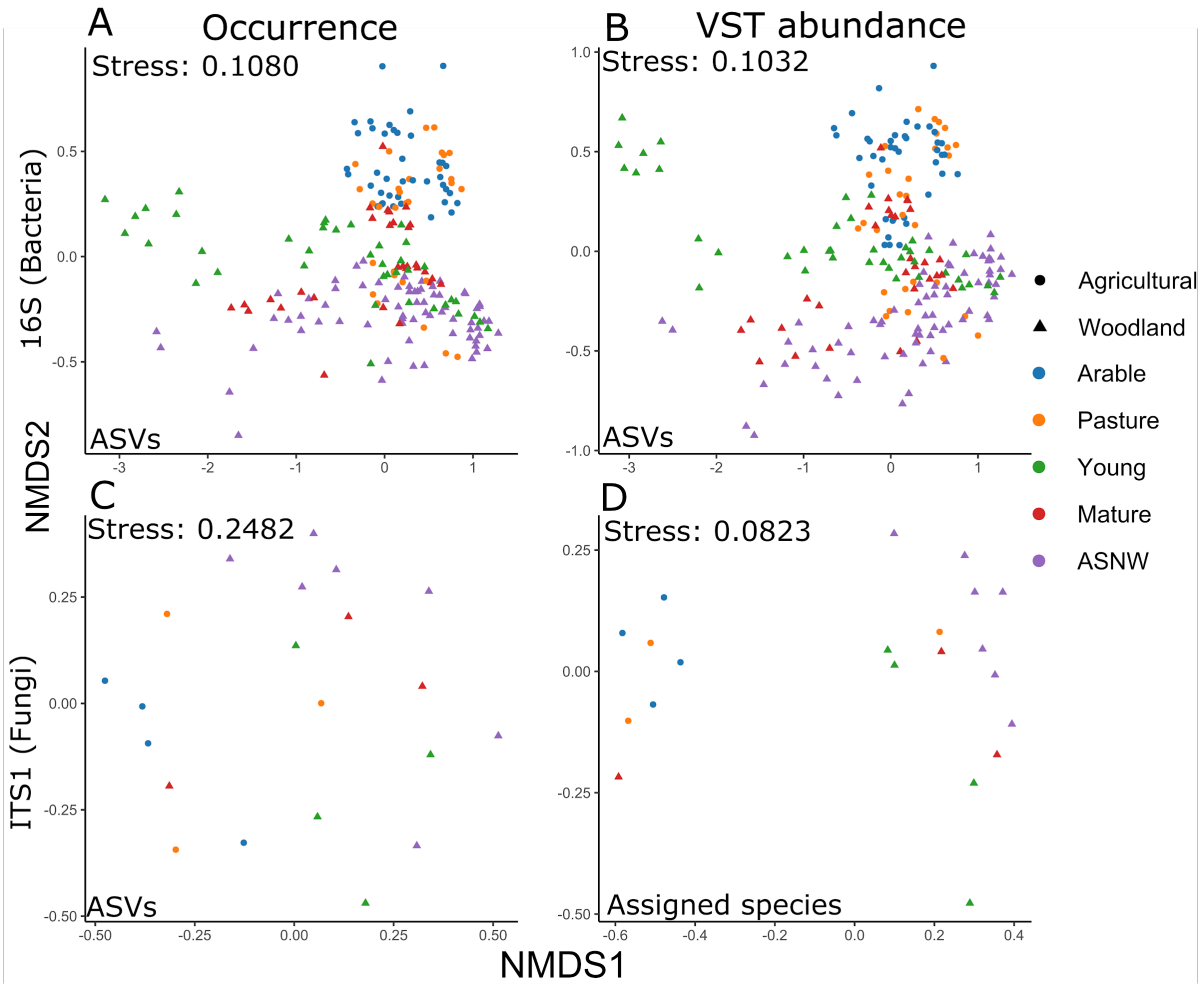


Figure 3.3 Bray-Curtis community NMDS ordinations of ASV occurrence (A and C), transformed ASV abundance (B) and transformed species abundance (D). Bacterial samples (A and B) are ordinated individually, but fungal samples are first pooled by site (C and D). Land-use is indicated by point colour, with circles representing all agricultural land-use types, and triangles representing all woodland land-use types. The direction and units of NMDS axes are arbitrary.

Communities of different land-use types significantly differed in their location along NMDS2 in bacterial models (Table 3.2, Figure 3.4 A and B), in both occurrence ordinations ($F_{(4, 16)}$ = 14.6, adjusted p = 0.0006) and VST abundance ordinations ($F_{(4, 16)}$ = 8.87, adjusted p = 0.008).

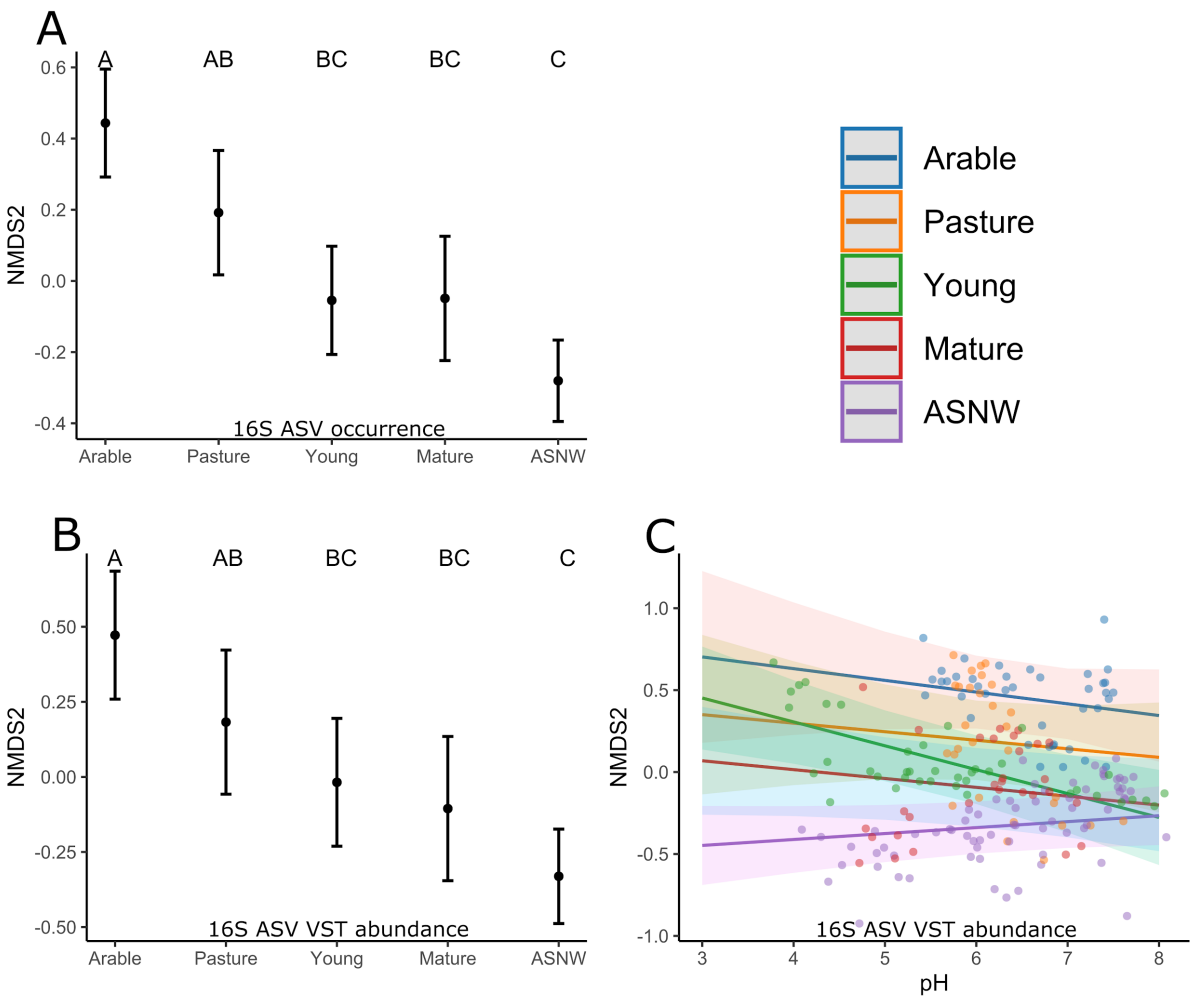


Figure 3.4 Effects remaining in the minimal adequate mixed-effect model of community similarity along NMDS2 the axis most associated with differences across land-use types. Model effects of ordinations of ASV occurrence (A) and transformed ASV abundance (B & C) plotted separately. I did not compare factor levels of mixed effect models with a p-value based approach due to known issues surrounding such methods. Instead, shared lettering indicates overlapping CIs. Error bars and shaded envelopes represent 95% confidence intervals around the coefficient.

Additionally, pH significantly influenced community position along NMDS2 in ordinations based upon bacterial occurrence ($F_{(1, 187)} = 7.29$, $p = 0.008$, adjusted $p = 0.057$), though the significance of this effect disappeared after p-value adjustment. For fungal ordinations based on VST abundance, pH alone did not significantly influence diversity, but I identified a significant interaction of pH and land-use prior to p-value adjustment ($F_{(4, 183)} = 3.20$, $p = 0.014$, adjusted $p = 0.078$) but lost significance after p-value adjustment (Table 3.2, Figure 3.4 C). In

both cases, higher, more neutral pH soils generally had communities shifted negatively along the axis, the same direction as the effect of woodland land-use types. In ancient woodlands, the shift of communities along the similarity axis was reversed. In these samples higher pH was associated with a generally more positive shift along the second NMDS axis, increasing in community similarity to agricultural sites.

Table 3.2 Coefficients and significance of effects in minimal NMDS axis models. Adjusted p-values generated the Benjamini & Hochberg adjustment method (correcting for 60 tests). Here bolding indicates a p-value < 0.05. Axes directions and units are arbitrary.

| Response Variable | Coefficient (95% CI) | d.f. | F-value | P-value | Adjusted p-value |
|---|---|---|---|---|---|
| Bacterial (16S) NMDS2 minimal models<br>ASV Occurrence (mixed-effect models) | | | | | |
| Land use | Figure 3.4 A | (4, 16) | 14.6 | **<.0001** | **0.0006** |
| pH | -0.045 (-0.08 to -0.01) | (1, 187) | 7.29 | **0.008** | 0.057 |
| VST ASV abundance (mixed-effect models) | | | | | |
| Land-use | Figure 3.4 B | (4, 16) | 8.87 | **0.0006** | **0.008** |
| pH | -0.071 (-0.20 to 0.06) | (1, 183) | 2.54 | 0.113 | 0.451 |
| Land-use:pH | Figure 3.4 C | (4, 183) | 3.20 | 0.014 | 0.078 |
| Fungal (ITS1) NMDS1 minimal models<br>ASV Occurrence (linear models) | | | | | |
| Land-use | Figure 3.4 D | (4, 15) | 5.85 | **0.005** | **0.042** |
| Bulk Density | -0.869 (-1.63 to -0.11) | (1, 15) | 5.98 | **0.027** | 0.137 |
| Species VST abundance (linear models) | | | | | |
| Land-use | Figure 3.4 E | (4, 16) | 8.04 | **0.0009** | **0.011** |

Land-use significantly influences the location of fungal communities along NMDS1 in ordination based upon both ASV occurrence ($F_{(4, 15)} = 5.85$, $p = 0.005$, adjusted $p = 0.042$) and species VST abundance ($F_{(4, 16)} = 8.04$, $p = 0.0009$, adjusted $p = 0.011$). The ASV richness ordination model also included a significant effect of bulk density that was lost after p-value adjustment ($F_{(1, 15)} = 5.98$, $p = 0.027$, adjusted $p = 0.137$). Higher soil bulk density shifted site communities negatively down the first NMDS axis (Table 3.2), in the same direction as the

effect of agricultural land. Meaning that more compacted woodland soils were generally more similar in community composition to agricultural communities.

Consistent qualitative trends of community change across land-use types were apparent in both bacterial and fungal selected community ordination axes (Figure 3.4). The models of bacterial community differences are mixed effect models, and I have discussed the inapplicability of post-hoc tests on mixed models in my previous chapter, preferring comparisons based upon 95% confidence intervals as indicated in the package literature (Pinheiro *et al.*, 2019). The ITS1 models presented are linear models and so post-hoc testing can be applied to this data. For occurrence-based fungal ASV data, p-value adjustment within the post-hoc test renders all land-use non-significant, despite identifying a significant overall effect of land use. For VST abundance ordinations of fungal species data, the post-hoc tests agree with the 95% confidence intervals, indicating significant differences between the locations of communities along NMDS1. Contrasts between ASNW and both agriculture land-use categories are significant, as well as contrasts between young woodland communities and arable communities, with p-values less than 0.05.

For mixed effect models, we must compare land-use categories by their 95% confidence intervals. In the axes associated most with community differences across land-use types, arable farmland and ASNW were positioned at opposite ends of a spectrum of community similarity (Figure 3.4 and Figure 3.5). These communities differed in 95% confidence intervals on the predicted effect coefficient for both bacterial and fungal communities, regardless of the method of ordination. Although post-hoc testing cannot identify which land-use types significantly differ along NMDS1 in ordinations of fungal ASV communities, the non-overlapping confidence intervals of only arable and ASNW communities (Figure 3.5A) suggests that this is the source of the overall significant effect of land use. Pasture, young woodland, and mature woodland never differed from one another along these axes for either bacterial or fungal communities, across all ordination methods. Neither did communities of different farmland land uses - arable or pasture - or different woodland communities - young woodland, mature woodland or ASNW. However, pasture communities significantly differed from ASNW communities in bacterial community composition (based on both occurrence and transformed

abundance-based metrics, Figure 3.4 A and B) and fungal communities when assessed using transformed abundance (Figure 3.5B) of identified species (but not when ordinated based upon ASV occurrence). Arable bacterial communities significantly differed from both young and mature woodland communities (Figure 3.4 A and B) in addition to previously noted differences from ASNW communities. Arable fungal communities differed from those of young woodlands when ordinated according to the transformed abundance of assigned fungal species (Figure 3.5B). The difference between occurrence and transformed abundance measures of community similarity is more influenced by the reduction of taxa included in the analysis than the transformation of the abundance data (Appendix Figure S3.1).

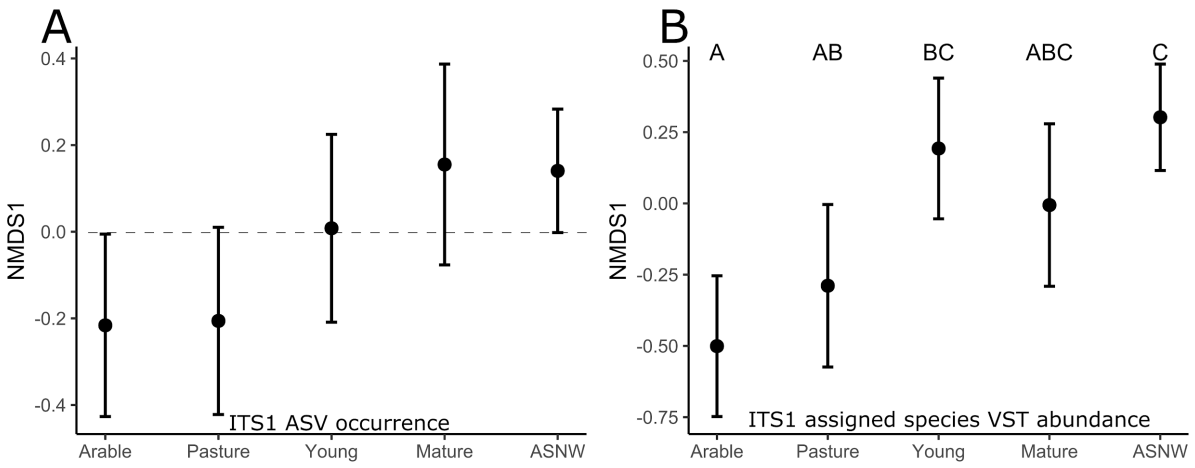


Figure 3.5 The effect coefficients of land-use on fungal community similarity along NMDS1 the axis most associated with differences across land-use types. Linear models indicated a significant land-use effect. Post hoc Tukey analysis did not have enough statistical power to detect which contrasts were significantly different for ordinations based upon ASV occurrence (A), but I detected significant differences for VST abundance of fungal species (B) and are indicated by lettering. Error bars represent 95% confidence intervals around the coefficient, the dotted line indicates that arable and ASNW CIs do not overlap for ASV occurrence ordinations.

## 3.5 Discussion

### *3.5.1 Soil communities vary across a wooded agricultural landscape*

Soil communities vary in composition, but not simplistic alpha diversity measures, across land-use categories in a gradient of afforestation and semi-natural woodland succession. I found insufficient evidence to support hypothesis 1; that significant differences in richness or Shannon diversity would be apparent between individual farmland and woodland land-use categories. While testing hypothesis 2, I did not detect an effect of land-use on alpha diversity after accounting for differences in soil chemistry between samples. Furthermore, I did not detect any significant predictor of fungal Shannon diversity across my samples. I had hypothesised that pasture sites would be more rich and diverse than arable sites and that woodlands that had been forested for longer would increase in richness and diversity either due to community assembly processes or due to increasingly hospitable soil conditions for woodland specialists. Molecular analysis of the soil diversity did not detect significant differences between the richness or Shannon diversity of any land-use types for both bacteria and fungi. Studies of soil microbial alpha diversity between land-use types remain uncommon, with a larger proportion of the literature focusing on PLFA analysis that targets functional subsets of communities. These studies have demonstrated the rapid change that occurs in afforested lands after only a few decades of tree growth. One study that took a similar approach to assessing bacterial diversity failed to detect differences in woodland soil alpha diversity but revealed differences in community composition of functional groups via PLFA analysis (Jangid *et al.*, 2011). But another that examined changes in the abundances of bacterial clades separately between sites, successfully identified changes in bacterial community composition in the first 8 years of reforestation, particularly in the phyla Acidobacteria and Firmicutes (Gellie *et al.*, 2017). In my analysis, bacterial community richness and Shannon diversity were remarkably consistent between land-use types and varied far more across sites within a single land-use category than between them. Fungal diversity appeared to be more variable between land-use types, but this difference was not significant and is likely to be affected by pooling the sample data to a small number of sites within each land-use replicate. Failing to detect an overall difference in bacterial or fungal richness is unexpected but not unprecedented.

Despite the remarkable similarity in the taxonomic richness and Shannon diversity of sites, the composition of communities differed strongly and significantly by land use, as indicated by hypothesis 3. I hypothesised that young woodland communities would differentiate from agricultural communities and would become even less similar in older woodlands. I evaluated significant differences using post-hoc testing for linear models and 95% CI around predicted coefficients for mixed-models. In all measures except for communities based upon fungal occurrence, ancient woodland differed from arable farmland significantly, for both bacteria and fungi. Although post-hoc testing failed to detect which land-use contrasts of fungal communities ordinated based upon ASV occurrence significantly differed from one another, the largest difference was between ASNW and arable farmland. Ancient woodland also differed from pasture for both methods of bacterial community construction, and for VST transformed abundance-based communities of assigned fungal species. Woodlands did not significantly differ from one another, neither did arable or pasture land. However, bacterial communities at young and mature woodland differed significantly from those of arable and pasture, whether constructed from ASV occurrence or VST read abundance. Finally, fungal communities constructed from VST fungal species abundance in young, but not mature woodlands significantly differed from arable and pasture communities, but not when based upon ASV occurrence alone. Previous studies have found rapid changes to afforested soil diversity and microbial composition in woodlands younger than any of those included in this study. Some studies have shown a dip in microbial diversity and activity immediately following planting, such as in orchards planted on rice paddies which dropped in microbial respiration after three years of growth (Wu *et al.*, 2020). However, older sites in the study showed a positive relationship between stand age and respiration rates in orchards up to 30 years old. In reforested riparian soils total PLFA concentrations decreased in the initial ten years after planting, then increased, returning to similar levels as those of remnant woodland patches after 23 years although no significant difference existed between age categories and no unambiguous trends were found between stand age and PLFA abundance or decomposition enzyme activity (Mackay *et al.*, 2016). Bacterial and fungal communities, assessed via metabarcoding rapidly differentiated from farmland communities after only 10 years, and continued to change over the next 20 years (Jiao *et al.*, 2018). Woodland soil communities can change rapidly following reforestation, approaching similar communities to remnant secondary woodland in ecological time scales as was shown in one study on Texan groves (Creamer *et al.*, 2016). Soil communities in these recovering groves on grassland in Texas, USA

showed rapid differentiation of bacterial PLFAs between wooded groves and the surrounding grassland matrix (Creamer *et al.*, 2016). In the same study, microbial biomass rapidly increased in the first half-century of grove growth and slowed but continued to increase in groves aged up to 140 years old. Across these studies, shifts in microbial communities have been indicated indirectly via changes in PLFA quantities and in microbial activity, such as respiration, and bacterial to fungal ratios. These measures strongly suggest community compositional change (Frostegård *et al.*, 2011) or changes in the activity of certain biological compartments that may still yield insights into community composition (Tiedje *et al.*, 1999; Barrios *et al.* 2007). These studies agree with our results, suggesting that the microbial communities of woodlands aged 50-150 years old are still changing, shifting away from those of agricultural landscapes. But the largest change has already occurred, and these woodlands often do not significantly differ from one another.

Studies that address long term differences between microbial communities of established, planted woodlands, and ancient woodlands are less common, possibly because of lack of suitable sites or confusion around terminology. Those that explicitly compare secondary woodland communities to those of ancient woodland, or woodland old enough to be classified as such indicate differences between them. Secondary woodland with a non-wooded history is either on a different successional trajectory than ancient woods, or it takes longer than a century to return to "natural" levels of microbial biomass and enzyme activity (Fichtner *et al.*, 2014). I identified a single study that compared arable farmland, young, mature, and ancient woodland (but not pasture) in a single study which indicated that microbial biomass increased rapidly following reforestation and slowed, but continued to increase as older woodlands aged, in agreement with the rest of the literature presented (Susyan *et al.*, 2011). They also detected decreases in the bacteria-fungi ratio of soils between young and mature woodland, with mature woodland having a higher proportion of fungi, but could not establish any general trends over long-term succession. Before my study, proxies for diversity, and functional diversity, such as respiration, biomass, and PLFA concentration have been employed to suggest short- and long-term microbial responses to afforestation. Studies that made use of biomass or activity measures do provide some insight into community composition, but changes in these measures do not necessarily indicate microbial community change. Most have only considered temporally short windows of woodland change, with inadequate tools

for assessing biodiversity. Of the existing studies, those with the best temporal extent have employed the least specific tools for assessing diversity changes and *vice versa*. No studies have explicitly included comparisons of long timescales with both arable and pasture farmland, and so none could demonstrate the intermediate composition of pasture microbial communities, sitting between arable farmland and ancient woodland. To my knowledge, no other genetic-based study of woodland creation and ageing has sequenced DNA from both bacteria and fungi. My study fills gaps in the literature, bringing next-generation methodologies to long-standing questions about the short-term and long-term effects of woodland creation.

The minimal models of alpha diversity and community composition highlighted the role of pH in structuring bacterial communities. Organic carbon and bulk density appeared important in structuring fungal communities. Nitrogen and total carbon correlated highly with organic carbon and so were removed from the models; however, they may still be large contributors to fungal community assemblage. The selection of these variables agrees with the previous literature. Previous work has highlighted the influence of pH in shaping British bacterial communities generally, with more acidic soils generally possessing lower bacterial diversity (Griffiths *et al.*, 2011; Kuramae *et al.*, 2012). The effects of carbon on fungal diversity are complex due to correlations with other soil properties such as nitrogen and C:N ratio, and because of the feedback of microbial diversity on soil nutrients. Fungi certainly change in response to the quality of carbon input (Hanson *et al.*, 2008) and soil carbon concentration (Thoms *et al.*, 2010). Waldrop *et al.* (2006) linked fungal diversity increase to high nutrient inputs. Other work in woodland systems has linked bacterial and fungal PLFAs and decomposition related enzyme activity to soil organic carbon (Mackay *et al.*, 2016). Bulk density is far less commonly linked to fungal diversity, but compaction has been linked to both a reduction in fungal richness (Rosas-Medina *et al.*, 2020) and its increase (George *et al.*, 2019). I did not find a directional effect on richness but did find it to alter the composition of communities. Infrequent reporting of bulk density effects may be because it is rarely tested for or may rarely contribute to diversity effects. In this work, the significance of bulk density as a predictor of fungal diversity disappeared after correction for multiple testing, so we cannot have high confidence in this result. Multiple factors have been shown to influence soil microbial community assemblage. Reviews of the literature across large spatial scales have

stressed the importance of pH, organic carbon, C:N ratio, redox state, soil moisture, nitrogen, phosphorus, soil texture, vegetation, topography, and geographic location in influencing bacterial communities (Griffiths *et al.*, 2011; Fierer, 2017). Fungal community assemblage also appears associated with soil C:N ratio, organic matter, pH, bulk density, soil moisture, and geography (George *et al.*, 2019). Effects for most of these variables did not make it into our minimal models. It is possible that I adopted too severe of a model selection criterion, or that these effects operate and interact across larger spatial scales than the extent of our limited sampling.

Additionally, I hypothesised that, as woodlands age, the importance of abiotic factors in shaping community structure would decrease, in favour of neutral or biotic processes. Evidence for this hypothesis may have arisen from the inclusion of additional interaction terms in the model construction, or from additional analyses of the relative importance of abiotic factors in pairwise analysis of community ordinations by land use. This hypothesis is testable, and I have some evidence indicating that this would be a fruitful area of research, in that the effect of pH on community structure in ASNW differs in direction from the other land-use categories. If interaction terms in similar models had shown that ASNW soil community structure responds differently to abiotic factors than younger woodlands or farmland I believe that would be of great interest. Or if pairwise CCA analysis of community structure and of abiotic variables indicated that the effect size of environmental variables on community structure significantly decreased across young woodland to ASNW and mature woodland to ASNW comparisons, that would be compelling evidence of the hypothesis. Currently, there is an absence of evidence that the importance of abiotic factors in shaping ecological communities decreases in ASNW.

Finally, I aimed to determine whether bacterial and fungal communities were more variable within or across sites of the same land use. Fungal samples were clearly highly heterogeneous within sites. Mean richness of fungal site aggregated samples was typically nine times higher than that of individual soil core samples, compared to around three times higher in bacterial assessments of the same soil cores. Fungal soil core samples typically had large numbers of ASVs with few reads. Further experimental analysis would be needed to determine if this is

due to actual community variation within sites, or due to methodological variability in identification of fungal ASVs via metabarcoding using the ITS1 genetic locus. Either way, I recommend sampling many soil cores to capture the fungal diversity of a site. It is unclear that mixing soils samples prior to DNA extraction would solve this problem, or if pooling samples following sequencing, as I have done, is a better solution. Existing research suggests that the problem does not stem from issues during amplification or sequencing as PCR replication has tended not to improve fungal sequence detection although different sequencing platforms may have improved taxon detectability (Smith & Peay, 2014; Singer *et al.*, 2019). Bacterial communities appear to be more homogenous across a site, and may require fewer soil samples to characterise, or be more amenable to characterisation by mixing samples before DNA extraction.

### *3.5.2 Limitations*

The bioinformatic methods employed in this work are robust but can struggle to handle community data with high numbers of zero values, as was the case for the fungal ASV data. The abundance distribution of ASVs within samples meant that pooling samples by site, following sequencing was necessary to characterise fungal communities intelligibly. The low read abundance of many taxa continued to hinder bioinformatic analysis even after pooling, affecting the transformation of read abundance for analysis and ordination. To allow for appropriate variance stabilisation and transformation of read abundance, I reduced the data

to fungal species that *DADA2* had assigned to species level. This alone changed the assessment of community structure for soil fungi but enabled the generation of fungal VST abundances that otherwise would have been seriously flawed. The improved transformation of fungal species abundance data over that of ASV abundance indicated that many low abundance ASVs had been removed or combined to produce the species dataset. I cannot see why the species assigned by *DADA2*, based upon the UNITE fungal taxonomic reference database would be biased in its composition, but readers may disagree. Including as much biological information as possible in assessment of alpha diversity is preferable whenever possible, but this was not an option for these data. The VST abundance-based ordinations of fungal communities presented in the results is striking. However, the reader must decide for themselves whether

the reduction in included taxa, which clearly affected ordinations, including decreasing NMDS stress, makes its conclusions less, more, or equally convincing.

The choice of statistical analysis in this chapter has been guided by two factors, stemming from the experimental design of the study. Firstly, during the initial analysis of the soil chemistry (Ashwood *et al.*, 2019), and in subsequent analyses, the importance of treating arable and pasture sites differently, as well as separating young and mature woodlands, became apparent. This changed the balanced, orthogonal design of seven replicates of each farmland, woodland, and ASNW, to a less balanced, non-orthogonal design with three, four, or seven replicates per group. Secondly, soil sampling was extensive, with ten replicates at each site leading to a large amount of data with potential for pseudo-replication. To make best use of these data I employed mixed effect models when model selection indicated that site effects were apparent in the data. Mixed effect models offer a powerful approach to handling non-independent data but also are easy to misuse. Criticisms of my approach might focus on my method of global model variable selection, its model refinement, and stopping short of multi-model inference or permutation testing. As outlined in the methods, my global model included all environmental variables included in the previous chapter. Well known issues to do with collinearity prevented my inclusion of both nitrogen and organic carbon concentrations in the model (Harrison *et al.*, 2018). Although alternatives, including incorporating variance inflation factors into model selection could have been implemented, I chose not to do so given the large number of variables. Perhaps this list of variables could have been further refined *a priori*, based upon the literature. But, as this approach to studying soil communities is relatively new, and might capture new aspects of microbial diversity, I chose not to make assumptions about the importance of variables like nitrate or ammonium concentrations on microbial community structure. I also chose not to fit random slopes for fixed effects as has been recommended for optimal modelling of random effects (Grueber *et al.*, 2011). This was due to the large number of fixed effects in the global model. The same reasoning prevented the inclusion of all interaction effects, although the inclusion of a land-use pH interaction was motivated by the importance of pH in NMDS ordinations which appeared to differ by land use, despite pH not significantly varying by land-use type as shown in the previous chapter. Model refinement by backwards selection and AIC is also contentious. Refinement by AIC alone would have led to highly complex final models, and by BIC would

have led to even more simplistic models than I have presented. By following the approach of Crawley (2013) I present models that are more simple, with the risk of overestimating the importance of variables in the candidate model. It is likely that some of the importance attributed to pH, organic carbon, and soil bulk density would be equally well accounted for in a more complex model or in an average of multiple models. However, the retention of these soil characteristics - pH, organic carbon, and bulk density - conforms with the literature outlined above. More complex approaches, including multi-model averaging or simulations (Harrison *et al.*, 2018), might better parameterise the effects of abiotic factors on soil communities. But for the purpose of this research, I am satisfied that the importance of land-use and woodland age category has been demonstrated in a manner that might be easily replicated by ecologists or capable industrial practitioners seeking to answer similar questions using metabarcoding data.

Previous work has made use of CCA to relate community structure to abiotic factors (Kuramae *et al.*, 2012). Analysis of the community data via CCA would have been inappropriate when the data was examined at the scale of the individual sample (i.e. the bacterial data) due to pseudo-replication that cannot be accounted for in the CCA structure, and at the site scale due to the sensitivity of CCA to low sample numbers. My method of analysis has demonstrated that significant differences exist between communities in different land-use types in at least one important measure of community structure. The analysis of soil microbial communities by isolating important NMDS axes follows the structure of the analysis of previous studies (e.g. (Rillig *et al.*, 2019)), I do not present analyses of both axes to lessen the possibility of false positives due to multiple testing and preserve experimental power. From my community ordinations, it is clear that significant aspects of soil communities are not shaped by the soil's land use, yet as NMDS plots do not assign variance to each axis, it is not possible to say the relative importance of each axis. In other words, there are differences based upon land use, but are these differences important to stakeholders and to the scientific community?

### *3.5.3 Does the uncertainty surrounding molecular approaches present new challenges to ecological interpretation?*

Sampling soil diversity and molecular methods more generally are not a panacea for assessing biodiversity, and that view is unlikely to be held by their implementers. Broadly, concerns around these approaches include difficulties in sampling representatively, methodological biases that may exaggerate the abundance of certain taxa making quantitative conclusions difficult, issues of taxonomic resolution such that taxonomic units may be over or under resolved compared to the level at which biologically important processes occur, issues in the accurate identification of taxa, limited taxonomy or reference databases (Winding *et al.*, 2005; Taberlet *et al.*, 2012; Bohmann *et al.*, 2014; Pompanon & Samadi, 2015). These criticisms of molecular approaches are only effective if we adopt standards for new methodologies that would preclude the adoption of the current best practices in all fields of ecological sampling. Each of them has parallels in traditional sampling techniques that has not prevented researchers from advancing ecological theory to its current point. Sampling methodologies of soils or indeed of any organisms must carefully consider replication and spatio-temporal scales. All survey techniques produce misclassifications due to human error, cryptic species, or limited keys or field guides, and taxa will always vary in their likelihood of detection in field surveys due to specific behaviours and biology (Durso *et al.*, 2011; Iknayan *et al.*, 2014). Extrapolating from observed organisms to landscape or habitat scale population sizes has challenges whichever survey methods is employed. Care must be taken when designing molecular experimentation with knowledge of the limitations of the approach. Molecular methods present advantages in cost-effectiveness, large scale data production, reproducibility, sensitivity, and taxonomic resolution that allow different experiments to be designed than would be feasible with traditional approaches (Baird & Hajibabaei, 2012; Zaiko *et al.*, 2016). Molecular methods will not replace traditional approaches in function, but the economies of scale associated with them, the demand for data needed to approach global problems, and the increasing potential or autonomous data collection may cause a seismic shift in the way ecology is done in the future (Derocles *et al.*, 2018).

### *3.5.4 Future work*

Further analysis might shed more light on the importance of these community differences but could not be undertaken in this project due to time constraints. Instead, I aim to determine in

the next two chapters if these differences in soil communities carry over to the communities of biologically impactful decomposers in woodland soils of varying age, resulting in differences in the rates of biogeochemical processes.

### *3.5.5 Synthesis and applications*

The microbial communities of woodland soils in this study generally differentiate from arable soils after creation within 50 years, broadly in agreement with previous findings (e.g. Jiao *et al.* 2018). Pasture soils microbial communities are more similar to woodland soils but significantly differ from those of ASNW soils in the composition and relative abundances of bacterial and fungal taxa. Individual pasture sites may have much more similar communities to woodlands than is normal, which may indicate that management decisions can have a large effect on the microbes found in pasture. The same is true for mature woodland, in that individual sites may have very similar soil microbial communities to farmland, indicating that management choices can cause woodland communities to shift to the more restricted communities more commonly found in arable systems. Woodlands of different age categories did not significantly differ in their microbial compositions, but woodlands continue to shift in their microbial diversity differentiating from agricultural sites even after hundreds of years of growth. This indicates that ASNWs have unique microbial communities and soils that cannot be replicated by the creation of new woodland sites.

## 3.6 References

Ashwood, F., Watts, K., Park, K., Fuentes-Montemayor, E., Benham, S. & Vanguelova, E.I. (2019) Woodland restoration on agricultural land: long-term impacts on soil quality. *Restoration Ecology*. 27 (6), 1381–1392.

Baas Becking, L.G.M. (1934) Geobiologie of inleiding tot de milieukunde. *Van Stockum & Zoon, The Hague, The Netherlands*. 263.

Baird, D.J. & Hajibabaei, M. (2012) Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular ecology*. 21 (8), 2039–2044.

Baldrian, P. (2017) Forest microbiome: diversity, complexity and dynamics. *FEMS microbiology reviews*. 41 (2), 109–130.

Balser, T.C., McMahon, K.D., Bart, D., Bronson, D., Coyle, D.R., Craig, N., Flores-Mangual, M.L., Forshay, K., Jones, S.E., Kent, A.E. & Shade, A.L. (2006) Bridging the gap between micro - and macro-scale perspectives on the role of microbial communities in global change ecology. *Plant and soil*. 289 (1), 59–70.

Barberán, A., Casamayor, E.O. & Fierer, N. (2014) The microbial contribution to macroecology. *Frontiers in microbiology*. 5203.

Barrios, E. (2007) Soil biota, ecosystem services and land productivity. *Ecological economics: the journal of the International Society for Ecological Economics*. 64 (2), 269–285.

Bates, D. (2006) *[R] lmer, p-values and all that*. [Online] [online]. Available from: https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html (Accessed 18 February 2020).

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, Articles*. 67 (1), 1–48.

Bennett, A.E., Daniell, T.J., Öpik, M., Davison, J., Moora, M., Zobel, M., Selosse, M.-A. & Evans, D. (2013) Arbuscular mycorrhizal fungal networks vary throughout the growing season and between successional stages. *PloS one*. 8 (12), e83241.

Błonska, E., Lasota, J., da Silva, G.R.V., Vanguelova, E., Ashwood, F., Tibbett, M., Watts, K. & Lukac, M. (2020) Soil organic matter stabilization and carbon-cycling enzyme activity are affected by land management. *Annals of Forest Research*. 63 (1).

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W. & de Bruyn, M. (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in ecology & evolution*. 29 (6), 358–367.

Bossio, D.A., Scow, K.M., Gunapala, N. & Graham, K.J. (1998) Determinants of Soil Microbial Communities: Effects of Agricultural Management, Season, and Soil Type on Phospholipid Fatty Acid Profiles. *Microbial ecology*. 36 (1), 1–12.

Brunet, J., Valtinat, K., Mayr, M.L., Felton, A., Lindbladh, M. & Bruun, H.H. (2011) Understory succession in post-agricultural oak forests: Habitat fragmentation affects forest specialists and generalists differently. *Forest ecology and management*. 262 (9), 1863–1871.

Burton, V., Moseley, D., Brown, C., Metzger, M.J. & Bellamy, P. (2018) Reviewing the evidence base for the effects of woodland expansion on biodiversity and ecosystem services in the United Kingdom. *Forest ecology and management*. 430366–379.

Callahan, B.J., McMurdie, P.J. & Holmes, S.P. (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*. 11 (12), 2639–2643.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A. & Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods*. 13 (7), 581–583.

Callahan, B.J., Sankaran, K., Fukuyama, J.A., McMurdie, P.J. & Holmes, S.P. (2016) Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses. *F1000Research*. 51492.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., 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 Suppl 14516–4522.

Chen, W., Jiao, S., Li, Q. & Du, N. (2019) Dispersal limitation relative to environmental filtering governs the vertical small-scale assembly of soil microbiomes during restoration Gaowen Yang (ed.). *The Journal of applied ecology*. 355.

Crawley, M.J. (2013) *The R Book*. Chichester: John Wiley & Sons.

Creamer, C.A., Filley, T.R., Boutton, T.W. & Rowe, H.I. (2016) Grassland to woodland transitions: Dynamic response of microbial community structure and carbon use patterns. *Journal of Geophysical Research: Biogeosciences*. 121 (6), 2016JG003347.

DEFRA (2018) *A Green Future: Our 25 Year Plan to Improve the Environment*.

Derocles, S.A.P., Bohan, D.A., Dumbrell, A.J., Kitson, J.J.N., Massol, F., Pauvert, C., Plantegenest, M., Vacher, C. & Evans, D.M. (2018) 'Chapter One - Biomonitoring for the 21st Century: Integrating Next-Generation Sequencing Into Ecological Network Analysis', in David A. Bohan, Alex J. Dumbrell, Guy Woodward, & Michelle Jackson (eds.) *Advances in Ecological Research*. [Online]. Academic Press. pp. 1–62.

De Vrieze, J., Ijaz, U.Z., Saunders, A.M. & Theuerl, S. (2018) Terminal restriction fragment length polymorphism is an 'old school' reliable technique for swift microbial community screening in anaerobic digestion. *Scientific reports*. 8 (1), 16818.

Dominati, E., Patterson, M. & Mackay, A. (2010) A framework for classifying and quantifying the natural capital and ecosystem services of soils. *Ecological economics: the journal of the International Society for Ecological Economics*. 69 (9), 1858–1868.

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N. & Wagner, H.H. (2020) *adespatial: Multivariate Multiscale Spatial Analysis*.

Dumbrell, A.J., Nelson, M., Helgason, T., Dytham, C. & Fitter, A.H. (2010) Relative roles of niche and neutral processes in structuring a soil microbial community. *The ISME journal*. 4 (3), 337–345.

Durso, A.M., Willson, J.D. & Winne, C.T. (2011) Needles in haystacks: Estimating detection probability and occupancy of rare and cryptic snakes. *Biological conservation*. 144 (5), 1508–1515.

Ettema, C. (2002) Spatial soil ecology. *Trends in ecology & evolution*. 17 (4), 177–183.

Fadrosh, D.W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R.M. & Ravel, J. (2014) An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*. 2 (1), 6.

Fichtner, A., von Oheimb, G., Härdtle, W., Wilken, C. & Gutknecht, J.L.M. (2014) Effects of anthropogenic disturbances on soil microbial communities in oak forests persist for more than 100 years. *Soil biology & biochemistry*. 7079–87.

Fierer, N. (2017) Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature reviews. Microbiology*. 15 (10), 579–590.

Forest Service (2006) *NI Forestry - a strategy for sustainable growth*.

Frelich, L.E. (2002) *Forest Dynamics and Disturbance Regimes: Studies from Temperate Evergreen-Deciduous Forests*. Cambridge University Press.

Frostegård, Å., Tunlid, A. & Bååth, E. (2011) Use and misuse of PLFA measurements in soils. *Soil biology & biochemistry*. 43 (8), 1621–1625.

Fuller, L., Fuentes-Montemayor, E., Watts, K., Macgregor, N.A., Bitenc, K. & Park, K.J. (2018) Local-scale attributes determine the suitability of woodland creation sites for Diptera Jörg Müller (ed.). *The Journal of applied ecology*. 55 (3), 1173–1184.

Gardes, M. & Bruns, T.D. (1993) ITS primers with enhanced specificity for basidiomycetes--application to the identification of mycorrhizae and rusts. *Molecular ecology*. 2 (2), 113–118.

Gellie, N.J.C., Mills, J.G., Breed, M.F. & Lowe, A.J. (2017) Revegetation rewilds the soil bacterial microbiome of an old field. *Molecular ecology*. 26 (11), 2895–2904.

George, P.B.L., Lallias, D., Creer, S., Seaton, F.M., Kenny, J.G., Eccles, R.M., Griffiths, R.I., Lebron, I., Emmett, B.A., Robinson, D.A. & Jones, D.L. (2019) Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. *Nature communications*. 10 (1), 1107.

Giller, K.E., Beare, M.H., Lavelle, P., Izac, A.-M.N. & Swift, M.J. (1997) Agricultural intensification, soil biodiversity and agroecosystem function. Applied soil ecology: a section of Agriculture, Ecosystems & Environment. 6 (1), 3–16.

Giller, P.S. (1996) The diversity of soil communities, the 'poor man's tropical rainforest'? Biodiversity and conservation. 5 (2), 135–168.

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M. & Whiteley, A.S. (2011) The bacterial biogeography of British soils. *Environmental microbiology*. 13 (6), 1642–1654.

Grueber, C.E., Nakagawa, S., Laws, R.J. & Jamieson, I.G. (2011) Multimodel inference in ecology and evolution: challenges and solutions. *Journal of evolutionary biology*. 24 (4), 699–711.

Hahn, C. & Lunt, D. (2019) *metaBEAT*. Github.

Hall, J.E., Whitbread, A.M., Kirby, K.J. & Joint Nature Conservation Committee (Great Britain) (2012) *National Vegetation Classification: Field Guide to Woodland*. Joint Nature Conservation Committee.

Handelsman, J. (2004) Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and molecular biology reviews: MMBR*. 68 (4), 669–685.

Hanson, C.A., Allison, S.D., Bradford, M.A., Wallenstein, M.D. & Treseder, K.K. (2008) Fungal Taxa Target Different Carbon Sources in Forest Soil. *Ecosystems* . 11 (7), 1157–1167.

Harrison, X.A., Donaldson, L., Correa-Cano, M.E., Evans, J., Fisher, D.N., Goodwin, C.E.D., Robinson, B.S., Hodgson, D.J. & Inger, R. (2018) A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*. 6e4794.

Haygarth, P.M. & Ritz, K. (2009) The future of soils and land use in the UK: Soil systems for the provision of land-based ecosystem services. *Land use policy*. 26, Supplement 1S187–S197.

Hermy, M. & Verheyen, K. (2007) Legacies of the past in the present-day forest biodiversity: a review of past land-use effects on forest plant species composition and diversity. *Ecological research*. 22 (3), 361–371.

Hosomichi, K., Mitsunaga, S., Nagasaki, H. & Inoue, I. (2014) A Bead-based Normalization for Uniform Sequencing depth (BeNUS) protocol for multi-samples sequencing exemplified by HLA-B. *BMC genomics*. 15645.

Iknayan, K.J., Tingley, M.W., Furnas, B.J. & Beissinger, S.R. (2014) Detecting diversity: emerging methods to estimate species diversity. *Trends in ecology & evolution*. 29 (2), 97–106.

Jangid, K., Williams, M.A., Franzluebbers, A.J., Schmidt, T.M., Coleman, D.C. & Whitman, W.B. (2011) Land-use history has a stronger impact on soil microbial community composition than aboveground vegetation and soil properties. *Soil biology & biochemistry*. 43 (10), 2184–2193.

Jiao, S., Chen, W., Wang, J., Du, N., Li, Q. & Wei, G. (2018) Soil microbiomes with distinct assemblies through vertical soil profiles drive the cycling of multiple nutrients in reforested ecosystems. *Microbiome*. 6 (1), 146.

Jolivet, P. & Foley, J.W. (2015) *Solutions for purifying nucleic acids by solid- phase reversible immobilization (SPRI)*. [Online] [online]. Available from: https://docslide.net/documents/spribuffersv23docx-view50-ml-conical-tubes-15-ml-microcentrifuge-tubes-disposable.html (Accessed 31 January 2019).

Kitson, J.J.N., Hahn, C., Sands, R.J., Straw, N.A., Evans, D.M. & Lunt, D.H. (2018) Detecting host-parasitoid interactions in an invasive Lepidopteran using nested tagging DNA metabarcoding. *Molecular ecology*.

Kuramae, E.E., Yergeau, E., Wong, L.C., Pijl, A.S., van Veen, J.A. & Kowalchuk, G.A. (2012) Soil characteristics more strongly influence soil bacterial communities than land-use type. *FEMS microbiology ecology*. 79 (1), 12–24.

Lladó, S., López-Mondéjar, R. & Baldrian, P. (2017) Forest Soil Bacteria: Diversity, Involvement in Ecosystem Processes, and Response to Global Change. *Microbiology and molecular biology reviews: MMBR*. 81 (2), .

Lladó, S., López-Mondéjar, R. & Baldrian, P. (2018) Drivers of microbial community structure in forest soils. Applied microbiology and biotechnology. 102 (10), 4331–4338.

Love, M.I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*. 15 (12), 550.

Mackay, J.E., Cunningham, S.C. & Cavagnaro, T.R. (2016) Riparian reforestation: are there changes in soil carbon and soil microbial communities? *The Science of the total environment*. 566-567960–967.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 17 (1), 10–12.

McMurdie, P.J. & Holmes, S. (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*. 8 (4), e61217.

Mitchell, R.J., Hester, A.J., Campbell, C.D., Chapman, S.J., Cameron, C.M., Hewison, R.L. & Potts, J.M. (2010) Is vegetation composition or soil chemistry the best predictor of the soil microbial community? *Plant and soil*. 333 (1), 417–430.

Moro, D. & Gadal, S. (2007) Benefits of habitat restoration to small mammal diversity and abundance in a pastoral agricultural landscape in mid-Wales. Biodiversity and conservation. 16 (12), 3543–3557.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., Knelman, J.E., Darcy, J.L., Lynch, R.C., Wickey, P. & Ferrenberg, S. (2013) Patterns and processes of microbial community assembly. *Microbiology and molecular biology reviews: MMBR*. 77 (3), 342–356.

Ng, K., Barton, P.S., Macfadyen, S., Lindenmayer, D.B. & Driscoll, D.A. (2018) Beetle's responses to edges in fragmented landscapes are driven by adjacent farmland use, season and cross-habitat movement. Landscape ecology. 33 (1), 109–125.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. & Wagner, H. (2019) *vegan: Community Ecology Package*.

O'Malley, M.A. (2008) 'Everything is everywhere: but the environment selects': ubiquitous distribution and ecological determinism in microbial biogeography. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 39 (3) p.314–325.

de la Peña, E., Baeten, L., Steel, H., Viaene, N., De Sutter, N., De Schrijver, A. & Verheyen, K. (2016) Beyond plant–soil feedbacks: mechanisms driving plant community shifts due to land-use legacies in post-agricultural forests. *Functional ecology*. 30 (7), 1073–1085.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team (2019) *nlme: Linear and Nonlinear Mixed Effects Models*.

Pompanon, F. & Samadi, S. (2015) Next generation sequencing for characterizing biodiversity: promises and challenges. *Genetica*. 143 (2), 133–138.

R Core Team (2019) *R: A Language and Environment for Statistical Computing*.

Remsen, J.V. (1994) Use and Misuse of Bird Lists in Community Ecology and Conservation. *The Auk*. 111 (1), 225–227.

Rillig, M.C., Ryo, M., Lehmann, A., Aguilar-Trigueros, C.A., Buchert, S., Wulf, A., Iwasaki, A., Roy, J. & Yang, G. (2019) The role of multiple global change factors in driving soil functions and microbial biodiversity. *Science*. 366 (6467), 886–890.

Rosas-Medina, M., Maciá-Vicente, J.G. & Piepenbring, M. (2020) Diversity of Fungi in Soils with Different Degrees of Degradation in Germany and Panama. *Mycobiology*. 48 (1), 20–28.

Schneider, T., Keiblinger, K.M., Schmid, E., Sterflinger-Gleixner, K., Ellersdorfer, G., Roschitzki, B., Richter, A., Eberl, L., Zechmeister-Boltenstern, S. & Riedel, K. (2012) Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. The ISME journal. 6 (9), 1749–1762.

Shi, Y., Delgado-Baquerizo, M., Li, Y., Yang, Y., Zhu, Y.-G., Peñuelas, J. & Chu, H. (2020) Abundance of kinless hubs within soil microbial networks are associated with high functional potential in agricultural ecosystems. *Environment international*. 142105869.

Shokralla, S., Spall, J.L., Gibson, J.F. & Hajibabaei, M. (2012) Next-generation sequencing technologies for environmental DNA research. *Molecular ecology*. 21 (8), 1794–1805.

Singer, G.A.C., Fahner, N.A., Barnes, J.G., McCarthy, A. & Hajibabaei, M. (2019) Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. *Scientific reports*. 9 (1), 5991.

Smith, D.P. & Peay, K.G. (2014) Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *PloS one*. 9 (2), e90234.

Strickland, M.S., Lauber, C., Fierer, N. & Bradford, M.A. (2009) Testing the functional significance of microbial community composition. *Ecology*. 90 (2), 441–451.

Stout, L.M., Joshi, S.R., Kana, T.M. & Jaisi, D.P. (2014) Microbial activities and phosphorus cycling: An application of oxygen isotope ratios in phosphate. Geochimica et cosmochimica acta. 138101–116.

Susyan, E.A., Wirth, S., Ananyeva, N.D. & Stolnikova, E.V. (2011) Forest succession on abandoned arable soils in European Russia – Impacts on microbial biomass, fungal-bacterial ratio, and basal CO2 respiration activity. *European journal of soil biology*. 47 (3), 169–174.

Sylvain, Z.A. & Wall, D.H. (2011) Linking soil biodiversity and vegetation: implications for a changing planet. *American journal of botany*. 98 (3), 517–527.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. & Willerslev, E. (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular ecology*. 21 (8), 2045–2050.

The Scottish Government (2019) *Scotland's Forestry Strategy 2019–2029*.

Thoms, C., Gattinger, A., Jacob, M., Thomas, F.M. & Gleixner, G. (2010) Direct and indirect effects of tree diversity drive soil microbial diversity in temperate deciduous forest. *Soil biology & biochemistry*. 42 (9), 1558–1565.

Tiedje, J.M., Asuming-Brempong, S., Nüsslein, K., Marsh, T.L. & Flynn, S.J. (1999) Opening the black box of soil microbial diversity. *Applied soil ecology: a section of Agriculture, Ecosystems & Environment*. 13 (2), 109–122.

Turley, N.E., Bell-Dereske, L., Evans, S.E. & Brudvig, L.A. (2020) Agricultural land-use history and restoration impact soil microbial biodiversity Gaowen Yang (ed.). *The Journal of applied ecology*. 57 (5), 852–863.

Uroz, S., Buée, M., Deveau, A., Mieszkin, S. & Martin, F. (2016) Ecology of the forest microbiome: Highlights of temperate and boreal ecosystems. *Soil biology & biochemistry*. 103471–488.

Waldrop, M.P., Zak, D.R., Blackwood, C.B., Curtis, C.D. & Tilman, D. (2006) Resource availability controls fungal diversity across a plant diversity gradient. *Ecology letters*. 9 (10), 1127–1135.

Wall, D.H., Nielsen, U.N. & Six, J. (2015) Soil biodiversity and human health. *Nature*. 528 (7580), 69–76.

Watts, K., Fuentes-Montemayor, E., Macgregor, N.A., Peredo-Alvarez, V., Ferryman, M., Bellamy, C., Brown, N. & Park, K.J. (2016) Using historical woodland creation to construct a long-term, large-scale natural experiment: the WrEN project. *Ecology and evolution*. 6 (9), 3012–3025.

Welsh Government (2018) *Woodlands for Wales*.

White, T.J., Bruns, T.D., Lee, S.B. & Taylor, J.W. (1990) 'Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics', in M. A. Innis, D. H. Gelfand, J. J. Sninsky, & T. J. White (eds.) *PCR Protocols: A Guide to Methods and Applications*. [Online]. New York: Academic Press. pp. 315–322.

Whytock, R.C., Fuentes-Montemayor, E., Watts, K., Barbosa De Andrade, P., Whytock, R.T., French, P., Macgregor, N.A. & Park, K.J. (2018) Bird-community responses to habitat creation in a long-term, large-scale natural experiment: Birds and Habitat Creation. *Conservation biology: the journal of the Society for Conservation Biology*. 32 (2), 345–354.

Whytock, R.C., Fuentes-Montemayor, E., Watts, K., Macgregor, N.A., Williams, L. & Park, K.J. (2018) Context-dependent colonization of terrestrial habitat ‘islands’ by a long-distance migrant bird. *Proceedings of the Royal Society B: Biological Sciences*. 285 (1885), 20181490.

Winding, A., Hund-Rinke, K. & Rutgers, M. (2005) The use of microorganisms in ecological soil classification and assessment concepts. *Ecotoxicology and environmental safety*. 62 (2), 230–248.

Wu, X., Xu, H., Tuo, D., Wang, C., Fu, B., Lv, Y. & Liu, G. (2020) Land use change and stand age regulate soil respiration by influencing soil substrate supply and microbial community. *Geoderma*. 359113991.

Zaiko, A., Schimanski, K., Pochon, X., Hopkins, G.A., Goldstien, S., Floerl, O. & Wood, S.A. (2016) Metabarcoding improves detection of eukaryotes from early biofouling communities: implications for pest monitoring and pathway management. *Biofouling*. 32 (6), 671–684.

Zak, D.R., Holmes, W.E., White, D.C., Peacock, A.D. & Tilman, D. (2003) Plant diversity, soil microbial communities, and ecosystem function: are there any links? *Ecology*. 84 (8), 2042–2050.

# Chapter 4. Changes in Leaf-Litter Decomposition Rate as Woodlands Age


## 4.1 Abstract

Decomposition is a major component of carbon cycling worldwide and has been suggested as an indicator of healthy soil functioning. Improving our understanding of the mechanisms of decomposition in woodlands is important to better understanding regional and global carbon models, as well as to making management decisions locally. Little attention has been paid to the role of long-term natural succession in shaping woodland ecosystems and their decomposer communities, with no existing studies comparing woodland decomposition across all stages of successional development with a suitable degree of replication. Furthermore, few existing studies demonstrate sufficient attempts to control or account for soil properties, understory characteristics, woodland structure or tree community composition in their understanding of decomposition. I conducted a full-replicated (n = 809), orthogonal, balanced study to test the effects of succession on the litter decomposition of five key UK tree species in 27 woodlands in the North East of England, over two seasons. I present a model of decomposition across newly created, mature, and ancient semi-natural woodlands; identifying significant predictors of decomposition that include soil properties and multivariate measures of ground cover and tree community. Litter mass loss was notably higher in ancient semi-natural woodland than woodland creation, but not significantly so, indicating directions for future work. Aspects of the study test implications of the Substrate-quality Matrix-quality Interaction hypothesis as well as the Home Field Advantage hypothesis. This study highlights the potential adaptations of existing leaf litter experimental designs to improve replication and better describe the mechanisms by which site variations drive differences in decomposition.


## 4.2 Introduction

### *4.2.1 Knowledge gaps surrounding decomposition and leaf litter experiments*

Decomposition regulates the flow of carbon through soil vegetation systems, controlling the flow of 90% of terrestrial plant production between sequestration or return to the rest of the carbon and nutrient cycle (Gessner *et al.*, 2010). Above- and below-ground diversity of vegetation and decomposers are the dominant influences on decomposition rate, along with

soil and site properties which exist in complex feedback loops with them (Trudgill, 1988). This is the case for multiple woodland systems processes, such as nitrogen cycling, that benefit humans and wildlife alike. A full understanding of the woodland floor would include detailed measuring of many processes and cycles that occur simultaneously within it. By focusing on a single process, such as decomposition, which itself requires multiple sub-process, such as lignin or cellulose digestion, we may produce a measure of the functioning of the woodland floor that can be easily replicated at multiple sites. Decomposition experiments using leaf litter bags have been a longstanding method for measuring litter decomposition rates, with research going back at least as far as the 1930s (e.g. (Falconer *et al.*, 1933)). This has fundamentally influenced our understanding of the impacts of climate change on regional carbon budgets (Aerts, 1997), the role of organisms in underpinning ecosystem processes (Petersen & Luxton, 1982), and the effects of climate change on ecological cycles (Coûteaux *et al.*, 1995). It has rarely fallen out of fashion as a study methodology; instead, it has proven to be highly adaptable and is recurrently applied to test new ecological theory. This classic approach remains prominent in the literature and has been used to assess the relative contributions of compartments of microbial diversity to decomposition (Bani *et al.*, 2018), to demonstrate the positive effect of grazing on grassland decomposition rate (Chuan *et al.*, 2018), and show how warming climates may lead to a decrease in grassland decomposition rate due to impacts on soil mesofauna (Yin *et al.*, 2019). Increasingly, litter studies are taking advantage of next-generation sequencing to better understand the variation of microbial diversity between samples and treatments at high taxonomic resolution (López-Mondéjar *et al.*, 2015; Purahong *et al.*, 2016; Asplund *et al.*, 2018; Buresova *et al.*, 2019; Xiao *et al.*, 2019; Ritter *et al.*, 2020; Štursová *et al.*, 2020). These approaches increase the capacity of researchers to generate ecological data and should encourage the implementation of large-scale ecological experiments relating diversity to ecosystem functioning.

### *4.2.2 Woodland succession influences decomposition*

Following afforestation, British woodlands change as they age, broadly following one of several potential patterns of succession (Frelich, 2002). Depending on the land-use history of the site, regeneration may start from a simple forest system, such as a uniform plantation that is never harvested, or natural regeneration may include complex structure and composition from the start, especially in cases of natural regeneration or when afforested sites adjoin

established woodland (Frelich, 2002; Brunet *et al.*, 2011). The successional pathway of a woodland, including its composition and structure, will depend on a variety of factors, including management, restocking, and site-specific factors. These may or may not eventually converge on a climax community over time. In the UK sufficient time has not passed since widespread intensive management of all ecosystem types to produce climax communities in even ASNWs, but these sites have changes in numerous characteristics beyond the tree composition and structure that distinguish them from afforested sites (Rackham, 2012). Whether or not created woodlands are on the same trajectory as naturally regenerated woodlands, and how management decisions influence these trajectories are major areas of research (Hermy & Verheyen, 2007). However, change is inevitable. In the second chapter I examined how the soils in a chronosequence of sites from unforested farmland to ancient semi-natural woodland change in several properties. These shifts included increasing organic carbon stocks, nitrogen stocks, and carbon to nitrogen ratios. Soil bulk density also varied across this gradient, perhaps due to mechanical disruption during the woodland creation process which persisted during early woodland development. The diversity and structure of the vegetation above-ground also develop over time. Changes to the structure and composition of woodland lead to new ecological niches that are filled over short or long time scales by flora and fauna depending on colonisation rate (Hermy & Verheyen, 2007). In the previous chapter, I demonstrated that the soil property changes observed in chapter 2 were accompanied by shifts in the bacterial and fungal communities of these same soils. This indicates that the processes that govern woodland succession above-ground have analogues below the soil surface. Colonisation, competition, and adaptation shape microbial community assemblage in woodlands. Although it may seem tautologous, the importance of trees to woodland microbial community composition cannot be understated. The microbial community and soils in the immediate metres surrounding a tree strongly are impacted by it (Jiao *et al.*, 2018; Dean *et al.*, 2020). Mutualistic and antagonistic feedback between microbes and trees in a location are likely to have played out for centuries under ancient trees. In a tree's lifespan, uncountable microbial generations will have experienced symbiotic feedback with it and its leaf litter and deadwood. Evidence suggests that decomposer communities adapt to the local conditions (resulting from selective pressure, physiological adaptation, taxonomic filtering, and evolutionary adaptation), decomposing leaf litter from a woodland faster where it is collected than in an alternative wood (Ayres *et al.*, 2009). This is known as the Home Field Advantage (HFA) effect and has been demonstrated in woodland as young as

11 years old (Sun & Zhao, 2016). The scales at which it operates, spatially and taxonomically, are unclear but has implications for the understanding of decomposer community assembly. It is unclear whether the spatial scale is limited to the area surrounding specific source trees or whether it operates equally across a whole woodland. We currently lack conclusive answers as to the scale and specificity of the HFA effect. Is a community adapted to decomposing the litter of a specific tree, to a population of trees, to the tree's species, to its larger taxonomic groups, or is the effect polyphyletic and based on leaf chemistry as is suggested by the substrate quality–matrix quality interaction (SMI) hypothesis (Freschet *et al.*, 2012)? The competing SMI hypothesis suggests that litter decomposes faster in a matrix of similar litter quality to itself, i.e. recalcitrant material decomposes fastest at a site where recalcitrant material makes up a large portion of the annual litter input and vice versa for easily decomposable material. Both hypotheses may be explained by adaption or alteration of the decomposer community over time. Resolving uncertainties surrounding these hypotheses could yield insights not only into the mechanisms of decomposition, but also the spatial scale at which soil microbial community assembly occurs and the robustness of microbial communities to tree species loss or land-use change. Few investigations exist into how the age of a woodland impacts the home-field advantage effect, or generally into how woodland age impacts decomposition rates of woodland species. We might expect the HFA or SMI to increase and plateau over some time scale, with a lag period between tree planting and the resulting effect. Over this period, colonising decomposers arriving at a woodland will supplement or replace resident decomposers. Existing microbial communities would adapt to capitalise on the decomposition of leaf litter specific to the site. However, this has not been tested over a range of successional stages. Work on Eucalyptus plantations showed declining rates of decomposition over a chronosequence representing the first eight years of plantation growth (Bargali *et al.*, 1993) and related results to changes in litter quality over that time. Another eucalyptus study observed similar trends of declining decomposition rate in mature forest (which we are assured is "probably several hundred years old"), with the younger 40, 9 and 6-year-old stands showing greater amounts of decomposition over the period relating to the difference to leaf litter quality (O'Connell, 1988). Similar, but non-significant, drops were observed in decomposition rates at rubber plantations; these were fastest in rubber aged 7 years but dropped in those aged 12 and 25 years (N'Dri *et al.*, 2018). In subtropical Chinese forest, leaf decomposition declined with stand age in stands aged from 22 to 116 years and was explained by changes in forest microclimate (Trogisch *et al.*, 2016). Douglas Fir

decomposed fastest in stands aged over 85 years than younger woodlands, and paper birch decomposed fastest in young woods (10-25ya) in a study in mixed stands of the two species in British Columbia and was shown not to be related to litter quality (Welke & Hope, 2005). Many other publications have explored differences in decomposition rate in various plantation species, along time scales relevant to commercial forestry. However, to my knowledge, no other study has measured leaf litter decomposition directly across newly created, mature and ancient woodlands. The most comparable study was by Barlow *et al.* (2007) conducted at three large sites: Amazonian eucalyptus plantation (4-9 years old), secondary woodland (14 - 19 years old), and primary Amazonian forest and detected significantly higher decomposition on the secondary and primary woodland than the plantation. The study has excellent within-site replication, but only examines a single site of each category and does not account for confounding effects of tree and understory vegetation composition. Although this work compared woodland of vast differing age, it still takes place in a silviculture plantation context, and the secondary woodland of 14-19 years is not comparable to what we would consider mature woodland of several hundred years in age in the UK. A study that examines young, mature, and ancient woodland as defined in a UK context is lacking, as is a study that separates the effects of succession on vegetation and soil properties from those of decomposer community succession. Studies also tend to use leaf litter collected from multiple sites, partially or completely confounding litter quality and site quality effects (although good examples of studies with the litter of multiple commonly occurring tree species do exist see Barlow *et al.* 2007). Studies that exist are rarely replicated across sites within age classes or replicated within sites. Often studies have no replication of sampling periods within sites. Even without molecular techniques, lack of replication is a long-standing issue in studies of this kind. A study considering ecologically relevant, rather than commercially relevant, time scales with an attempt at replication of sites and sampling period, is absent from the literature and greatly needed.

Decomposition is influenced by multiple drivers. Temperature, moisture availability, leaf litter quality, surrounding tree communities, and decomposer community diversity and composition for both mesofauna and microbes are powerful influences on decomposition rate (Bradford *et al.*, 2016; Xiao *et al.*, 2019). These factors are often considered in leaf litter decomposition projects, yet soil conditions are infrequently examined. This is surprising given

the well-documented associations between soil conditions such as pH, organic matter, nutrient availability and moisture and microbial composition and diversity (Lladó *et al.*, 2017). The soil may be a reservoir of decomposer diversity from which seasonal decomposers colonise fallen litter. Recent publications that review the drivers of microbial decomposition discuss how litter inputs shape soil properties but there has been less discussion so far of how soil chemical and physical properties influence the decomposers of leaf litter (Baldrian, 2017; Bani *et al.*, 2018; Porre *et al.*, 2020). Soil properties may be more important at sites lacking a thick organic or litter layers, where falling litter is decomposed or dispersed into the soil quickly rather than accumulating and forming its own microclimate. A recent meta-analysis has made efforts to include soil chemistry data to characterise the conditions in which the non-additive effects of mixing leaf litter of species appear (Porre *et al.*, 2020). They identify several soil characteristics, soil C, N, and C:N, as significantly affecting decomposition when data from across studies are analysed in varying combinations. Bacterial community structure has been demonstrated to be shaped by multiple soil properties including pH, C:N ratio, and soil texture (Griffiths *et al.*, 2011; Fierer, 2017). In fungi, the importance of soil properties on shaping community structure appears to be less strong, but soil properties still impact the abundances of different fungal groups, such as observed effects of pH on the abundance of fungal taxa (Rousk *et al.*, 2010), or the observed impact of C:N ratio or N on fungal community composition (Lindahl *et al.*, 2007). It is possible that soil properties also impact fungal communities by altering associated and co-occurring bacterial communities (Seaton *et al.*, 2020). Organic input is a major driver of these soil properties, and litter quality, tree composition, and tree diversity are often included in models of decomposition for this reason. However, these are not the sole determinants of soil chemistry and Bradford *et al.* (2016) recommend direct measurements of soil chemistry at study sites to improve current models of decomposition rate. It remains uncommon in the literature to directly measure soil chemistry, perhaps due to the lack of site replication in many leaf litter experiments. Studies with higher site replication are needed to compare the effects of soil properties on community structure.

### *4.2.3 Implications of land-use change effects on ecosystem functioning*

Decomposition may be an indicator of healthy functioning in other soil processes for land managers (Doran & Zeiss, 2000). Understanding how ecosystem functioning varies between

habitats would be of applied interest to land managers seeking to quantify the cost and benefits of converting land. In the UK woodland creation and the conservation of ancient semi-natural woodland (ASNW), along with other woodland of lower conservation priority is of great national interest and public concern. However, our understanding of soil ecosystem functioning across this successional gradient is poor. Woodland planting is becoming part of governmental strategies to combat biodiversity loss and climate change in all of the national and devolved governments, which also highlight the importance of protecting ASNW (Forest Service, 2006; DEFRA, 2018; Welsh Government, 2018; The Scottish Government, 2019). Although ASNW is classified as an irreplaceable habitat, when its loss is unavoidable the possible mitigation strategies include woodland creation and habitat translocation, although it is noted that this does not replace the lost environment (Parliamentary Office of Science and Technology, 2014). Notably, recent infrastructure projects, such as the building of the channel tunnel and the High Speed Two (HS2) rail project have impacted or grubbed out multiple ancient woodlands and these mitigation strategies have consequently been employed (Helliwell *et al.*, 1996; HS2 Ltd, 2017). If we wish to assess the success of woodland creation or translocation as a mitigation for the loss of mature or ancient woodland in the context of soils, the successional changes in soil microbial biodiversity and soil ecosystem functions between created and ancient woodland must be well understood.

Tree selection for planting projects will also play a key role in shaping soil communities at new sites. Conservation woodland, such as The Woodland Trust's creation sites that dominate this study are species-rich, with the goal of allowing the site conditions to select successful trees from a broad range of species. They also aim to encourage the establishment of species of conservation value, including ancient woodland indicator species; and to allow for unexpected climate and invasion events to select winners and losers from a broad stock. Yet the impacts of tree selection in planting projects on microbial communities is unstudied. The history of silviculture in Northumberland and across the British Isles means that anthropogenic as well as natural influences have shaped the tree community composition in older mature, plantation, and ASNW sites too. Because of this tree community composition at sites remains important but may be less strongly tied to successional processes and regional variation than the ground vegetation of a site. Complex histories of land management bring into question

the over-reliance in the literature on tree community data, rather than ground vegetation data which may better indicate woodland quality.

It is important to recognise that high or low decomposition rates are not, in and of themselves, indicators of woodland soil health. One view of soil carbon cycling would suggest that lower decomposition rates would be preferred to maximise carbon sequestration (Jandl *et al.*, 2007) as decomposition converts organic material to $CO_2$. Alternatively, increased decomposition rate could be seen as a sign of a healthy soil that can rapidly turnover the nutrients needed by plants (Arias *et al.*, 2005). Because of this, conclusions relating to optimal decomposition rates or carbon sequestration management strategies are not meant to be drawn from this work. In this work, I consider decomposition as a multi-functional process (Wagg *et al.*, 2019) - one that requires a range of organisms capable of performing many functions under a range of environmental conditions. Dispersal and competition ensure that over time woodlands either acquire the organisms best adapted to decomposing the litter found at that site or the populations at the site adapt to decompose the litter found there to outcompete newly arriving competitors (Ayres *et al.*, 2009). It is unclear how long this process takes to occur. By comparing the decomposition rates of woodlands of differing age, while accounting for other influences on decomposition rate, we can detect differences in the specialisation of the decomposer community to decomposing leaves from trees commonly found in and around those woodlands. I hope to demonstrate whether the specialisation of microbial communities to decomposing woodland leaves occurs quickly over a few decades or slowly over hundreds of years of woodland succession. I also can examine how the abundance and basal area of a tree at a site causes changes to the soil ecosystem that influence the decomposition of its litter. In some ways, this might be thought of as an as-yet-undefined parallel successional process to complement aboveground succession of trees and ground flora. The time it takes for the microbial community to adapt to decompose woodland tree leaves, rather than litter from the previous land-use type, may be indicative of the time it takes for the microbial community to adapt in other ways. Changes in decomposition provide one potential measure of microbial ecological adaptation and succession.

In the previous chapter I demonstrated that the bacterial and fungal community composition of ASNW is not fully restored in 50-60ya woodland creation sites. In this chapter I will go on

to assess whether this impacts upon the ecosystem functioning of those communities, in a large scale highly replicated study in the North East of England. Numerous variables are known to impact microbial diversity across the UK (Griffiths *et al.*, 2011). I explored the normal variations of woodland soils over successional stages in the first chapter of this thesis, albeit in the different landscape of the English midlands. I found that woodland soils vary in properties that are important for predicting woodland diversity, such as C:N ratio, organic carbon content, and pH. Although it was not analysed, they also differed in soil texture. Although I did not find significant differences across the 15 woodlands in that analysis, low sample sizes (four woodland creation, three mature secondary, seven ASNW) may have precluded this. Other studies have found characteristics such as pH (Susyan *et al.*, 2011; Kurganova *et al.*, 2018), and soil carbon (Kurganova *et al.*, 2018), and C:N ratio (Kurganova *et al.*, 2018) to vary across woodland successional stages. Previous studies have often not described if or how they control for these confounding variables between study sites. In this study I suggest a method for controlling for these factors that requires only simple laboratory equipment and techniques. I group sites by pH and soil texture, which both influence microbial diversity directly (Seaton *et al.*, 2020), but also may influence other soil properties that further structure soil microbial communities. Each group contained a single site of the three treatment categories of my study, woodland creation, mature woodland and ASNW, allowing site and this group triplicate variable to be used as a random effect to control for multiple site similarities.

### *4.2.4 Hypotheses*

This study establishes a fully-replicated experiment to compare leaf-litter decomposition rates (and vegetation characteristics) between newly created, mature secondary and ancient semi-natural woodlands in NE England using leaf-litter bags containing material from five common British tree species. We demonstrate the value and feasibility of studying long chronosequences and highlight the need to monitor and control for differences in soil properties during analysis and site selection.

I hypothesise H1) Woodland age category will be a strong predictor of the composition of ground vegetation but not tree composition, due to the different roles of anthropogenic

forces in shaping tree and ground cover communities.H2) Mass lost to decomposition will differ by woodland age, older woods will have less leaf material remaining at the point of sampling for each sampling period as microbial communities there have had a longer time to adapt to decomposing litter from trees present at these sites. H3) Tree community composition will impact leaf litter decomposition as proposed in the HFA and SMI hypotheses. H4) Decomposition rate will be higher when the leaf litter taxa have high importance at the decomposition site, that is to say that microbial communities will adapt to decompose the commonest, largest trees at a site which produce the most litter.H5) Ground vegetation will not significantly impact decomposition rate, so far it has rarely been incorporated into HFA studies, perhaps because the material is less recalcitrant and falls in a lower mass.

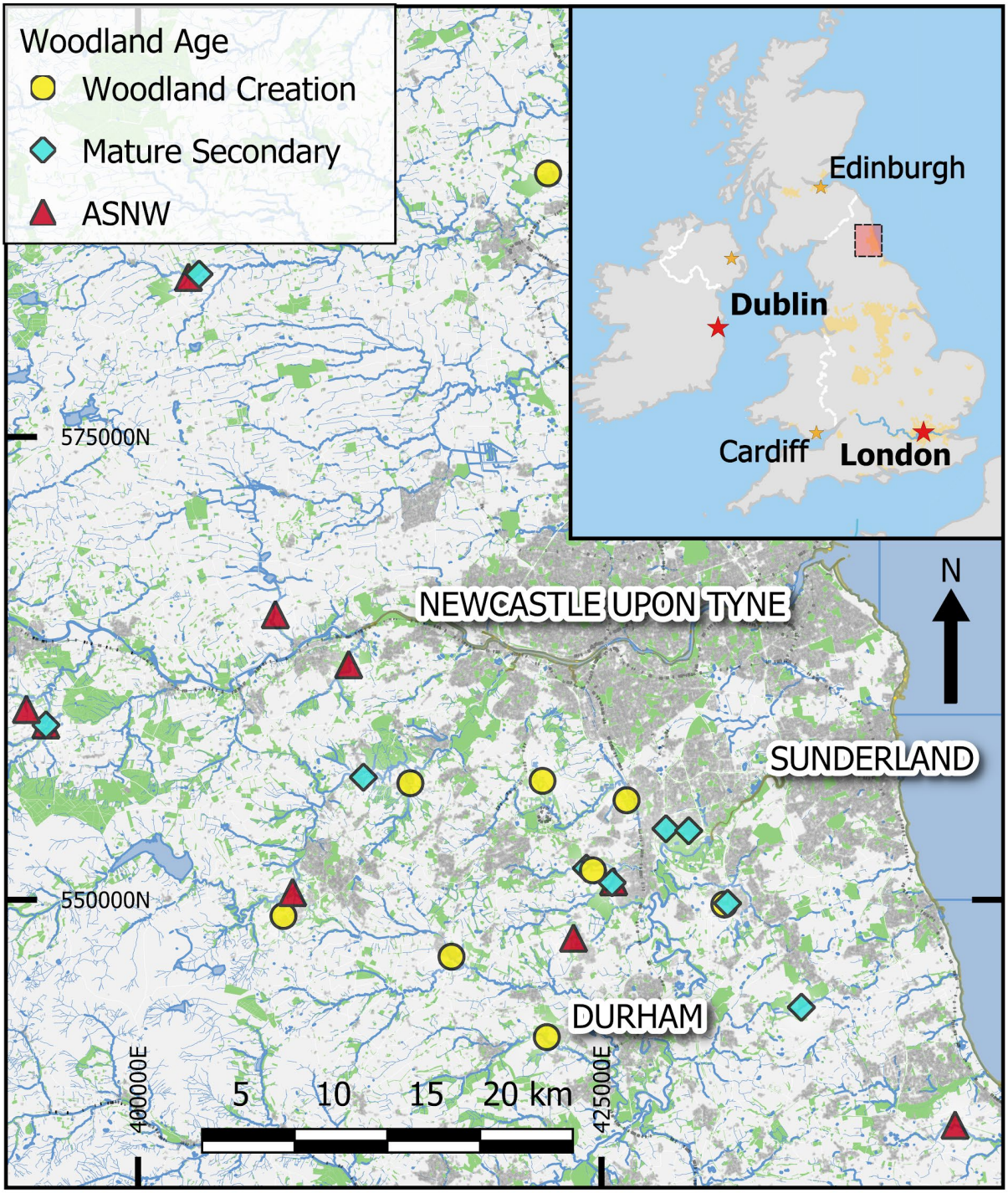


Figure 4.1 The 27 broadleaf woodland sites in the NE of England selected for litter decomposition experiments. Woodland age category is indicated by colour and shape for woodland creation sites (0 - 50ya), mature Woodland (50-400ya), and ancient semi-natural woodland (established before 1600 CE). Full data in Table S4.1.

## 4.3 Methods

### *4.3.1 Fieldwork summary*

Briefly, I selected 27 woodlands in NE England (Figure 4.1), nine woodland creation (10-50ya), nine mature secondary woodland (approx. 50-150ya), and nine ASNW (extant pre-1600), and grouped them into site triplicates, as described below, to minimise differences in pH and soil texture that are known to impact microbial populations. I buried single species 1g (dry weight) leaf litter bags from five common British tree species at these sites, just below the soil surface. I collected three replicate litter bags of each species from each site in the spring and summer, totalling 809 leaf litter bags (having not placed out one elm litter replicate due to shortage of material).

### *4.3.2 Site selection*

Potential candidate woodlands were suggested based upon local knowledge from The Woodland Trust and Durham County Council, to ensure that the woodlands selected were representative of the local landscape. Woodlands with atypical land-use histories were excluded at this stage, such as woodland planted on brownfield sites, and woodlands were required to be over two hectares in size. I used the ancient woodland inventory to confirm the presence and extent of Ancient Semi Natural Woodland (ASNW) at sites (Natural England, 2011). I frequently had access to the woodland planting year of young woodlands, but for older woodlands I had only approximate dates. The woodlands identified were all located on publicly accessible land; including sites owned by The Woodland Trust, Durham County Council, Gateshead Council; or owned by The National Trust. In the summer of 2017, I collected 160 soil cores across 29 candidate woodlands, with some sites consisting of compartments of adjacent woodland of differing age that allowed for 36 potential locations. While visiting woodlands, I made note of the site's characteristics, such as the presence of woodland indicator species, disturbance to the soil or evidence of previous disturbance (such as the dominance of bracken or bramble in the understory, or evidence of fire), uniform stands of trees (which may be evidence of planting), ancient or veteran trees, and other indicators (such as the aspect, locations, wood banks, previous management)to substantiate the woodland age category classification. To streamline the section process, I rejected sites that appeared to be misclassified, plantation sites on ancient woodland, and sites without suitable, accessible sampling locations. In the North East, Ancient woodland indicator understory

species include the abundant, and easy to identify, Dog's Mercury (*Mercurialis perennis*) and Ramsons (*Allium ursinum*); along with many fern species such as Broad Buckler Fern (*Dryopteris dilatate*) and Male Fern (*Dryopteris filix-mas*). Additionally many of the less common woodland specialist plants are ancient woodland indicators and I identified them when possible, during these surveys. Younger sites are often characterised by an abundance of grass cover, bracken (*Pteridium aquilinum*) or brambles (*Rubus fruticosus*). Indicator tree species were useful to describe sites, but the planting of ancient indicator trees such as field maple (*Acer campestre*) and others on woodland creation sites means that once reliable indicators may now mislead visitors as to the age of a site.

To characterise and group sites, I collected soil cores from the top 15cm of the soil with an auger from around 5 locations at each site. This provided multiple potential locations with a large range of soil properties for pairing sites. Potential locations for sampling were selected on-site based upon flat site relief, accessibility, and suitable space for burying leaf litter. I also tried to sample widely across the woodland to maximise the variations in soil properties sampled. The location of the soil core was recorded physically on a site map, the coordinates we recorded on a GPS device, a written description of the site was created, and photographs of the area and approach were taken. Soil cores were stored in polyethene bags, hereafter referred to as sample bags and transferred to -20°C storage at the end of each day of sampling. Soil cores were analysed to determine pH (aqueous), soil organic carbon (LECO RC612 Multiphase Carbon and Water analyser), total nitrogen (CHN628 Series Elemental Determinator), and soil texture (DEFRA, 2006). Triplicates of sites were assembled based on these characteristics (Table S4.1); one woodland creation, one mature secondary woodland, and one ASNW. Triplicate sites had to be within 0.5 pH units of one another, no more than one step away from each other on a soil texture triangle (DEFRA, 2006) and of similar hydrological profiles. Preference was given to including multiple compartments of differing age from a single woodland in the same triplicate when possible, to maximise site similarity (such as for Nunsbrough, Congburn, Wallington, and Newfield woods). I was successful in ensuring that the proportion of soil organic carbon was also similar across woodland types ($F_{(2, 24)} = 0.905$, $p = 0.42$), and C:N ratio ($F_{(2, 24)} = 0.351$, $p = 0.71$), although C:N ratio did vary considerably within triplicates. Other soil variables such as total carbon, and total nitrogen correlated strongly with organic carbon and so were not modelled.

### *4.3.3 Accounting for spatial autocorrelation*

In line with the previous chapters, I quantified spatial autocorrelation using Distance-based Moran's eigenvector maps (dbMEM) (Borcard & Legendre, 2002; Dray *et al.*, 2006), selecting important components via forward selection that were then incorporated in later models as fixed effects. I undertook all analyses described below in R (v3.6.0). The easting and northing coordinates of the sites were used to calculate dbMEM using the package *adespatial* (Dray *et al.*, 2020). Forward selection indicated that the first MEM would significantly improve model fit for models of C:N ratio and some measures of total nitrogen. Including additional MEMs beyond the first never significantly improved model fit. As such, I brought forward this first MEM component for inclusion in further models to represent spatial autocorrelation of sites.

### *4.3.4 Vegetation characteristics*

In July 2018, I surveyed sites to estimate tree density, tree species composition and diameter at breast height (DBH) of trees (at 1.3m height). I also surveyed ground flora cover using the Braun-Blanquet scale. I surveyed 10 random points for vegetation surveys within 20m of each of the 27 locations chosen for decomposition experiments. Each vegetation survey point was separated by a minimum of 5m. The first point chosen for surveying vegetation was always located directly on top of the location where leaf litter was buried for decomposition. The rest were located randomly within the 20m radius survey area. I identified vegetation survey points using uniformly random compass directions, combined with a distance out from the centre decomposition location to that point (drawn from a square root of a uniform distribution from 0 to 1, multiplied by the radius of the survey area to avoid clustering). At these points I collected tree vegetation data by surveying four trees per point, for a total of 40 trees per site, and assessing ground cover within four 0.25m x 0.25m quadrats, for a total of 40 quadrats per site. Following the point-centred quarter method (Sutherland, 2002), at each of these vegetation survey points I defined four quadrants along the cardinal compass directions. I identified the nearest tree to the survey point within each quadrant and measured its distance from the point, the species, and the DBH. The four ground flora quadrats were located in the previously defined quadrants at random distances from the point of up to 1m.

There I centred the quadrat and assessed the percentage ground cover of different vegetation types into six categories (Braun–Blanquet scale: <1%, 1-5%, 6-25%, 26-50%, 51-75%, 76-100%).

The metrics produced from the point-centred quarter method data are commonly used and calculated using established methods (Dahdouh-Guebas & Koedam, 2006), I have repeated them here for clarity. I pooled data by genus. I converted the tree community data to density in trees per $m^2$ by calculating the inverse mean distance from each point to its nearest four trees. I calculated the genus' density as the product of the density of all trees at a site and the genus' proportion. I calculated genus' relative density as the quotient of the number of individuals of the genus divided by the total number of trees recorded (40), expressed as a percentage. I calculated the basal area for each tree by converting DBH to cross-sectional area at breast height, assuming the cross-section is approximately circular. I determined dominance as the product of genus' density and genus' mean basal area at the site. I calculated relative dominance as the quotient of the genus' dominance divided by the sum of the dominance of all genera at the site, expressed as a percentage. I calculated frequency as the quotient of the number of sample points the genus occurred at divided by the number of points, expressed as a percentage. I calculated the relative frequency of each genus at a site as the quotient of the genus' frequency of that species at the site divided by the sum of the frequency of all species at the site, expressed as a percentage. The sum of the relative frequencies for each site approximated to 100. I calculated the importance of a tree species at a site as the sum of the relative density, the relative dominance, and the relative frequency of the species at each site.

### *4.3.5 Leaf litter collection and burial*

I collected leaf litter from 5 species of trees for the decomposition experiment: Common Ash (*Fraxinus excelsior*), Pedunculate Oak (*Quercus robur*), Common Beech (*Fagus sylvatica*), Common Hazel (*Corylus avellana*), Wych Elm (*Ulmus glabra*). From September to December 2017, I collected leaf litter from trees outside of the study area using a $1m^2$ trap fitted with a net that hung clear from the ground (Figure 4.2).

Figure 4.2 Examples of leaf litter collection using netted leaf litter traps. Air-dried litter was homogenised before filing litter bags to simulate natural physical disruption (bottom-left).

For all species, excluding ash, I sampled from a single location and when possible, a single tree. I collected oak leaves and a portion of the ash leaves from a private residence at Lat: 54.9380, Long: -1.7268. I collected the remaining leaves from Gosforth Park Nature Reserve at approximately the following locations: beech (Lat: 55.0283, Long: -1.5933), ash, (Lat: 55.0265, Long: -1.5917), elm (Lat: 55.0265, Long: -1.5917), hazel (Lat: 55.0247, Long: -1.5933) . I sorted out non-target species and stored leaves to dry in breathable bags inside a constant temperature room (20°C) with a dehumidifier for storage and air drying. When handling samples I attempted to minimise the potential for contamination by clearing then cleaning all working areas with 0.5% bleach, followed by 80% ethanol and wore gloves whenever working with samples. I dried leaves for at least one week before homogenising litter with a food processor for 5-10 seconds (Figure 4.2). I cleaned all equipment with bleach and ethanol between uses. I filled leaf litter bags with 1g (± 0.01g) of this dry, homogenised leaf litter. I indicated the species in each bag with a coloured plastic tag corresponding to the species.

Once assembled, I stored the leaf litter bags by species until they were assembled with other bags in a rigid leaf litter frame (Figure 4.3).

Figure 4.3. Preparing a site for litter burial (left) involved mapping the location of bags, photographing sites, and preallocating the sampling order. Leaf litter bags were assembled on frames (right) containing a tagged litter bag from each species such that an entire sampling replicate could be easily found and removed.

### *4.3.6 Leaf litter burial and sampling*

I buried leaf litter frames in January 2018. To minimise spatial variation, I aimed to bury replicate frames of leaf litter bags in a line, with the nine frames oriented perpendicular to the line and separated by 30cm (Figure 4.3). However alternative layouts were sometimes necessary, and all layouts were recorded. I avoided burying leaves within a metre of trees and chose flat areas of ground to work on. The sampling order of frames was randomly allocated in advance to prevent systematic spatial variation differences and I noted where each frame was buried to ease collection. Frame burial was conducted to minimise the disturbance of the soil by wedging open the ground with a spade and placing the frames upright in the trench created, closing it with pressure on either side of the trench. Photos and drawing of the site layout were taken. At Elemore woods, the exact site had been disturbed by the removal of a clump of elm, so the study site was shifted approximately 5-10m away from the original sampling location to the nearest available spot.

Litter from alternative sites origins was distributed as follows: at 21 sites one ash replicate (63 samples) was included from Gosforth Nature Reserve in each sampling season and was tracked with coloured tags to monitor for differences in decomposition. Missing litter bags due to insufficient collection of elm were distributed as follows: 16 sites missed one replicate

of elm in the third sampling season (48 samples - this season has not been analysed in this work). At one site one elm replicate was missed in the second sampling season (i.e. only two replicates were placed out at that site for collecting that season). I randomised species order within frames and labelled the frames with an ID indicating if they differed from the standard litter replicate frame as described above. These markings were visible when frames were collected.

I removed sample frames over three sampling seasons, separated by 13 weeks. Sampling periods fell during April, July, and October of 2018, each lasting three weeks. I returned to sites in the same order. I selected frames for sampling in advance of site visits to ensure that the sampling design was followed as outlined in the previous section, for generic replicates specific, random frames were selected. I removed these frames and immediately transferred leaf litter bags to individual sample bags (Figure 4.3). I wore nitrile gloves and minimised contact with the litter bags. If I uncovered the incorrect frame, I replaced it immediately. Additionally, I took a soil core at the site during every sampling period, cleaning the auger with a wipe then decontaminating it with a *Chemgene* wipe (1:20 concentration). The soil core was then stored in a polythene bag and frozen that day at -20°C. These cores have not been analysed as part of this study.

### *4.3.7 Leaf litter handling - preparations for later molecular work*

All leaf litter samples from the spring and summer sample collections were taken forward for dry weight and molecular analysis. While being transferred or handled I recorded the weights of material to a thousandth of a gram to monitor for the potential accidental loss of material. It was essential to prevent contamination of material with DNA from the laboratory or from other samples to conduct molecular work for the subsequent chapter. To prevent contamination, I handled material wearing gloves, using single-use sterile spatulas, and worked over sterilised aluminium sheets which were refreshed as needed. On the day of collection from the field, and thereafter when not being handled, I stored samples at -20°C. I removed approximately 0.1g of leaf litter material from each sample for molecular analysis to determine biodiversity for the following chapter. I weighed the filled Eppendorf tube before and following the removal of this material. Samples were then randomly organised into pre-libraries of between 89 and 92 samples for molecular analysis.

### *4.3.8 Leaf litter dry weight*

I transferred the entire remaining leaf litter sample to a dry, weighed envelope and recorded the combined weight of the sample and the envelope. In some samples, particularly those at high clay content sites, soil transferred into the bags leaving some samples notably covered in soil. These were recorded as such (21 samples). I chose not to wash samples before weighing. I had homogenised material prior to the decomposition experiment, and I believe it would have been difficult to prevent the loss of fine, crushed litter material during washing, or to measure the amount lost. Doing so for 809 samples would have taken more time than was available. Finally, washing material would have removed DNA from the surface of the leaf and modified sequencing results. The envelope was sealed and dried at 80°C for four days and reweighed, allowing for the calculation of the sample dry weight. I assumed that the removed material for molecular work had an identical ratio of wet to dry weight. I multiplied the percentage moisture content by the wet weight of the sample to calculate the dry weight of the original samples.

Some samples' dry weight increased to over 1g. There are several possible explanations. I suspect that the litter accumulated weight from soil coating the leaves which I believe was widespread within the data. In most cases, the weight increase was not large. But, at nearly all sites some litter bags contained greater than 1g in weight. Recalcitrant litter was more likely to show an increase in total weight. Additionally, the added weight to leaves of quick decomposing species likely meant that the amount of decomposition was underestimated. . Presumably, there may be some relationship between the surface area of the leaf material remaining and the area available to be covered by soil. At one site, nearly all leaf litter bags had substantially increased in weight over the study. The site likely has a high proportion of clay content that may have adhered to the litter and confounded measures of mass loss. Because of this, I have excluded this site from the models. After removing these samples only litter from 10 bags was recorded as containing leaf litter with a notable covering of soil. However, it is clear from the distribution of dry weights that multiple other litter samples were also covered in a disproportionately large quantity of soil but not indicated as such during dry weight calculations. I chose not to remove these outliers as there is no simple way to modify the data or remove samples that entirely resolves this problem. This reduced the precision at which we were able to estimate the remaining dry weight of leaf litter material but was

strongly influenced by site. If some sites are more likely to produce soil-covered leaf litter, this effect can be accounted for in a mixed-effect model structure. Removing data, or modifying the values, would reduce the ability of the model to account for site-level differences in soil properties that cause this to be accounted for in the model's random effect. Overall, this is an additional producer of noise in the data that reduced precision, but it accounted for in models.

### *4.3.9 Statistical methods*

*Constrained Correspondence Analysis of Ground Vegetation and Tree Communities*

To test H1, that woodland age category would influence vegetation communities but be a better predictor of ground cover than tree cover, I compared the percentage of inertia explained by woodland age category in canonical correspondence analyses of the ground and tree vegetation community matrices. Important components of these CCAs were brought forward to mixed-effect models as fixed effects to test H3 and H5. In producing this data, I calculated the importance of each tree genera for each site. I included the importance (density x frequency x dominance - as described above it incorporates measures of spatial evenness and basal area) of trees of the genera of the decomposing leaf species at the decomposition site in mixed-effect models to test H4.

To produce the descriptive statistics of vegetation communities, I first pooled the Braun-Blanquet ground vegetation data grouping rare cover classes (such as trees of different species, deadwood and leaf litter, bare earth and bare rock) to improve ordinations. As Braun-Blanquet classes are on an ordinal scale, but class bin size increases non-linearly, grouping cover types by summing or averaging scores produces misleading results. I pooled data assuming scores represented the minimum percentage cover for the score. In this way, cover scores for a quadrant may total more or less than 100% but may be pooled without overestimating the percentage cover. I pooled the data into the broad cover categories: exposed earth, plant litter, non-vascular plants, trees and saplings, debris, and vascular plants (e.g. forbs, including ancient woodland indicators such as *Mercurialis perennis*). I ordinated ground cover community data and tree community importance data separately using the package *vegan* (Oksanen *et al.*, 2019). Tree genera were ordinated based upon their importance at each site. As outlined above, importance is derived from the density (number

of individuals), evenness (number of points observed at) and basal area of trees and so may better represent the influence of a tree at a site than any of those individual measures alone. Although I have not directly measured the canopy structure or cover of sites, importance is influenced by tree DBH and so will differ for sites with similar composition but different canopy structures. I analysed community data using Constrained Correspondence Analysis (CCA), producing axes constrained by the age category of the woodland, and unconstrained axes explaining the remaining variation. I extracted scores scaled to best represent the position of sites, as I later used this site-level data in modelling decomposition. Constrained axes were selected for modelling if their eigenvalue was similar to that of important unconstrained axes. Unconstrained axes were selected when they explained large amounts of the variation and fit into an intelligible ecological gradient.

*Linear Mixed-Effect Modelling*

My remaining hypotheses relate to the effect of site characteristics on decomposition rate. I produced a minimal mixed effect model using the package *nlme* (Pinheiro *et al.*, 2019) to determine which characteristics were present and significant after model refinement. Effect significance was assessed using its 95% confidence intervals in line with longstanding advice regarding the misuse of p-values from mixed effect models (Bates, 2006). I produced candidate models using backward selection of variables, removing variables by comparing model AIC and testing for a non-significant difference in deviance; they included a random effect of site (or site compartment when the same woodland was sampled in different aged compartments) nested within triplicate. This recognises that unmeasured effects at each site will influence all samples within it, and accounts for similarities within triplicates of pH, soil texture, and other unmeasured variables correlated with those characteristics. I was able to produce similar models in the *lme4* package (Bates *et al.*, 2015), in which I was also able to consider more complex crossed random effect structures that included season and subject species as random effects, rather than fixed effects. However, testing *lme4* model significance, which requires the package *lmerTest* (Kuznetsova *et al.*, 2017), produced less conservative estimates of variable significance, although converging on a closely aligned model structure. I chose the more conservative approach of using the *nlme* constructed models. The global model contained sampling season; the leaf litter's tree species of origin; woodland age

category; the first component of a dbMEM of spatial autocorrelation; soil organic carbon concentration; C:N ratio; CCA1, CA1, and CA2 from the tree cover and importance CCA; CCA1 and CA1 from the vegetation CCA; the importance of the tree at the site; and whether the leaf litter material was noted as soil covered.

## 4.4 Results

### *4.4.1 Vegetation communities*

I first tested H1 - that woodland age category would be a better predictor of ground cover than tree cover with a pair of CCAs. The correspondence analysis of the ground vegetation data constrained by woodland age indicated that age category significantly affected ground cover ($F_{(2,24)}$ = 5.49, p = 0.001), explaining 31.4% of the total inertia (Figure 4.4). It produced two constrained axes, only the first of which explained a large amount of the variation in the data. The first axis (CCA1) separates woodland creation sites from mature and ancient semi-natural woodland, with high values present in sites dominated by grass or thistle, intermediate values for sites with large amounts of bramble, nettle and leaf litter, and low values for sites with ancient woodland indicator and woodland specialist plants such as dog's mercury, mosses, ramsons, and ferns. The second constrained axis (CCA2) differentiated ancient from mature woodland, with high values for ground vegetation more common at mature secondary woodland, such as nettles, ferns, and brambles and low values for ground vegetation including mosses, ramsons, dog's mercury, and ivy. However, this axis explained far less of the variation in the data (with an eigenvalue of only 0.021, compared to 0.293 for CCA1). Thirteen unconstrained axes were produced, but only the first explained large amounts of the variation (eigenvalue of 0.230). High values of CA1 were associated with sites high in bramble, leaf litter, ferns and with dense trees or saplings. Low values were associated with dog's mercury, mosses, nettles and exposed earth. The axis may be capturing variation between bright and open woodlands, and shady closed woods.

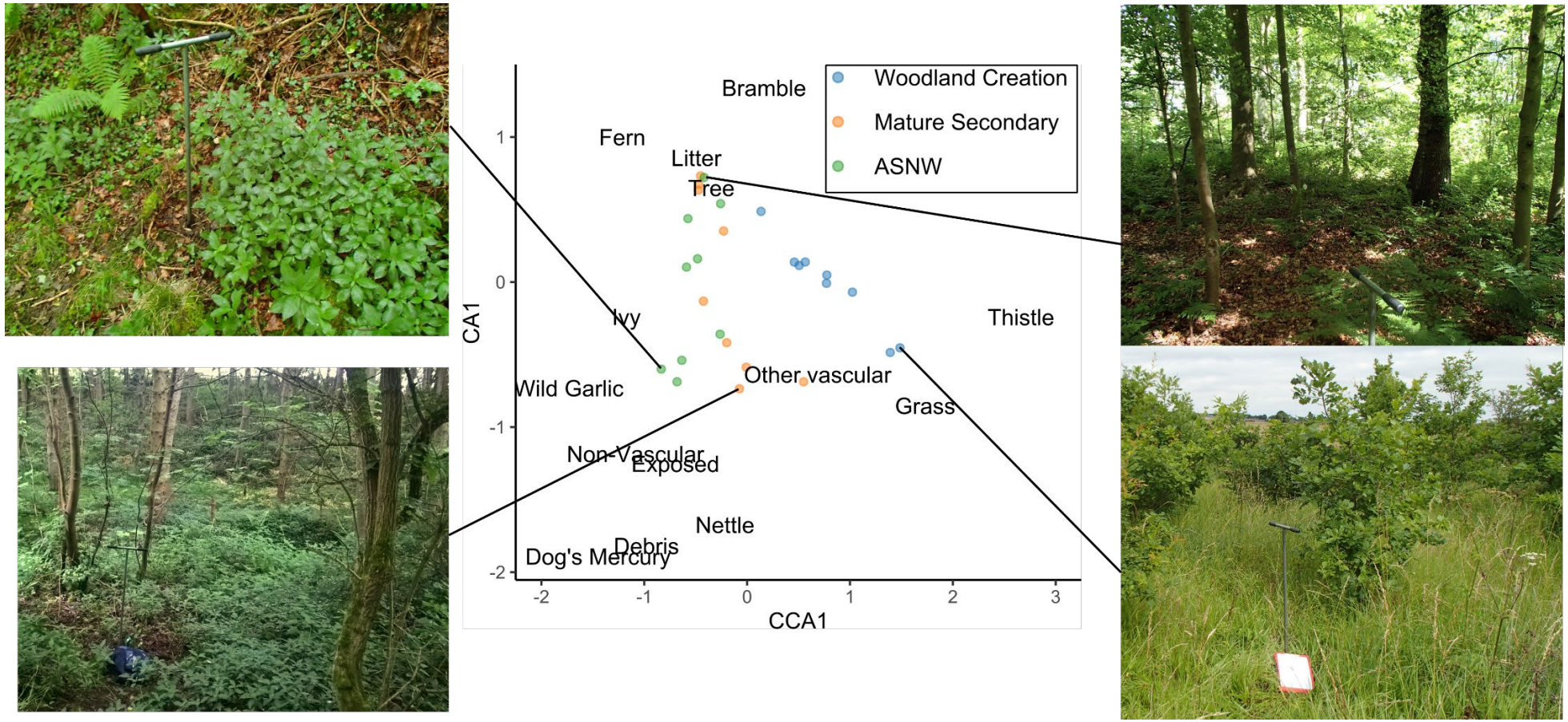


Figure 4.4 Key axes from a CCA on ground vegetation surveys of recently created, mature and ancient broadleaf woodlands in NE England. The data are collected from within 20m of buried leaf litter experiments. CCA1 is explained by the unordered woodland age category, CA1 is independent of the woodland age category.

Canonical Correspondence analysis of tree communities and importance identified two constrained axes explained by woodland age category. These axes indicated that age category significantly affected tree cover ($F_{(2,24)}$ = 1.90, p = 0.003) and explained 13.7% of the total inertia, though only the first explained a large proportion of the data (Figure 4.5). This axis (CCA1, with an eigenvalue of 0.313) was associated with alder, lime, hornbeam, elder, spruce, holly, and beech at low values, and poplar (*Populus*), guelder rose (*Viburnum*), rowan (*Sorbus*), cherry (*Prunus*), oak (*Quercus*), and ash (*Fraxinus*) at high values. Pine (Pinus) was also more common at these higher value sites. Several unconstrained axes were produced, two of which were carried forward for modelling the first unconstrained axis (CA1, with an eigenvalue of 0.4) captured variation related to wet woodlands, with low values associated strongly with willow (Salix) and alder (Alnus), and less so with pear (*Pyrus*), wild apple (*Malus*), hawthorn (*Crataegus*), and elder (*Sambucus*). Pine trees (*Pinus*) were also more common in these sites, perhaps reflecting the difficulty of extracting timber from river valleys. High values were associated with elms (*Ulmus*), limes (*Tillia*), hornbeam (*Carpinus*), and maples (*Acer*). The second axis (CA2, with an eigenvalue of 0.363) was associated with apple (*Malus*), elder (*Sambucus*), ash (*Fraxinus*), elms (*Ulmus*), birches (*Betula*), poplars (*Populus*), maples (*Acer*), and spruce (*Picea*) at high values. At low values, sites were more likely to contain guelder rose (*Viburnum*), beech (*Fagus*), holly (*Ilex*), oaks (*Quercus*), pear (*Pyrus*), cherry (*Prunus*), and rowan (*Sorbus*). Rare species such as guelder rose(*Viburnum*), pear(*Pyrus*), hornbeam(*Carpinus*), and lime (*Tillia*) were often strongly associated with an axis precisely because they were rare, each occurring at only a single site. Equally, planting preferences of organisations like the Woodland Trust have caused ancient-woodland indicators such as guelder rose to be common in these planted woods.

In agreement with H1, woodland age category explained a higher proportion of inertia in CCA models of understory vegetation communities (31.4%) than tree communities (13.7%).

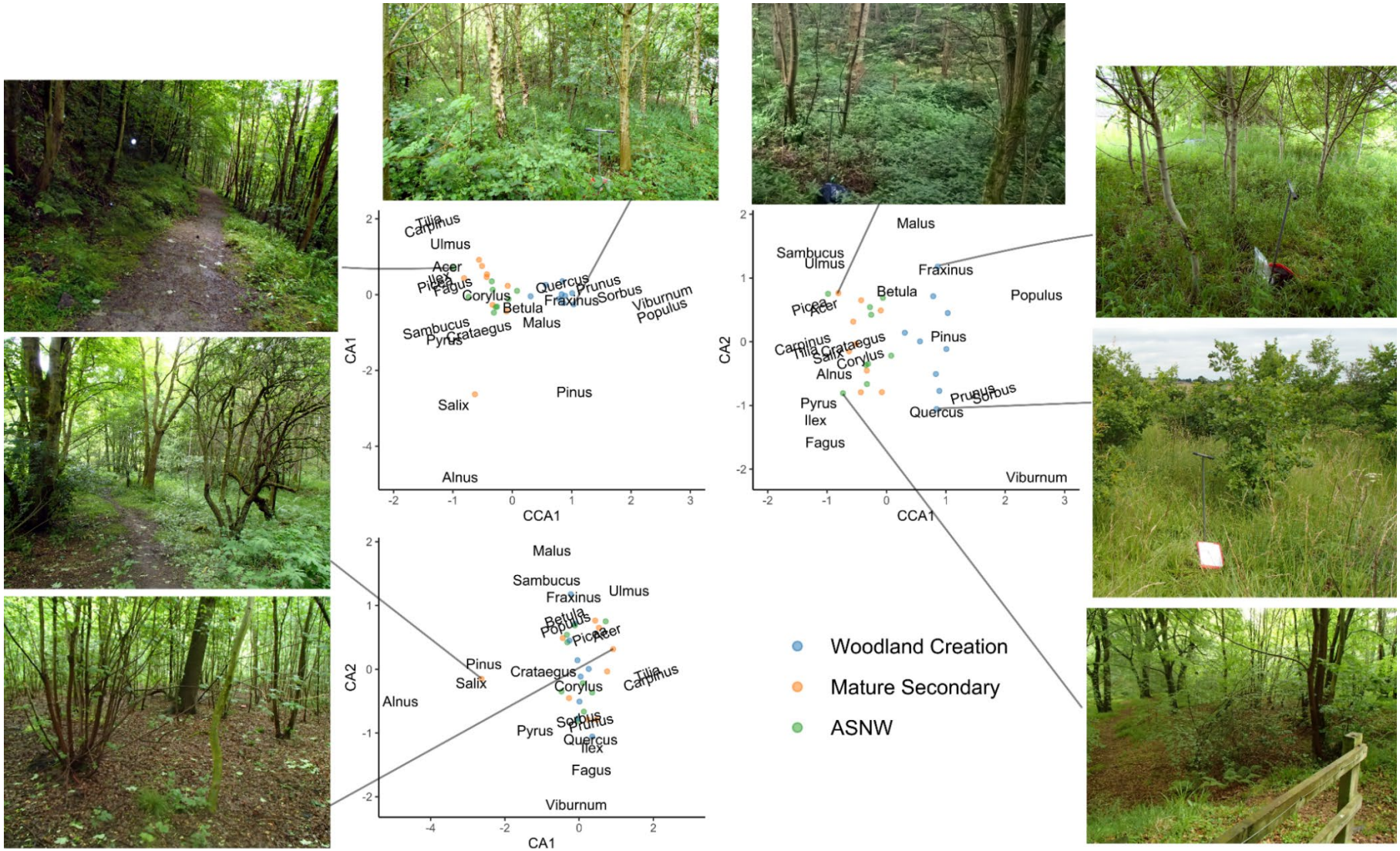


Figure 4.5 Key axes of a CCA on tree importance data collected within 20m of buried leaf litter. Constrained axis CCA1 is explained by woodland age category (recently created, mature, or ancient), and other unconstrained axes are independent of woodland age.

### *4.4.2 Decomposition models*

Multiple variables were identified as significantly affecting the remaining mass of leaf litter of litter bags placed in woodland sites in winter, when sampled in the spring and summer (Table 4.1). Expectedly, the sampling season ($F_{(1,748)} = 151$, $p < 0.0001$) and subject species ($F_{(4,748)} = 271$, $p < 0.0001$) significantly influenced the remaining leaf litter dry weight. H2 was not supported as woodland age category was not identified as a significant predictor of remaining leaf litter dry weight. However, after taking into account the other predictor, the model predicted a 0.010g lower weight of remaining leaf litter in bags collected from mature woodland than woodland creation and 0.047g less material in bags at ASNW than woodland creation sites but did so with low confidence (Figure 4.6). We can conservatively predict the increase in mass loss as a percentage increase by comparing these figures against the highest predicted mass loss for a species and season combination in woodland creation. If we assume a 1g initial weight and use the predicted second season elm mass loss - the most conservative reference level - that translates to 2% greater loss in mature woodland than woodland creation, and 9% greater loss at ASNW woodland than woodland creation (A predicted decline in elm mass of 0.59g in ASNW, 0.55g in mature woodland, and a loss of 0.54g in woodland creation all other variables being equal - Figure 4.6). For all other species and season combinations the predicted effect-size of woodland age category would be greater, and across the study averages at an approximate 4% and 16% increase in mass loss for all species across both seasons for mature and ancient woodland respectively. Additionally, this is likely an underestimation of the mass loss because of the weight of soil added to some leaves.

In agreement with H3, I identified an unconstrained axis of the tree community and importance CCA as a significant predictor of decomposition rate ($F_{(1,12)} = 23.2$, $p < 0.0004$). Figure 4.5 shows that sites with high tree CA2 (high in crab apple, ash, elder, elm, and birch; low in oak, beech, holly and guelder rose) had lower remaining leaf litter (higher decomposition rate). No evidence was found for H4 in that the importance of the tree genus at the site of decomposition was non-significant and was eliminated from models in the early stages of model refinement. In disagreement with H5, an unconstrained axis of the ground cover community ordination was a significant predictor of decomposition rate ($F_{(1,12)} = 8.24$, $p < 0.0141$). Sites with a high vegCA1 value (high in Dog's Mercury, Nettle, or non-vascular

plants; low in ferns, litter, and bramble) had lower remaining leaf litter (higher decomposition rate).

Table 4.1 Backwards selected model coefficients and significance values for a Wald test explaining the difference in mass from decomposed leaf litter bags collected in spring and summer of 2018 from broadleaf woodland in the NE of England.

| | Coefficient (change in mass) | DF | F value | P |
|---|---|---|---|---|
| Intercept | | (1,747) | 1548 | **<0.0001** |
| Woodland age | | (2, 12) | 3.65 | 0.0576 |
| Season | | (1, 747) | 151 | **<0.0001** |
| Litter's species | | (4, 747) | 271 | **<0.0001** |
| Leaf soiled | | (1, 747) | 26.38 | **0.0003** |
| Soil C:N | -0.003 (-0.005 to -0.001) | (1, 12) | 12.7 | **0.0039** |
| treeCA2 | -0.068 (-0.097 to -0.039) | (1, 12) | 25.97 | **0.0003** |
| vegCA1 | -0.064 (-0.109 to -0.019) | (1, 12) | 10.06 | **0.0081** |

Additionally, the soil C:N ratio of the site was a significant predictor of decomposition ($F_{(1,12)}$ = 11.8, p = 0.0050) and sites with a higher C:N ratio had less decomposition over the study period, with a predicted average difference of 0.09g (9% of initial weight) between the lowest and highest C:N ratio sites. Leaf litter bags that were identified to have notably soiled leaves had significantly higher weights (Table 4.1). Decomposition over the period was highest in elm, then ash, followed by hazel and oak, which generally had lost similar weights at each sampling period. Beech leaves decomposed the least over the time period.

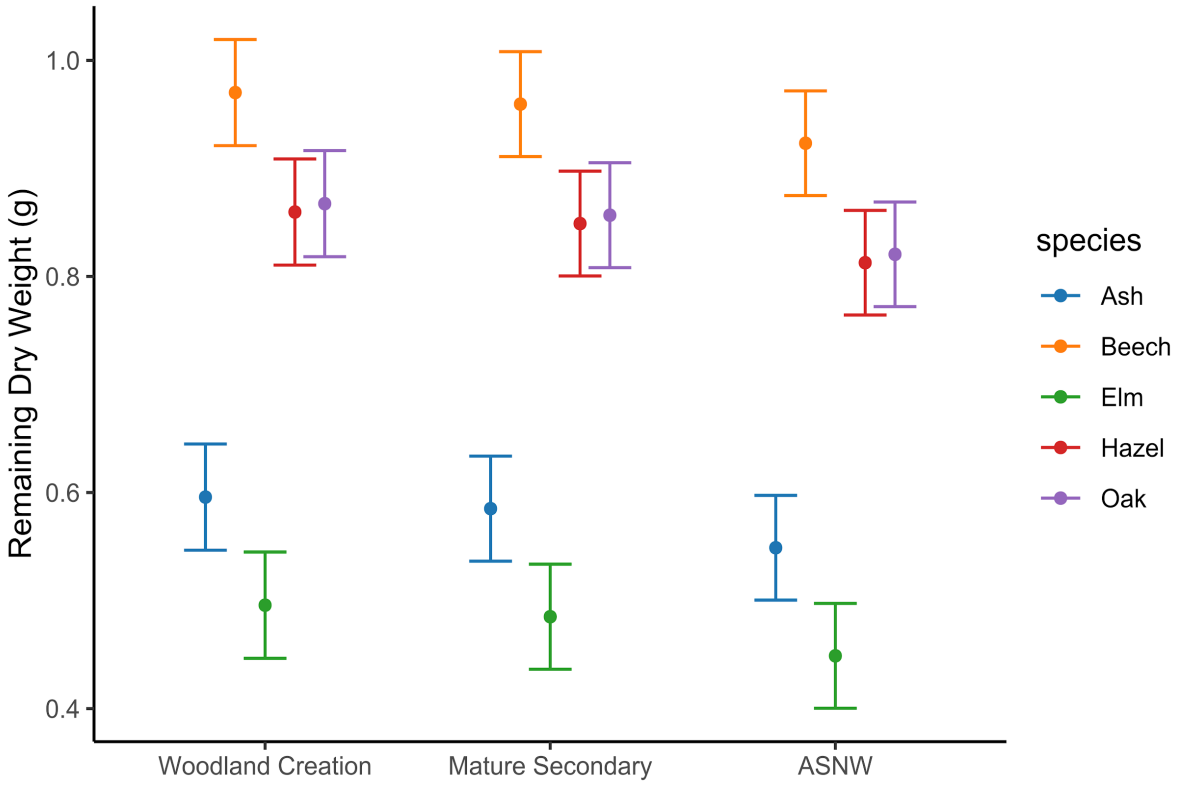


Figure 4.6 Predicted effects and 95% confidence intervals of tree species and woodland age category on remaining leaf litter weight after sampling litter bags in the summer season. The choice of season to be displayed does not affect the relationship between species or age categories as there are no interaction terms in the model. .

## 4.5 Discussion

### *4.5.1 Variations in decomposition across woodlands in the North East of England*

I have shown that vegetation community composition and soil C:N ratio of the decomposition site predict variation in decomposition rates in British woodland. Additionally, I present evidence that suggests long term successional processes that are independent of the tree, ground vegetation changes, or soil property changes, increase the rate of decomposition in woodlands. Further work with adjusted methods will be well placed to verify this effect. My first hypothesis was accepted: woodland age category was a significant predictor of understory and tree vegetation communities but explained a greater proportion of the variation in ground cover than tree cover. Woodland age category explained nearly a third of the variation present in the vegetation community data, but only 13.7% of the variation present in the tree community data. Generally, young woodland creation sites were less similar to mature woodland and ASNW than these categories were to each other along the

principal constrained axis, supporting H1. This likely reflects this greater influence of land-use history on tree communities when compared to ground cover communities which have shorter generation times and are less likely to have been intensively managed (Hermy & Verheyen, 2007). Trees in UK woodlands have a strong history of silviculture that still influences the composition and structure of trees in all woodlands (Rackham, 1986, 2012).

In line with H2, decomposition rate did differ by age category (but was not significant statistically, p = 0.0576); ASNW were 0.047g lighter than woodland creation sites across the study, and mature woodland were 0.01g lighter after accounting for influential site factors. If we assumed an initial mass of 1g, across the study the predicted reduction in leaf litter bag dry weight was higher by around 16% and 4% respectively in ASNW and mature woodland when compared to woodland creation. These results agree with previous work that has compared decomposition at young and primary Amazonian forests finding significant increases in decomposition amounts of seasonally sampled litter bags (Barlow *et al.*, 2007). Given that some litter bags may have additional weight from adhered soil, these are conservative estimates. These results could be substantiated in future work, either the addition of the third season's data to these models or independent studies on this phenomenon. If decomposition rates vary consistently with woodland age that would appreciably modify our understanding of carbon cycling in the UK. The study design effectively controlled for differences in soil chemistry between woodland age categories, but C:N ratio remained a significant predictor of decomposition rate which has been suggested in a previous meta-analysis of litter decomposition (Porre *et al.*, 2020). High soil C:N sites, which were usually young or mature sites rather than ASNW, had faster rates of decomposition. This may be due to the normally higher abundance of fungi at sites with high C:N ratio soils, caused by a generally lower requirement of fungi for nitrogen (Fierer *et al.*, 2009). Alternatively, in growth limited, high C:N ratio soils the addition of litter material with a low C:N ratio with comparatively labile nitrogen may have stimulated microbial activity.

As hypothesised, tree community composition was also a significant predictor of decomposition rate with an unconstrained axis of the tree vegetation significantly influencing remaining leaf litter weight (H4). This is not surprising given our understanding of how tree

diversity and composition influences decomposition via changes to soil stoichiometry and litter moisture that have knock-on effects on decomposer communities (Xiao *et al.*, 2019). However, existing work on decomposition tends to describe site vegetation in terms of alpha-diversity, or with community categories (e.g. “pure” species stands vs “mixed” stands), rather than taking a multivariate approach to vegetation gradients (Martini *et al.*, 2019). Contrary to our hypothesis, the importance (the index of a trees relative frequency, relative density, and relative basal area) of the decomposing leaf species at the site was not a significant predictor of decomposition rate (H5). As predicted, ground vegetation variation was a significant predictor of remaining leaf litter (H6), but only the differences in ground cover unrelated to woodland age category were included in the final model structure after model refinement. The role of understory vegetation is less well understood in the process of tree leaf decomposition, but in this case, it may also be indicating a natural gradient of temperature or moisture that is more closely tied to decomposition rate.

Multiple variables included in the global model confound the effect of woodland age, though none of them are present in the minimal model. As such it remains difficult to determine the mechanism by which woodland age might influence the decomposition rate. If decomposer communities continue to adapt to woodland leaf litter inputs over the centuries, we would expect to see increased decomposition rates in older woodland categories over younger ones. Depending on your interpretation of the HFA and SMI hypotheses this could be limited to the rates of decomposition of tree species present at a site (Ayres *et al.*, 2009), or of leaf litter with similar phenology (Freschet *et al.*, 2012). However, there was low confidence in these estimates and wide confidence intervals. In reality, the decomposition rate may not vary with woodland age or may decrease as woodlands get older. Without more precise measurements of leaf litter loss, we cannot reject the null hypothesis that decomposition rate does not vary with woodland age. Though these effect sizes may appear small in absolute mass, a difference of this size could account for a few percentage points of previously unexpected variance worldwide, if identified as significant. As such they might meaningfully and appreciably alter models of carbon cycling at a global or national level in such a way that has not been considered previously (Jandl *et al.*, 2007). Further work into carbon cycle modelling and experimental research should incorporate more information about site age and vegetation into analysis and study design.

The importance of the subject leaf species at the site - which takes into account the density of the tree species, the evenness throughout the site, and the size of the tree - did not impact decomposition rate. This may indicate that the presence of a tree species at a site does not influence decomposer community assembly to favour its leaf litter decomposition. If tree importance does not relate to preferential decomposition of its litter then the adaption of a decomposer community to specific important species at a site mat a site may not be the mechanism by which the HFA is produced effect (Ayres *et al.*, 2009). Inconsistency support for the HFA hypothesis is not unusual, and the strength and direction of the effect has been associated with tree cover, litter quality, and the dominance of the study subject species at sites (Veen, Freschet, *et al.*, 2015). For nearly 45% of the leaf litter bags collected, the trees of the same species as those that produced the litter were not observed in the 20m surrounding the study area, and even when present, these trees tended to be of low importance at the site. A better test of the effect of tree species importance at a site on the decomposition of leaves from the same species could be achieved by setting out to select sites with a greater range of importance values for the subject tree species (i.e. sites that vary in the subject's density, evenness, and basal area). By selecting beech woodlands, hazel coppice stands, or broadleaf plantations one could design such an experiment, but it was not the main purpose of this study.

I aimed to control the effects of soil properties on decomposition rate and microbial composition by grouping sites into triplicates with similar soil properties. Soil carbon and C:N ratio did not significantly vary with age categories, although C:N ratio did vary within triplicates they did not do so consistently or significantly. Generally, this method of grouping sites *a posteriori* to minimise the effects of confounding variables worked effectively.

### *4.5.2 Vegetation gradients and their impacts on decomposition*

Five vegetation gradients were incorporated into my models of decomposition. Two constrained axes - representing the primary explanatory axis of the tree cover and ground cover data explained by woodland age category - and three unconstrained axes that are intended to capture natural gradients within woodlands. I am aware of no other

decomposition experiments that have sought to describe site vegetation cover using multivariate methods then incorporate these values into decomposition models. Although simply describing vegetation with an NMDS but not incorporating axes into models has been done (Larkin *et al.*, 2014), as has describing the similarity of the sites community to the tree species that produced the litter using trait-based distance indexes (Jewell *et al.*, 2015). This may be because the method required large site-level replication. Studies sometimes include tree species diversity in models (Larkin *et al.*, 2014; Trogisch *et al.*, 2016; Joly *et al.*, 2017), but it is more common to see categorical or species-specific descriptions. Multivariate descriptions of diversity offer ways of describing diversity that do not depend on the *a priori* beliefs of researchers as to which aspects of diversity are important. Diversity on its own is not a mechanism of determining decomposition rate; certain aspects of diversity, certain tree species, may have a disproportionate influence on the decomposition rate (Joly *et al.*, 2017) and multivariate methods are capable of describing diversity gradients in the selected sites and relating them to changes in decomposition.

In this study, the effects of constrained axes may be thought of as indirect effects of woodland ageing. Had they appeared in models, it would suggest that all or part of the effect of woodland age on decomposition resulted from changes in this aspect of woodland ecology. Unconstrained axes appearing in the minimal model may either directly influence the decomposition of leaf litter as a result of soil inputs from the plants, or may reveal natural gradients of temperature, moisture, or sites characteristics that impact decomposition rate but were unmeasured. The low first unconstrained axis of the treeCCA (treeCA1) was associated with riparian trees - willows and alder - that grow well in wet soils. The axis may be capturing differences in soil moisture between sites. The underlying gradient of the second unconstrained tree axis (treeCA2) is less apparent. However, low values of treeCA2 included trees with difficult to decompose leaves - oak, beech, holly, whereas high values indicated trees with faster decomposing leaves - e.g. ash, elm. This axis (treeCA2) appeared in the final model. Sites with more important (i.e. more frequent, evenly spaced, larger) trees that had easy to decompose leaves had higher decomposition rates. This result may agree with a modification of the HFA effect, the Substrate-quality Matrix-quality Interaction (SMI) effect (Freschet *et al.*, 2012). This SMI hypothesis suggests that litters decompose fastest at sites with natural litter input of similar quality due to the availability of decomposers already

adapted to similar litter. By definition, the greatest amount of decomposition in my study was experienced by quickly decomposing litter, it may be expected under the SMI hypothesis that the decomposition amount of this labile material was lessened at sites dominated with recalcitrant litter input (Freschet *et al.*, 2012). However, no species interaction effects were supported in models indicating that the influence of treeCA2 on decomposition was not affected by litter decomposition recalcitrance, which may have been supposed from Frechet *et al.*'s description of the SMI effect. If treeCA2 is an important predictor of decomposer community structure this may substantiate the effect. Otherwise, it is possible that this association is brought about by a separate mechanism. Being an unconstrained axis, this effect is not likely to affect decomposition via manipulating the microclimate of the woodland as extreme values along this axis occur in young woodlands with juvenile trees as often as in older woodlands with mature trees and closed canopies.

The unconstrained ground vegetation axis (vegCA1) largely explained differences between lightly shaded, moist woodlands at low values - containing many woodland indicators, grass, and nonvascular plants; and closed canopy, litter covered sites at high values - high in bramble, leaf litter, dense with saplings and large adult trees. These closed sites experience higher decomposition rates, which agrees with some previous work (Joly *et al.*, 2017). Naively, one might expect less shaded woodlands to have higher decomposition rate due to potentially higher daytime temperatures, but alterations to litter decomposition microclimate caused by canopy closure are not well described. The increase in decomposition at more shaded sites may be due to more stable environmental conditions or may be acting via a different mechanism such as a priming effect on decomposer communities from the accumulated leaf litter. The vegCA1 variable has a slight negative correlation with treeCA2, as leaf litter normally accumulates at sites with tree species that produce difficult to decompose leaf litter, such as beech or oak. However, as these effects act in the same direction, but have a low, negative correlation, their impact on decomposition likely represents different mechanisms. These are speculative explanations as the measures do not directly include information about canopy structure and canopy cover. But as the basal area of trees is included in the calculation of tree importance at a site, the structure of these woodlands affects the CCA. From my experience surveying the sites I suspect that canopy cover and structure were strongly related to the woodland age category as has been found in previous studies (Trogisch *et al.*, 2016). This

would mean that canopy effects are likely to be incorporated into the age category or treeCCA1 model variables, only the former of which was included in the most parsimonious model structure.

Overall, this suggests a higher decomposition rate at closed-canopy sites, perhaps indicating that these sites provide a more hospitable environment for decomposers. Additionally, decomposition rate was higher at sites dominated by trees with fast decomposing leaf litter, irrespective of woodland age. Finally, it suggests that, although changes in tree and ground vegetation caused by succession are present and will likely contribute to the effect of woodland age on decomposition rate, neither is the main mechanism by which woodland age impacts decomposition.

### *4.5.3 Assessing site selection and site triplicates*

Selecting sites with similar pH and soil texture to form triplicates usefully and effectively controlled for differences between woodland age categories, with non-significant differences in organic carbon, C:N ratio, and pH between age categories. These variables may still produce within age category effects but would not be confounding with the main study question. Challenges of this approach surround finding appropriate sites to group together, which will vary region to region. In the northeast of England, woodland creation and mature woodland sites of over 2 hectares in size are relatively easy to locate and often publicly accessible (working with the permission of the landowner is still required but many publicly accessible sites are owned by the same organisations easing accessibility). Ancient woodland sites are less common, often fragmented, often smaller, often optimally shaped for avoiding edge effects, due to centuries of pressure from land-use change (Rackham, 2008). It is fair to say that the typical Northumbrian ancient woodland exists in a wonky strip, in a valley, within a few hundred metres of a stream or river, on a field margin. At these sites there are limited spaces that are more than 50m from a site edge and that are safely accessible in all seasons. If edge effects are to be common throughout the entirety of the majority of these ancient woodlands, does that cease being an edge effect and become the more accurate representation of Britain's ancient woodland stocks? Recent work suggests that edge effects of soil properties and microbial diversity appear to be extremely spatially limited (<15m) in

British woodlands (Błonska *et al.*, 2020). Selecting from these limited potential ancient sites such that they can be paired with similar mature and creation sites further limits the possible study locations. I had intended to study 30 sites in ten triplicates but was unable to create a tenth triplicate from the available sites. Researchers must consider their own site availability, and time constraints during study design before attempting to group sites prior to analysing them, but the benefits can be substantive. These methods may require more advanced statistical methods to analyse that can be off-putting, but they are not much more advanced than accounting for site pseudoreplication via random effects that should be undertaken in many studies as it is.

### *4.5.4 Strengths and weaknesses of the methodology*

To achieve the goals of this study, I needed to decompose leaves from multiple species, with high degrees of replication, while handling samples with care so that decomposer DNA could be extracted without contamination. Above, I advocated for additional leaf litter experiments with high degrees of site and within-site replication, modelled to account for pseudo-replication to improve the statistical and explanatory power of decomposition studies. One difficulty anecdotally experienced in leaf litter studies, especially ones where leaf litter bags are buried, is the loss of a large proportion of leaf litter bags. Decomposition experiments seek to describe subtle effects and as much statistical power as possible. By securing leaf litter bags to a rigid, plastic frame I managed to recover all the leaf litter bags I placed out. I used site measurements photographs, drawn maps, and GPS coordinates. With these methods I have successfully completed a large-scale, fully replicated decomposition experiment, and recommend them for future work.

The greatest drawback of my technique has been the lack of precision in calculating leaf-litter loss. Sampled leaf litter sometimes gained mass over the course of the experiment, due to the addition of soil onto the leaf material. I have not calculated mass loss but remaining leaf weight, which is not uncommon in the literature (e.g. Barlow *et al* 2007). Samples that were identified as especially soil covered weighed a predicted 0.24g more than their counterparts. Leaves with less notable soil cover will have added noise to the dataset. Differences in soil that might influence how much material adhered to leaves will vary between but not within sites.

By not removing high weight data from the model, the random effect from the site has more information regarding this noise to build estimates from. This would not be possible if high weight samples were removed.

Some studies have cleaned leaf material to remove soil (Barlow *et al.*, 2007; Sundsdal *et al.*, 2020). But these methods are missing from the many other articles leaving some ambiguity as to whether cleaning is commonly undertaken and how thorough this cleaning is. Cleaning was not an option in my experiment. To more accurately simulate natural decomposition, while excluding insects and molluscs from the leaf litter bags, I homogenised the leaf material. Leaf litter must also be homogenised to avoid bias when collecting DNA from the leaf. If I had washed material, I would risk losing fine leaf litter fragments, additionally I risked contaminating or washing away the eDNA signature of the sample. Future experiments may benefit from modifying this experimental design to reduce the noise in the data attributable to soil clinging to the leaf.

The question remains: was this approach suitable for answering the questions of this study? More precision in estimating leaf litter weight change may have identified the differences between woodland age categories in leaf litter loss as significant. However, this study offers a direction for future work. A non-molecular study, or a study that could overcome this issue of precision will be well placed to establish if the differences suggested in this study regarding woodland age are significant. Despite issues in precision, I was able to identify impactful, yet subtle effects of vegetation and soil characteristics on decomposition, as well as unambiguous differences in decomposition rate by subject species. As such, I believe that most aspects of my approach represent important improvements over a more traditional, lower-replication design.

### *4.5.5 Towards a mechanistic understanding of decomposition*

Producing mechanistic explanations of key soil processes is essential to improving our understanding of carbon cycling at local, landscape, and global scales (Bani *et al.*, 2018; Porre *et al.*, 2020). Traditional approaches of conducting leaf litter experiments and assessing site

characteristics have yielded key insights, particularly highlighting fungi as the key agents of decomposition with bacteria an important additional role (Schneider *et al.*, 2012). The quality and recalcitrance of the input material to decomposition has also been explored extensively (Lummer *et al.*, 2012; Veen, Freschet, *et al.*, 2015; Veen, Sundqvist, *et al.*, 2015). The mechanism by which site characteristics are believed to alter decomposition rate is often suggested to be via shaping decomposer community composition (Bani *et al.*, 2018) but varying degrees of evidence are presented. While it is clear that these properties do modify soil communities (Griffiths *et al.*, 2011; Seaton *et al.*, 2020), their effect of their drivers of decomposition drivers are often left unexplored, particularly those of microclimates of temperature and moisture. I raised concerns that insufficient attention is being paid to tree community composition, rather than diversity, which means that effects of tree identity on decomposition rate may be being missed (Martini *et al.*, 2019). The structure of woodland canopies will also affect these microclimates, as well as altering shade (Joly *et al.*, 2017). As the evidence of powerful photolytic decomposition effects mounts in aboveground litter decomposition experiments (Berenstecher *et al.*, 2020), it becomes increasingly clear that decomposition experiments must better control for site factors of soil properties, woodland structure, and tree composition in their studies; or collect data on these variables while increasing across-site replication in order to avoid misleading results.

The approach taken in this study, of exploring vegetation data using multivariate statistics has identified natural gradients that exert significant influences on decomposition rate, that can be accounted for while trying to answer a central question around decomposition. This has allowed for speculation as to how vegetation differences influence decomposition, above that which may be achieved by examining site diversity alone. Much of the same can be said regarding the examination of soil properties at a site and incorporating this information into models. Notably, the decomposition of tree litter frequently ignores the effects of understory and ground cover on decomposition. These may be directly influencing decomposer communities or may manipulate the woodland microclimate. Studies that fail to capture ground cover data may also fail to capture data on important drivers of decomposition.

Traditional decomposition experiments will benefit from molecular approaches that can directly link ecological processes to biodiversity. Although next-generation molecular approaches are becoming common in soil diversity studies, decomposition studies are less likely to make use of them. Early implementers have demonstrated fine-scale effects of fungal community composition on decomposition rates (Asplund *et al.*, 2018); have highlighted that complex interactions and feedbacks between soil properties, tree communities, litter quality, and microbial diversity underpin decomposition in woodland environments (Xiao *et al.*, 2019); and have called into question the assumption of functional equivalence in microbial decomposer communities varied soil inoculums (Strickland *et al.*, 2009). However, the capacity of high-throughput, next-generation molecular approaches to produce voluminous outputs, allowing for high degrees of replication in ambitious study designs has not been fully taken advantage of. Future experiments can push boundaries to sequence hundreds or thousands of samples to quantify subtle effects on diversity with high statistical power. Ambitious projects with high degrees of replication within and across sites are critical to understanding the spatial scales at which soil processes play out in natural and managed systems (Bradford *et al.*, 2016).

Future litter experiments may benefit from the site selection and description methods employed here. I conducted initial surveys of potential sites and selected them to control for the confounding variables suspected to importantly influence soil processes, in this case pH and soil texture. I then made sure to include descriptions of site characteristics that could not be controlled for in models. Decomposition is a complex process with numerous drivers, many of which are poorly understood. To improve our understanding of the mechanisms of decomposition, fewer assumptions must be made as to the importance of unexpected drivers on decomposition rate, specifically regarding soil chemistry, tree composition, understory vegetation, and woodland structure.

### *4.5.6 Conclusions*

In a novel decomposition experiment, I have demonstrated how a highly replicated leaf litter project can reveal subtle drivers of decomposition rate, including effects of soil properties, tree community composition and ground cover. I provide evidence that is strongly suggestive

of differing rates of decomposition in differently aged woodlands that provides a fruitful direction for future studies. I demonstrated the effectiveness of moving beyond alpha diversity measures of tree composition to improve our understanding of the influences on decomposition by including multivariate descriptions of tree communities in my modelling approach. Additionally, I highlighted the benefits of including understory vegetation data in decomposition models due to the potential for these data to reveal direct and indirect drivers of decomposition rate in these systems. Although the method presented led to imprecision in determining decomposition rate that may have reduced its capacity to detect results, the approach still proved to be effective. Notably, it improves the recovery of litter bags that can cost an experiment its statistical power. Future work can easily build upon this approach to asking and resolving specific questions surrounding decomposition while altering the handling of samples to overcome any specific issues encountered here. Overall, these results are suggestive of differing rates of key ecological processes in woodlands at different stages of succession, adding to our understanding of ancient habitats in managed landscapes.

### *4.5.7 Synthesis and applications*

My previous chapters demonstrate many nutrient pools of interest to policy makers, such as carbon stocks, reach similar levels to ancient woodland in comparatively short time frames. However, using a narrow set of metrics to judge the restoration of woodland functioning risks missing significant differences between even mature secondary woodland and ASNW. This chapter adds credence to this view by suggesting a possible increase in decomposition rate in older woodlands. Although I have not demonstrated that this result is significant, these results suggest that more work must be done to evaluate ecosystem service provision in similar woodlands of differing ages.

Replicating mature or ancient woodlands, or mitigating their loss to land-use change, is looking increasingly infeasible the better we understand how these ecosystems operate. Time changes an ecosystem by causing the accumulation or depletion of resource pools, by modifying its structure via vegetation growth, and through the neutral and niche dynamics of biotic interactions. Time can only move in one direction and so cannot be thought of in the same way as other ecological gradients when making management decisions. This work agrees

with general conservation concepts, that ecosystem age measurably alters ecosystem function and therefore an ecosystem's value in ways that may be impossible to replicate at short temporal scales.

The vegetation of a site and its soil's carbon to nitrogen ratio, and potentially other soil properties not included in this study, are significantly associated with the amount decomposition occurring over the spring and summer seasons in woodlands and possibly other seasons too. This suggests that management and planting decisions that modify the composition of woodland trees and the understory composition will lead to changes in key soil processes over time.

**4.6 References**

Aerts, R. (1997) Climate, Leaf Litter Chemistry and Leaf Litter Decomposition in Terrestrial Ecosystems: A Triangular Relationship. *Oikos* . 79 (3), 439–449.

Arias, M.E., Gonzalez-Perez, J.A., Gonzalez-Vila, F.J. & Ball, A.S. (2005) Soil health -- a new challenge for microbiologists and chemists. *International microbiology: the official journal of the Spanish Society for Microbiology*. 8 (1), 13–21.

Asplund, J., Kauserud, H., Bokhorst, S., Lie, M.H., Ohlson, M. & Nybakken, L. (2018) Fungal communities influence decomposition rates of plant litter from two dominant tree species. *Fungal ecology*. 321–8.

Ayres, E., Steltzer, H., Simmons, B.L., Simpson, R.T., Steinweg, J.M., Wallenstein, M.D., Mellor, N., Parton, W.J., Moore, J.C. & Wall, D.H. (2009) Home-field advantage accelerates leaf litter decomposition in forests. *Soil biology & biochemistry*. 41 (3), 606–610.

Baldrian, P. (2017) Forest microbiome: diversity, complexity and dynamics. *FEMS microbiology reviews*. 41 (2), 109–130.

Bani, A., Pioli, S., Ventura, M., Panzacchi, P., Borruso, L., Tognetti, R., Tonon, G. & Brusetti, L. (2018) The role of microbial community in the decomposition of leaf litter and

deadwood. *Applied soil ecology: a section of Agriculture, Ecosystems & Environment*. 12675–84.

Bargali, S.S., Singh, S.P. & Singh, R.P. (1993) Patterns of weight loss and nutrient release from decomposing leaf litter in an age series of eucalypt plantations. *Soil biology & biochemistry*. 25 (12), 1731–1738.

Barlow, J., Gardner, T.A., Ferreira, L.V. & Peres, C.A. (2007) Litter fall and decomposition in primary, secondary and plantation forests in the Brazilian Amazon. *Forest ecology and management*. 247 (1), 91–97.

Bates, D. (2006) *[R] lmer, p-values and all that*. [Online] [online]. Available from: https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html (Accessed 18 February 2020).

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, Articles*. 67 (1), 1–48.

Berenstecher, P., Vivanco, L., Pérez, L.I., Ballaré, C.L. & Austin, A.T. (2020) Sunlight Doubles Aboveground Carbon Loss in a Seasonally Dry Woodland in Patagonia. *Current biology: CB*.

Błonska, E., Lasota, J., da Silva, G.R.V., Vanguelova, E., Ashwood, F., Tibbett, M., Watts, K. & Lukac, M. (2020) Soil organic matter stabilization and carbon-cycling enzyme activity are affected by land management. *Annals of Forest Research*. 63 (1), 71–86.

Borcard, D. & Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological modelling*. 153 (1), 51–68.

Bradford, M.A., Berg, B., Maynard, D.S., Wieder, W.R. & Wood, S.A. (2016) Understanding the dominant controls on litter decomposition. *The Journal of ecology*. 104 (1), 229–238.

Brunet, J., Valtinat, K., Mayr, M.L., Felton, A., Lindbladh, M. & Bruun, H.H. (2011) Understory succession in post-agricultural oak forests: Habitat fragmentation affects forest specialists and generalists differently. *Forest ecology and management*. 262 (9), 1863–1871.

Buresova, A., Kopecky, J., Hrdinkova, V., Kamenik, Z., Omelka, M. & Sagova-Mareckova, M.

(2019) Succession of Microbial Decomposers Is Determined by Litter Type, but Site Conditions Drive Decomposition Rates. *Applied and environmental microbiology*. 85 (24), .

Chuan, X., Carlyle, C.N., Bork, E.W., Chang, S.X. & Hewins, D.B. (2018) Long-Term Grazing Accelerated Litter Decomposition in Northern Temperate Grasslands. *Ecosystems* . 21 (7), 1321–1334.

Coûteaux, M.M., Bottner, P. & Berg, B. (1995) Litter decomposition, climate and liter quality. *Trends in ecology & evolution*. 10 (2), 63–66.

Dahdouh-Guebas, F. & Koedam, N. (2006) Empirical estimate of the reliability of the use of the Point-Centred Quarter Method (PCQM): Solutions to ambiguous field situations and description of the PCQM+ protocol. *Forest ecology and management*. 228 (1), 1–18.

Dean, C., Kirkpatrick, J.B., Doyle, R.B., Osborn, J., Fitzgerald, N.B. & Roxburgh, S.H. (2020) The overlooked soil carbon under large, old trees. *Geoderma*. 376114541.

DEFRA (2018) *A Green Future: Our 25 Year Plan to Improve the Environment*.

DEFRA (2006) *Technical Advice Note 52 – Soil texture*.

Doran, J.W. & Zeiss, M.R. (2000) Soil health and sustainability: managing the biotic component of soil quality. *Applied soil ecology: a section of Agriculture, Ecosystems & Environment*. 15 (1), 3–11.

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N. & Wagner, H.H. (2020) *adespatial: Multivariate Multiscale Spatial Analysis*.

Dray, S., Legendre, P. & Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological modelling*. 196 (3), 483–493.

Falconer, J.G., Wright, J.W. & Beall, H.W. (1933) The Decomposition of Certain Types of Forest Litter Under Field Conditions. *American journal of botany*. 20 (3), 196–203.

Fierer, N. (2017) Embracing the unknown: disentangling the complexities of the soil

microbiome. *Nature reviews. Microbiology*. 15 (10), 579–590.

Fierer, N., Strickland, M.S., Liptzin, D., Bradford, M.A. & Cleveland, C.C. (2009) Global patterns in belowground communities. *Ecology letters*. 12 (11), 1238–1249.

Forest Service (2006) *NI Forestry - a strategy for sustainable growth*.

Frelich, L.E. (2002) *Forest Dynamics and Disturbance Regimes: Studies from Temperate Evergreen-Deciduous Forests*. Cambridge University Press.

Freschet, G.T., Aerts, R. & Cornelissen, J.H.C. (2012) Multiple mechanisms for trait effects on litter decomposition: moving beyond home-field advantage with a new hypothesis: Substrate-matrix quality interactions in decay. *The Journal of ecology*. 100 (3), 619–630.

Gessner, M.O., Swan, C.M., Dang, C.K., McKie, B.G., Bardgett, R.D., Wall, D.H. & Hättenschwiler, S. (2010) Diversity meets decomposition. *Trends in ecology & evolution*. 25 (6), 372–380.

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M. & Whiteley, A.S. (2011) The bacterial biogeography of British soils. *Environmental microbiology*. 13 (6), 1642–1654.

Helliwell, D.R., Buckley, G.P., Fordham, S.J. & Paul, T.A. (1996) Vegetation succession on a relocated ancient woodland soil. *Forestry* . 69 (1), 57–74.

Hermy, M. & Verheyen, K. (2007) Legacies of the past in the present-day forest biodiversity: a review of past land-use effects on forest plant species composition and diversity. *Ecological research*. 22 (3), 361–371.

HS2 Ltd (2017) *High Speed Two Phase One: London-West Midlands Ancient Woodland Strategy*.

Jandl, R., Lindner, M., Vesterdal, L., Bauwens, B., Baritz, R., Hagedorn, F., Johnson, D.W., Minkkinen, K. & Byrne, K.A. (2007) How strongly can forest management influence soil carbon sequestration? *Geoderma*. 137 (3), 253–268.

Jewell, M.D., Shipley, B., Paquette, A., Messier, C. & Reich, P.B. (2015) A traits-based test of the home-field advantage in mixed-species tree litter decomposition. *Annals of botany*. 116 (5), 781–788.

Jiao, S., Chen, W., Wang, J., Du, N., Li, Q. & Wei, G. (2018) Soil microbiomes with distinct assemblies through vertical soil profiles drive the cycling of multiple nutrients in reforested ecosystems. *Microbiome*. 6 (1), 146.

Joly, F.-X., Milcu, A., Scherer-Lorenzen, M., Jean, L.-K., Bussotti, F., Dawud, S.M., Müller, S., Pollastrini, M., Raulund-Rasmussen, K., Vesterdal, L. & Hättenschwiler, S. (2017) Tree species diversity affects decomposition through modified micro-environmental conditions across European forests. *The New phytologist*. 2141281–1293.

Kurganova, I.N., Lopes de Gerenyu, V.O., Mostovaya, A.S., Ovsepyan, L.A., Telesnina, V.M., Lichko, V.I. & Baeva, Y.I. (2018) Effect of Reforestation on Microbiological Activity of Postagrogenic Soils in European Russia. *Contemporary Problems of Ecology*. 11 (7), 704–718.

Kuznetsova, A., Brockhoff, P. & Christensen, R. (2017) lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, Articles*. 82 (13), 1–26.

Larkin, D.J., Steffen, J.F., Gentile, R.M. & Zirbel, C.R. (2014) Ecosystem Changes Following Restoration of a Buckthorn-Invaded Woodland. *Restoration Ecology*. 22 (1), 89–97.

Lindahl, B.D., Ihrmark, K., Boberg, J., Trumbore, S.E., Högberg, P., Stenlid, J. & Finlay, R.D. (2007) Spatial separation of litter decomposition and mycorrhizal nitrogen uptake in a boreal forest. *The New phytologist*. 173 (3), 611–620.

Lladó, S., López-Mondéjar, R. & Baldrian, P. (2017) Forest Soil Bacteria: Diversity, Involvement in Ecosystem Processes, and Response to Global Change. *Microbiology and molecular biology reviews: MMBR*. 81 (2), e00063–16.

López-Mondéjar, R., Voříšková, J., Větrovský, T. & Baldrian, P. (2015) The bacterial community inhabiting temperate deciduous forests is vertically stratified and undergoes seasonal dynamics. *Soil biology & biochemistry*. 8743–50.

Lummer, D., Scheu, S. & Butenschoen, O. (2012) Connecting litter quality, microbial community and nitrogen transfer mechanisms in decomposing litter mixtures. *Oikos* . 121 (10), 1649–1655.

Martini, F., Xia, S.-W., Yang, X. & Goodale, U.M. (2019) Small-scale and multi-species

approaches for assessing litter decomposition and soil dynamics in high-diversity forests. *Applications in plant sciences*. 7 (4), e01241.

Natural England (2011) *Ancient Woodland Inventory*.

N'Dri, J.K., Guéi, A.M., Edoukou, E.F., Yéo, J.G., N'Guessan, K.K. & Lagerlöf, J. (2018) Can litter production and litter decomposition improve soil properties in the rubber plantations of different ages in Côte d'Ivoire? *Nutrient Cycling in Agroecosystems*. 111 (2), 203–215.

O'Connell, A.M. (1988) Decomposition of leaf litter in karri (Eucalyptus diversicolor) forest of varying age. *Forest ecology and management*. 24 (2), 113–125.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. & Wagner, H. (2019) *vegan: Community Ecology Package*.

Parliamentary Office of Science and Technology (2014) *Ancient Woodland*.

Petersen, H. & Luxton, M. (1982) A Comparative Analysis of Soil Fauna Populations and Their Role in Decomposition Processes. *Oikos* . 39 (3), 288–388.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team (2019) *nlme: Linear and Nonlinear Mixed Effects Models*.

Porre, R.J., van der Werf, W., De Deyn, G.B., Stomph, T.J. & Hoffland, E. (2020) Is litter decomposition enhanced in species mixtures? A meta-analysis. *Soil biology & biochemistry*. 145107791.

Purahong, W., Wubet, T., Lentendu, G., Schloter, M., Pecyna, M.J., Kapturska, D., Hofrichter, M., Krüger, D. & Buscot, F. (2016) Life in leaf litter: novel insights into community dynamics of bacteria and fungi during litter decomposition. *Molecular ecology*. 25 (16), 4059–4074.

Rackham, O. (2008) Ancient woodlands: modern threats. *The New phytologist*. 180 (3), 571–586.

Rackham, O. (1986) *The History of the Countryside*. London: J.M. Dent.

Rackham, O. (2012) *Woodlands*. London: William Collins.

Ritter, C.D., Dunthorn, M., Anslan, S., Lima, V.X., Tedersoo, L., Nilsson, R.H. & Antonelli, A. (2020) Advancing biodiversity assessments with environmental DNA: Long-read technologies help reveal the drivers of Amazonian fungal diversity. *Ecology and evolution*. 186281.

Rousk, J., Bååth, E., Brookes, P.C., Lauber, C.L., Lozupone, C., Caporaso, J.G., Knight, R. & Fierer, N. (2010) Soil bacterial and fungal communities across a pH gradient in an arable soil. *The ISME journal*. 4 (10), 1340–1351.

Schneider, T., Keiblinger, K.M., Schmid, E., Sterflinger-Gleixner, K., Ellersdorfer, G., Roschitzki, B., Richter, A., Eberl, L., Zechmeister-Boltenstern, S. & Riedel, K. (2012) Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. *The ISME journal*. 6 (9), 1749–1762.

Seaton, F.M., George, P.B.L., Lebron, I., Jones, D.L., Creer, S. & Robinson, D.A. (2020) Soil textural heterogeneity impacts bacterial but not fungal diversity. *Soil biology & biochemistry*. 144107766.

Strickland, M.S., Lauber, C., Fierer, N. & Bradford, M.A. (2009) Testing the functional significance of microbial community composition. *Ecology*. 90 (2), 441–451.

Štursová, M., Šnajdr, J., Koukol, O., Tláskal, V., Cajthaml, T. & Baldrian, P. (2020) Long-term decomposition of litter in the montane forest and the definition of fungal traits in the successional space. *Fungal ecology*. 46100913.

Sundsdal, A., Graae, B.J., Speed, J.D.M., Bukombe, J., Mtweve, P.J., Arneberg, M.K., Haukenes, V.L., Grevskott, R.T. & Smith, S.W. (2020) Teatime in the Serengeti: macrodetritivores sustain recalcitrant plant litter decomposition across human-modified tropical savannahs. *Plant and soil*.

Sun, Y. & Zhao, S. (2016) Leaf litter decomposition in urban forests: test of the home-field advantage hypothesis. *Annals of forest science*. 73 (4), 1063–1072.

Susyan, E.A., Wirth, S., Ananyeva, N.D. & Stolnikova, E.V. (2011) Forest succession on abandoned arable soils in European Russia – Impacts on microbial biomass, fungal-bacterial ratio, and basal CO2 respiration activity. *European journal of soil biology*. 47 (3), 169–174.

The Scottish Government (2019) *Scotland's Forestry Strategy 2019–2029*.

Trogisch, S., He, J.-S., Hector, A. & Scherer-Lorenzen, M. (2016) Impact of species diversity, stand age and environmental factors on leaf litter decomposition in subtropical forests in China. *Plant and soil*. 400 (1), 337–350.

Trudgill, S.T. (1988) *Soil and Vegetation Systems*. Oxford: Oxford University Press.

Veen, G.F. (ciska), Freschet, G.T., Ordonez, A. & Wardle, D.A. (2015) Litter quality and environmental controls of home-field advantage effects on litter decomposition. *Oikos* . 124 (2), 187–195.

Veen, G.F. (ciska), Sundqvist, M.K. & Wardle, D.A. (2015) Environmental factors and traits that drive plant litter decomposition do not determine home-field advantage effects. *Functional ecology*. 29 (7), 981–991.

Wagg, C., Schlaeppi, K., Banerjee, S., Kuramae, E.E. & van der Heijden, M.G.A. (2019) Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. *Nature communications*. 10 (1), 1–10.

Welke, S.E. & Hope, G.D. (2005) Influences of stand composition and age on forest floor processes and chemistry in pure and mixed stands of Douglas-fir and paper birch in interior British Columbia. *Forest ecology and management*. 219 (1), 29–42.

Welsh Government (2018) *Woodlands for Wales*.

Xiao, W., Chen, H.Y.H., Kumar, P., Chen, C. & Guan, Q. (2019) Multiple interactions between tree composition and diversity and microbial diversity underly litter decomposition. *Geoderma*. 341161–171.

Yin, R., Eisenhauer, N., Auge, H., Purahong, W., Schmidt, A. & Schädler, M. (2019) Additive effects of experimental climate change and land use on faunal contribution to litter decomposition. *Soil biology & biochemistry*. 131141–148.

# Chapter 5. Changes to Soil Decomposer Communities Revealed by DNA Metabarcoding as Woodlands Age

## 5.1 Abstract

Woodland succession in the UK has been associated with changes in the composition and structure of vegetation aboveground but changes belowground are poorly characterised. The microbial communities in woodland soil and litter layers are predicted to vary over centuries of woodland ageing as they adapt to changes in the soils and vegetation, but also as bacteria and fungi colonise the site and compete for nutrients. Variation in the microbial diversity of decomposing leaves has been explored in relation to changes in leaf litter quality, seasonal change, and the Home Field Advantage hypothesis. As of yet, they have not been described across a chronosequence of woodlands of different ages that include newly created, mature, and ancient semi-natural woodland. It is unclear whether key decomposers of woodland tree leaf litter take years, decades, or centuries to colonise a woodland after afforestation. Differences between woodlands of different ages in the active diversity on decomposing leaf litter may relate to differences in ecosystem functioning between sites or may support calls for woodland creation or conservation. In this study I compare the bacterial and fungal diversity of decomposing single-species leaf litter bags (n = 809) from 5 common UK broadleaf tree species, at 27 woodland locations in North East England. These sites are organised in triplicates of similar woodland creation, mature woodland and ancient semi-natural woodland. I sample litter at 13 and 26 weeks following burial under the soil surface, in the spring and summer, respectively. I used next-generation amplicon sequencing to characterise fungal and bacterial diversity of leaf litter. Differences between litter producing tree species and sampling seasons were the dominant determinants of community composition for fungi and bacteria and for the alpha diversity of bacteria. The tree species of the litter but not seasonal influences significantly affected fungal alpha diversity. Generally, recalcitrant litter had higher alpha diversity and bacterial diversity increased in the summer sampling season. Next in explanatory power, woodland age category significantly influenced bacterial and fungal composition, best explained by variations in site ground cover associated with woodland age. Fungal community composition was also significantly associated with direct effects of woodland age category that were not explained by changes to the vegetation or soil properties of the site. Other vegetation gradients relating to ground cover and tree

composition and structure explained aspects of bacterial and fungal community composition, possibly supporting the substrate quality-matrix quality interaction hypothesis of decomposition. These also suggest that site management and planting choices influence the decomposer communities of woodlands. Overall, these results suggest that the active decomposer diversity of woodlands takes many decades to stabilize following afforestation, and that smaller but significant differences exist in the composition of mature and ancient woodland in the UK.

## 5.2 Introduction

### *5.2.1 Research needs*

The global and regional distributions of soil microbes are starting to be revealed with ambitious biogeographic next-generation sequencing (NGS) studies which identify variations in diversity by habitats and site characteristics (Fierer & Jackson, 2006; Griffiths *et al.*, 2011; George *et al.*, 2019). These studies have highlighted regional differences in UK bacterial communities between the north-west and south-east (Griffiths *et al.*, 2011), between crop-grassland and woodland-heath sites (George *et al.*, 2019), and correlations with numerous site and soil characteristics, particularly soil pH. Fungal biogeography, however, has been less extensively researched, but appears to be influenced by similar drivers within the UK (George *et al.*, 2019). The classical view that fungal taxa were present more-or-less ubiquitously but selected for by the environment has been challenged; early studies involving NGS have shown that dispersal barriers such as oceans and mountains also affect fungal distributions (Peay *et al.*, 2010). However, the extreme variability of fungal genetic barcoding regions, along with other peculiarities of the locus, present additional challenges to precisely identifying and describing fungal taxonomic variability using amplicon sequencing methods (Bengtsson-Palme *et al.*, 2013; Tedersoo *et al.*, 2016), which may slow progress in this area.

Broad scale descriptions of soil microbial diversity will be foundational for microbial ecology, but detailed comparisons across important ecological gradients will be essential for informed management and conservation. However, convincing land managers and policy makers of the value of microbial diversity may be required to effect applied outcomes from this work. Uncovering the relationships between microbial diversity and ecosystem functioning may be

one way to achieve this. Reviews have highlighted that relationships exist between soil carbon dynamics and both richness and microbial community composition (Nielsen *et al.*, 2011; Schimel & Schaeffer, 2012; Lladó *et al.*, 2017). These have highlighted that carbon cycling processes in general seem highly redundant but that specialised processes such as lignin and cellulose breakdown may be less robust to species loss. Recent experimental work has lent support to the assumption that increased richness may produce increasingly multifunctional soil microbial communities, demonstrating that increased richness leads to increasingly multifunctional bacterial and fungal communities when examined separately or together (Wagg *et al.*, 2019). These studies consistently recommend additional detailed, experimental work be conducted into the microbial support of soil processes, such as decomposition, that fuses bacterial and fungal genetic datasets (Uroz *et al.*, 2016; Bani *et al.*, 2018; Wagg *et al.*, 2019), explores the role of time and succession in soil microbial systems (Schimel & Schaeffer, 2012; van der Putten *et al.*, 2013), and attempts to separate active diversity from inactive or dormant microbes (Baldrian, 2017). In the previous chapters, I presented a study that incorporated the first two of these goals into its design by examining differences in bacterial and fungal diversity over a successional gradient of farmland to Ancient Semi-Natural Woodland (ASNW). I found that soil biodiversity of bacteria and fungi alters greatly following land-use change from arable to woodland systems and continues to differentiate away from arable systems in composition and relative abundances of bacteria and fungi for centuries. In this chapter I present the results of a study that incorporates all these features by comparing the bacterial and fungal diversity found on decomposing leaves over a woodland successional gradient from recently planted woodland to ASNW. The experiment moves beyond the question: does microbial diversity vary between woodlands at different successional stages? To ask the question: does the active diversity on decomposing leaves vary between woodlands of different ages? For the first time, I use NGS to explore the roles of seasonal change and the species of leaf litter on decomposer communities in a fully-replicated, orthogonal experiment that controls for or quantifies the confounding effects of vegetation and soil properties. By following the decomposition of litter placed out in winter over multiple seasons, I mimic natural decomposition of leaves as closely as possible. Importantly, the experiment explores the role of site characteristics on shaping microbial decomposer communities, including soil properties, tree community composition and dominance, and understory vegetation, which have all received little attention or have been ignored completely in studies to date.

### *5.2.2 Leaf litter from classical ecology to modernity*

Leaf litter studies have long been used to tie ecosystem properties or leaf characteristics to ecosystem functioning going back at least to the interwar period (Falconer *et al.*, 1933). They have been employed in a broad range of habitats including woodlands (Yue *et al.*, 2018), savannah (Sundsdal *et al.*, 2020), and aquatic systems (Pu *et al.*, 2014). The simplicity and versatility of leaf litter work has allowed research into numerous pure and applied ecological questions. Examinations of the Home Field Advantage effect (HFA) and the Substrate quality–Matrix quality Interaction (SMI) hypothesis (Freschet *et al.*, 2012) have allowed researchers to explore the mechanisms of community specialisation to ecosystem properties, a fundamental question in community assembly. Other studies have demonstrated changes to ecosystem functioning resulting from changes to grazing management (Chuan *et al.*, 2018) or plantation growth (Trogisch *et al.*, 2016). Lessons from nearly a century of leaf litter studies have improved our methodologies. For example, fine mesh litter bags have been shown to be capable of excluding larger invertebrates allowing the effects of microinvertebrates, bacteria, and fungi to be studied in isolation, at the cost of modifying moisture microclimates within bags (Bokhorst & Wardle, 2013). Mesh size in particular has been identified as an important factor in influencing litter experimental results. Additionally, non-additive effects on decomposition rate stemming from combining leaf litter of different species in a single bag have led to widespread investigation and discussion of the role of diversity in ecosystem functioning (Wardle *et al.*, 1997; Jonard *et al.*, 2008). Non-additive effects have not been observed in all cases (Jacob *et al.*, 2010), and meta-analysis has failed to confirm their importance in the majority of systems (Porre *et al.*, 2020). These studies and other reviews have highlighted the breadth of factors that can influence decomposition rate, including climate, leaf characteristics, site conditions, and soil chemical and physical properties (Aerts, 1997; Preston *et al.*, 2009; Porre *et al.*, 2020). Despite the recognised importance of a range of factors in modifying decomposition rates, disproportionate attention has been paid to different drivers and they are really all measured or controlled for in individual studies of decomposition (Porre *et al.*, 2020). In particular, recent studies that I am aware of have often not measured or modelled the effects of site, soil and vegetation properties, other than climate, on decomposition (but see (Bayranvand *et al.*, 2020) for a soil microbiome example of good site descriptions with high replication).

### *5.2.3 Methodological advances*

*Turn-of the century approaches in soil microbiology*

Recent advances in molecular science and computing have enabled research into the microbial communities found on decomposing leaves (Handelsman, 2004). The earliest of these culture-independent advances enabled researchers to characterise the litter microbiome based upon enzyme expression or molecular indicators such as PhosphoLipid Fatty Acids (PLFA). Reduced barriers to sequencing and analysis later enabled researchers to identify microbial taxa to a higher resolution based upon genetic sequences. Without metagenetic methods, modern techniques had already improved our understanding of the mechanisms of decomposition. For example, a combined PLFA, enzyme activity and metaproteomics study conducted by Schneider *et al.* (2012), confirmed fungi as the major producers of leaf degrading enzymes and that decomposer associated microbial communities varied by season and litter quality. Strickland *et al.* (2009) indicated that decomposition rate of a litter is fastest in soils inoculated with microbial communities taken from the same site, although site and dominant litter effects on inoculum communities were not separated. This complemented previous research demonstrating effects of dominant tree species on soil communities as assessed by PLFA profile (Hackl *et al.*, 2005) and metagenetic methods (Urbanová *et al.*, 2015), suggesting that tree communities exert influence on microbial communities that go on to impact decomposition. Work in this area has contributed to the general discussion around whether microbial communities adapt to decompose the tree litter present at their sites - the HFA hypothesis - or whether qualitatively recalcitrant or labile litter inputs specialise soil communities to decomposing litter of a similar quality - the SMI hypothesis (Freschet *et al.*, 2012).

*Recent advances in litter microbiome dynamics*

Previous work leads us to expect to find different microbial communities in a site's soil and the more dynamic leaf litter layer (Baldrian *et al.*, 2012). As such, the relationship between the active decomposer microbial communities found on litter and the broader pool of soil microbial diversity needs further attention. For example, it is clear that leaf properties influence decomposer communities (Purahong *et al.*, 2015) and that woodland soil properties

influence both bacterial and fungal soil communities (Griffiths *et al.*, 2011; Hamonts *et al.*, 2017; George *et al.*, 2019), but it is unclear how the effects of soil properties carry over to the leaf litter layer by influencing the pool of decomposers available at the site. Additionally, we have good reason to believe that leaf litter decomposer communities alter in their composition over the course of a leaf's decomposition (López-Mondéjar *et al.*, 2015; Purahong *et al.*, 2016; Buresova *et al.*, 2019), with the early stages of decomposition dominated by fungi and fungal leaf endophytes with bacteria and other fungal decomposers colonising over time (Baldrian, 2017; Buresova *et al.*, 2019). However, little attention has been paid to reconstructing the temporal dynamics of decomposer communities in a realistic multispecies litter community. We know little of the seasonal dynamics of resource switching in decomposer communities over the year as few studies include microbial community data from multiple commonly found tree species' litter of the study system in individual litter bags in their experiments. Yue *et al.* (2018) demonstrated that such experiments are feasible with an ambitious, replicated, multi-species experimental design, but the molecular community profiling method employed did not provide high enough taxonomic resolution to answer this question and it was not the focus of the study. Buresova *et al.* (2019) examines litter from beech, sedge and milkvetch at two sites with 70 litter bags per litter type, sampling five replicates at each of seven time periods over 13 months. However, they were only able to detect differences in the microbial communities of sedge and beech, which diverged in the summer months. Pereira *et al.* (2019) showed differences between actinobacterial abundance, fungal abundance, fungal to bacterial ratio and actinobacteria to fungi ratios between *Quercus pubescens, Quercus ilex, and Pinus halepensis* litters under various conditions across three woodland sites, but did not present data exploring bacterial or fungal beta-diversity in greater detail.

*Next-generation sequencing*

Only now, towards the second decade of the 21st century are decomposition researchers regularly making use of NGS and bioinformatic approaches that provide the volume of reads and the taxonomic resolution needed to answer these questions. Overall these approaches have generally validated expectations that fungal alpha-diversity, evenness and functional diversity positively influence decomposition rate (Xiao *et al.*, 2019), but the rate of

decomposition may be more strongly influenced by site properties than microbial community (Buresova *et al.*, 2019). Considerable advances have been made in our understanding of decomposition since the introduction of molecular methods, but that is not to say that the future of decomposition research relies on next-generation approaches or even molecular ones. Fundamental mechanisms of decomposition are still poorly understood and complex techniques are not required to improve our understanding of them, for example the crucial role of photodegradation in aboveground decomposition has been recently quantified with the application of biocidal solutions to soils (Berenstecher *et al.*, 2020).

### *5.2.4 Challenges and their resulting knowledge gaps*

Numerous obstacles must be overcome in conducting leaf litter experiments that are time consuming, labour intensive, and logistically challenging. These experiments require the collection, cleaning, processing, and bagging of leaf litter; burying or laying out litter bags and then recovering them; and calculating mass loss. At the same time sites must be accessed which may be remote, may require maintaining good relationships with landowners, may need to be visited at very specific times, and may require adjustments to be visited safely in all seasons. These factors inevitably lead researchers to make decisions about the scale of their experiments and which potential predictors to collect data on. While recognising these serious challenges, the literature makes clear the need for increasingly replicated, large scale studies of decomposition (Bradford *et al.*, 2016) that include information on a wider range of predictors (van der Wal *et al.*, 2013). Specifically, incorporating a wider range of tree species' litter in decomposition experiments has been suggested (Hättenschwiler *et al.*, 2005), and better characterising site vegetation beta diversity differences rather than alpha diversity difference when testing for diversity effects of site vegetation (Gessner *et al.*, 2010). Meta-analyses of existing work may be considered to meet these research gaps, but this approach may be hampered by the strong influence of small decisions in study methodology on the assessment of microbial diversity. Descriptions of the microbial composition of soil samples are heavily dependent on sample handling and storage (Delavaux *et al.*, 2020), laboratory approaches (Dopheide *et al.*, 2018), and bioinformatic pipelines can all influence the results of soil microbiome studies (Pauvert *et al.*, 2019). As such reviews or meta-analyses that combine results from multiple genetic studies risk misinterpreting differences in study

methodology as differences in habitat, treatment, or biology. Large scale studies with identical sample handling and analysis are required.

To my knowledge no study has compared the active microbial diversity found on decomposing leaves between woodlands of different ages. In previous chapters I demonstrated that microbial diversity of both bacteria and fungi significantly vary across a gradient of a forestation and succession. Although I did not find significant differences between woodlands of different ages, woodland communities continued to differentiate from agriculture ones over time, with ancient woodland communities being the most dissimilar to arable ones. In the literature, much of the work is limited to the first half-century of woodland growth following restoration on agricultural land. There are also conflicting results regarding the role of woodland age on microbial diversity; for example, significant differences in diversity and composition between very young woodlands less than around 15 years and more established restoration woodlands that have been established for several decades (Jangid *et al.*, 2011; Mackay *et al.*, 2016; Creamer *et al.*, 2016; Jiao *et al.*, 2018; Wu *et al.*, 2020). These differences appear to carry through to older mature secondary woodlands of around a century or two in age, and to ancient or remnant woodlands in some studies (Zhu *et al.*, 2010; Creamer *et al.*, 2016; Mackay *et al.*, 2016). However, other studies have found no significant differences in microbial diversity or composition between woodland of different ages (Ma *et al.*, 2019) and reported differences between mature and ancient woodland microbial diversity vary inconsistently in direction and significance between studies and taxonomic groupings (Zhu *et al.*, 2010; Jangid *et al.*, 2011). As of yet we lack a clear understanding of the general trends in microbial changes during broadleaf woodland succession. It is unclear if differences in the response of microbial communities to woodland succession are caused by soil chemistry, biogeographic differences, management, or land use history. To resolve some of these outstanding questions, I conducted a large-scale study to determine trends across multiple similar sites, as provided in the previous chapters.

### *5.2.5 Hypotheses*

I hypothesised that (H1) alpha diversity will vary across tree species' litter, season, and woodland age. Litter microbial alpha diversity of both bacteria is expected to be higher in fast

decomposing species like ash and elm and lowest in recalcitrant species such as beech and oak, while the reverse is expected to be true for fungal alpha diversity. I predict it will increase over the course of the year and will be higher in mature secondary woodlands that woodland creation or ASNW sites. I hypothesise that (H2) ecological communities will significantly differ by the tree species of the litter, sampling season, and woodland age. Litter of tree species that decompose at similar rates will be most similar, and woodland communities will become increasingly dissimilar as the age difference between woodland increases. H3 Tree community diversity and dominance will strongly influence tree leaf decomposer communities, more so than understory vegetation.

## 5.3 Methods

### *5.3.1 Site descriptions and decomposition experiment*

The previous chapter described a large-scale highly replicated decomposition experiment in woodlands situated across the North East of England (see Chapter 4 for full details). Briefly, in this experiment I surveyed 27 woodlands in the summer of 2017 and grouped them into triplicates with young woodland, mature woodland, and ASNW such that no sites within a triplicate differed by more than 0.5 pH or had dissimilar soil textures. At these sites I buried replicates of 1g of leaf litter from five tree species (Pedunculate Oak - *Quercus robur*, Common Ash - *Fraxinus excelsior*, Common Beech - *Fagus sylvatica*, Wyche Elm - *Ulmus glabra*, and Common Hael - *Corylus avellana*) in December 2017 producing a highly replicated, balanced decomposition experiment. I designed the experiment planning to sample litter bags every 13 weeks, the first sampling in spring (April 2018, the second in July in the summer (July 2018). I also buried and sampled an autumn (October 2018) replicates but did not have time to analyse them for this work. I buried three replicates of each of the five species, for each sampling season, at each site (3 replicates x five species x two seasons x 27 sites = 810 leaf litter bags). At each site, I collected soil cores to determine edaphic properties, including soil texture, total carbon, organic carbon, total nitrogen, and C/N ratio. I recorded the coordinates of the location to model spatial autocorrelation within the data. While sampling litter bag replicates for the summer season (July 2018) I surveyed the tree communities and understory communities of each sampling location while collecting samples. A multivariate modelling approach using canonical-correlation analysis (CCA) identified variation within these data that were best explained by the age category of the woodland, identifying one constrained axis

that accounted for ground vegetation and one for tree community composition and importance (dominance x frequency x density) that were associated with older or younger woodlands. These models also identified several unconstrained axes that captured naturally occurring environmental gradients within the data that correlated with certain trees or understory plants. These constrained and unconstrained gradients are more fully explored in the previous chapter, along with a detailed exploration and critique of the methods. Prior to determining the loss of dry weight, I removed approximately 0.1g of wet leaf material from each sample from which to extract microbial DNA. I sequenced DNA from these samples and calculated alpha and beta diversity metrics for analysis as described below.

### *5.3.2 Summary of molecular and bioinformatic workflows:*

Briefly, I extracted, purified and amplified DNA from each replicate (n=809) of each of every seasonal sample of each tree species' litter at each site (site x species x season = 270) sample at two genetic loci, the 16S rDNA region for bacteria, and the ITS1 rDNA loci for fungi. I organised and processed samples in libraries with up to 92 samples, two positives and two negatives. The first two sample libraries consisted of the ITS1 and 16S amplifications of samples mostly from the first sampling season and were sequenced separately to test the efficacy of the extraction method. The remaining samples were sequenced subsequently and contained samples from both sampling seasons. Each sample, negative and positive was identified with within-library molecular tags that were applied during this initial amplification. I normalised the concentrations of all samples within libraries and performed a second, short amplification to add library level tags. Finally, I pooled all libraries together for sequencing on either an *Illumina MiSeq* or *Illumina Hiseq* platform at Northumbria University or Durham University, respectively. I filtered, trimmed and assigned taxonomy to the demultiplexed samples using DADA2 (Callahan, McMurdie, *et al.*, 2016) in R (v3.6.0). I then accounted for differences in sequencing depth in the package DESeq2 (Love *et al.*, 2014), applying a Variance Stabilising Transformation (VST) and rlog transformation to the read abundance data. Finally, I generated alpha diversity statistics in the package *phyloseq* to test H1 and analysed beta-diversity with a CCA model in the package vegan (Oksanen *et al.*, 2019) to test H2 and H3.

### *5.3.3 Molecular Methods*

*Extraction and purification*

Prior to DNA extraction I placed 0.1 of leaf litter material into a 5ml Eppendorf Tube with 0.8g of crushed, acid-washed garnet. I randomly allocated sample tubes into libraries with between 89 and 92 samples. The first library consisted of a subset of samples to be sequenced on a MiSeq as a pilot of the molecular methods. As such this initial library contained mostly samples from the first sampling season. Samples within a library were extracted and purified simultaneously, along with a sample negative. I added a second negative to the library prior to the initial amplification PCR, along with a positive DNA sample of either extracted separately. I added a second positive sample prior to the second PCR that added library level tags to the data. In fungal amplifications I used *Flammulina velutipes* DNA the first positive (demonstrating contamination control prior to the first PCR) and *Lentinula edodes* for the second positive. For bacteria, I used DNA extracted from isolates of *Shewanella oneidensis* for the first DNA positive, and *Planococcus alkanoclasticus* for the second PCR positive.

I added 900µL of lysis solution A (an aqueous solution of 147 mM guanidine thiocyanate, 228 mM trisodium phosphate, 26 mM sodium chloride, 67 mM Tris HCl, and 27 mM EDTA at 9.0 pH) to the tube containing the 0.1g of leaf material and 0.8g of sterilised, crushed garnet. I shook this mixture at 1750 RPM in a Geno/grinder 2010 for four minutes, then added 350µL of lysis solution B (an aqueous solution of 90 mM aluminium ammonium sulphate and 1.25% (w/v) SDS). I centrifuged the sample at 4,000 x g for 1 minute, transferred 1ml of supernatant to a 2.2ml deep-well plate and then centrifuged it again at 4,000 x g for 10 minutes. I transferred 250 µl of supernatant to a fresh 1.2ml reaction tube plate. To this tube I added 100 µl volume of protein flocculant solution (aqueous 5M ammonium acetate solution) and incubated the sample over ice for at least 10 minutes before centrifuging at 4,000 x g for 10 minutes. I added 50µL of amplification inhibitor flocculant solution made of equal parts of freshly aqueous 180 mM aluminium ammonium sulphate and 204 mM calcium chloride dihydrate. I then centrifuged the solutions again at 4,000 x g for 10 minutes. I added 800µl of binding solution (5.5 M aqueous guanidine HCl) to each sample and mixed the sample by pipetting. In three successive steps I transferred 600µl of the sample solutions into a 96-well silica-membrane spin plate (Bio Basic, #SD5007), sealed the plate with a breathable

membrane, and centrifuged the samples at 4,000 x g for 5 minutes over a 2.2ml deep-well plate to collect the flow-though. After completing this three times I added a wash solution (80% ethanol) to each well to remove excess contaminates and inhibitors, applied a breathable seal to the plate, then centrifuged the samples at 4,000xg for 15 minutes. Following this I discarded the flow through and transferred the spin plate to a 0.6ml DNA collection plate. I added 200 µl of elution buffer (1mM Tris at pH 8) heated to 70°C directly to the silica filter and left it to incubate for two minutes. Finally, I centrifuged the samples at 4,000xg for 5 minutes, collecting the eluted, purified DNA in the collection plate. This plate was sealed with a polypropylene seal and stored at -20°C.

*Amplification*

Prior to amplification, I diluted all samples to 1:20 concentration with 1mMol tris which tended to improve PCR success. I diluted especially recalcitrant samples further to 1:200 in 1mMol tris which normally allowed for successful amplification. As with the chapter 3, I used nested tagged (Kitson *et al.*, 2018) primer pairs ITS1F-ITS2 and 515F-806R for ITS fungal and 16S bacterial loci respectively (White *et al.*, 1990; Gardes & Bruns, 1993; Caporaso *et al.*, 2011). For ITS1 amplification I used 35 cycles (95°C for 30s, 53°C for 30s, and 72°C for 60s) of 20µl reactions using the MyFi Mix polymerase enzyme (Bioline), 2 µl of DNA template and the primers (reaching 0.25 µM concentration in PCR). For 16S, I PCR amplified samples in 30 cycles (95°C for 15s, 49°C for 30s, and 72°C for 60s) but were otherwise identical to those of the ITS1 amplification. I sealed wells of the plates with a drop of mineral oil to reduce the risk that wells would contaminate one another. I sealed the plate with a polypropylene seal during amplification on a PrimeG gradient thermal cycler, 96 x 0.2ml (5PRIMEG/02) and confirmed amplification success with an electrophoresis gel. I re-amplified samples that failed extractions.

I used Solid-Phase Reversible Immobilisation (SPRI) carboxylated paramagnetic beads to clean amplifications of primer dimers, size select the targeted region, and normalise sample concentration. I combined dilute SPRI beads (0.1mg/ml) with 15µl of sample in a 0.8 ratio of beads to sample, which I selected as it enables the selective binding of fragments larger than c.200 bp in testing while allowing unused primers and primer dimers to be washed off. I

conducted these SPRI bead clean-up and size selections by hand for the initial library of each locus, which contained the same set of samples destined for sequencing on a single MiSeq run. For all later libraries, destined for HiSeq sequencing, I conducted SPRI bead clean-ups and size selections on an OT-2 liquid handling robot (*Opentron*) to achieve high sample throughput and precise timings. As with previous work (Chapter 3), I used a magnetic plate to pellet samples while I washed them off primer dimers and with two successive 80% ethanol washes. I then allowed the samples to air dry for 15 minutes, removed samples from the magnetic plate and eluted samples in 10mMol tris. I pooled 15µl of each sample separately for each library. I increased the concentration of the pre-libraries of pooled samples with an SPRI bead concentration step. Here I combined 200µl of each pooled pre-library with undiluted bead solution (1mg/ml) in a 1.8 ratio of beads to sample, following the same steps as above.

I amplified the concentrated, pooled pre-libraries with a second PCR amplification to attach Illumina adapters and library specific tags. The conditions were 12 cycles (98°C for 20 s, 72°C for 30 s) in 20µl reactions using MyFi Mix (Bioline), with 5 µl of template DNA and library primers (at 0.25 µM reaction concentration). As with previous work. I conducted these amplifications in triplicate and pooled and concentrated from the combined 60µl (108µl 1mg/ml full strength bead solution) to 30µl. I checked the amplification's success by comparing the library concentration via Qubit analysis and fragment size via gel electrophoresis before and after amplification. One library that did not amplify successfully and so I diluted this library to 5ng/µl, facilitating successful amplification. Following the second PCR, I conducted an additional SPRI bead size selection to remove unused primers and primer-dimers as described previously in this paragraph. For the initial MiSeq libraries, I then sent these libraries in separate tubes to be sequenced on a MiSeq v3 (2x 300 bp). For later libraries, I diluted all libraries to 5nMol (around the concentration of the most dilute library) and then pooled all libraries into a final sample for HiSeq 2500 (2x250b) sequencing. This final multi-library was concentrated using SPRI beads again to 11.35 nMol to satisfy the requirements of the sequencing facility regarding sample concentrations.

### *5.3.4 Bioinformatics*

I received data from the sequencing facilities demultiplexed to the library level. I demultiplexed these using *metaBEAT* (Hahn & Lunt, 2019) and processed the MiSeq and HiSeq runs separately until taxonomic assignment, prior to which I merged sequencing runs for each locus. I trimmed and filtered the demultiplexed sequence data using *DADA2* in *R* (v3.6.0), removing primers with *cutadapt* v1.18 (Martin, 2011). DADA2 filters samples based upon read quality, then merged paired-end reads, removed any chimeras, inferred Amplicon Sequence Variants (ASVs) as recommended by the package authors (Callahan *et al.*, 2017). The summary data for this process can be found in supplementary materials Table S5.1. I then merged sequencing runs for each locus in DADA2 and assigned taxonomy using the UNITE and SILVA databases (UNITE v 8.0, SILVA v132). I then removed ASVs that had fewer than 10 reads across all samples. Samples that had appeared to fail to amplify in gel images of PCR products had still been sequenced. At this stage, the read totals of the initial PCR and the PCR repeat, or repeats were compared and the one with the highest read total was carried forward for analysis, the others were discarded. Later multivariate analysis required a completely balanced study design, with equal sample numbers. As there is some unevenness in the number of replicates for each sample, I also produced a version of the data merging replicates of each tree species' litter at a site for each season (up to three replicates) before normalising samples for uneven read depth. This reduced the dataset by two thirds, from 809 sample replicates to 270 merged samples.

I normalised for differences in sampling depth with *DESeq2*. AS a result, one bacterial sample with low read depth (hazel litter in spring, replicate 2, Newfield woods mature secondary woodland<2700 reads) was discarded due to the requirement of VST transformations that at least one gene or ASV must appear in all samples to allow for normalisation. This sample was the only regular sample (excluding positives and negatives) that lacked ASV3, which had been assigned to an unknown species within the genus *Galbitalea* in the family *Microbacteriaceae* (however DADA2 assignments should be treated cautiously). I also removed 16S samples that had not been assigned to at least the phylum level, as recommended in existing protocols for loci with good phylogenetic coverage (Callahan, Sankaran, *et al.*, 2016). I used the remaining adjusted read counts to calculate ASV richness and Shannon diversity in *phyloseq* (McMurdie & Holmes, 2013). *Phyloseq* takes methods from the *vegan* package to calculate these metrics (Oksanen *et al.*, 2019). I then VST and rlog transformed the normalised abundance data,

merged by replicate, to reduce heteroscedasticity and minimise differences in sample ASVs with few reads, as recommended in analysis pipelines (Love *et al.*, 2014; Callahan, Sankaran, *et al.*, 2016). I used this replicate-merged, transformed data to analyse beta-diversity for H2 and H3.

### *5.3.5 Statistical Analysis*

*Alpha Diversity - testing Hypothesis 1*

For each locus, I modelled differences in Shannon diversity between the litter of different tree species, between seasons and between woodland age categories using GLMMs from the *nlme* package (Pinheiro *et al.*, 2019). I refined models down from a global model containing all first and second order interactions, to a minimal adequate model where the removal of any fixed term significantly decreased model fit tested by the anova.lme function of *nlme* which compares models producing a likelihood-ratio p-value. I also removed significant interaction terms between non-significant variables. As the age of the woodland was the central research question, I did not remove it from models so that a p-value could be generated. I used a samples site and the assigned triplicate as random effects. As a post-hoc test on significant model variables, I compared the 95% confidence intervals on the predicted effects for each factor level for overlap. Fungal models had somewhat skewed residuals, influenced by outliers with particularly low Shannon diversity. These outliers are otherwise representative of the rest of the data, and had normal sequencing depths, simply very low diversity. Exploration of transformations or modifications that resolve this skew did not importantly change the conclusions of the model. Guided by a preference for parsimony, I have chosen to report on this interpretable model below that includes these outliers (whilst acknowledging important caveats).

*Beta Diversity - Testing Hypotheses 2 and 3*

I used the VST and rlog transformed data to construct CCA models of differences in transformed abundance between samples using the package *vegan*. As described above, permutation testing required the data first be merged to remove difference in sample number based upon numbers of replicates, so I used a dataset merge by triplicate before *in silico* read

depth normalisation. I incorporated sample triplicate as a conditioning variable, and the global model included the tree species of the litter, season, woodland age category, 1st order interactions of these three variables, soil organic carbon, soil C:N ratio, the importance (relative frequency x relative density x average basal area) of the leaf litter's genus at the site, the first constrained axis of each of the understory and tree vegetation importance CCAs, the first unconstrained axis of the understory vegetation CCA and the first two unconstrained axes of the tree vegetation CCAs. Vegetation CCA axes were selected based upon explanatory power within the vegetation data and have been described in more detail in the previous chapter. I refined the model using the ordistep function with both forward and backwards selection at each step, between a minimal model with the conditioning variable of triplicate alone, to the global model as the maximum scope. I constrained the ordistep permutations, and future permutations, within the nested experimental structure of site within the site's triplicate. I determined significant axes using permutation testing, restricted as described above, and identified the significance of model terms in the same way. I produced biplots of significant model axes, using the 95% confidence intervals on the model estimate to identify significantly different effects of levels within categorical variables. I calculated the correlations between significant continuous variables and axes, to allow readers to discern the factors contributing to significant axes and the axes at which significant variables act.

Each axis represents a propensity for a sample to include different fungal or bacterial taxa that are strongly or weakly associated with either the negative or positive end of the axis. I assigned over 2,500 fungal and over 10,000 bacterial taxonomic sequence variants in *DADA2*, each of which had a score for each axis of their corresponding CCA model I could not present the scores for each taxonomic unit, nor would this be recommended as the accuracy of the taxonomic assignment of both fungi and bacteria to species or genus level are uncertain. I summarised the three fungal and bacterial classes with the highest median scores and the lowest median scores along each axis, excluding taxonomic classes with fewer than three ASVs which often have very high median scores due to their size. I also presented the median score of the three most common fungal phyla - Ascomycota, Basidiomycota, and Mortierellomycota, which together made up nearly 95% of ASVs.

## 5.4 Results

### *5.4.1 Bioinformatics*

*Contamination control: sample positives and negatives*

Positive and negative samples demonstrated generally effective contamination control and were analysed prior to any filtering of ASVs based on ASV prevalence, taxonomic misassignment, or the removal of eukaryotic ASVs. Negative samples, especially PCR negatives, lacking large quantities of purified DNA, may be more likely to amplify trace contaminants that would not be detected in real samples, or would be amplified at lower abundance. Read data for negatives during the DADA2 pipeline can be found in the supplementary materials Table S5.2. In 16S samples, in no instances did the PCR positive occur in negatives or regular samples, and the DNA positive occurred in one non-positive sample, at an abundance of 3 reads. There was evidence of small levels of contamination on this plate, in that the PCR negative contained 11 ASVs with 195 total reads, the highest number of ASVs found in any negative sample. However, the extraction negative on the same plate contained only a single ASV, and this ASV was not found in any other samples. The negative with the highest number of reads contained 265 reads across 3 ASVs. Other negatives contained low numbers of ASVs (<10 ASVs) and reads (<160 reads). The median contaminant read abundance in negative samples was 6 reads, and the mean 17. This indicates limited contamination with low numbers of ASVs at very low read abundances that are relatively small in comparison to the signal present in extracted samples.

For ITS samples, the PCR positive occurred in a single non-positive sample, at an abundance of 2 reads. The DNA positive indicated an issue with contamination within certain rows and columns of the plate containing repeated amplifications of samples that had weak or no bands in a gel image of their plates. Due to the gradual assembly of samples in this library as samples failed to be detected in gels, the contamination is spatially contained within the library to a single row and column on the PCR plate. All contaminated samples were removed. In these cases, sequence data may still be available from the “failed” samples in their original library. When this was the case, these initial samples were included in the analysis instead. The twelve of the twenty ITS negatives contained 0 reads. Negatives from the MiSeq library contained 19 and 15 ASVs in the extraction and PCR negatives respectively, each having 341 total reads. The

negative with the most reads and ASVs (the PCR negative of library 11) contained 23 ASVs with 8929 total reads. Thirteen of these ASVs were found in no other samples, indicating some form of general laboratory contamination with a small number of commonly found soil ASVs. The PCR negative of library 13 also contained 7 ASVs, with 2511 total reads, four of these ASVs, representing the majority of the reads, were found in no other samples also indicating a general laboratory contamination. The remaining four negatives with any reads contained one extraction negative with 3 ASVs and three negatives with a single ASV. The read abundances of contaminants were far higher in the ITS samples than the 16S samples, however, biological and phylogenetic differences between these loci, such as differences in total detected diversity, makes direct comparisons of read abundance spurious. Any contamination in samples is concerning. However, the contamination found in these samples is limited and generally of multiple orders of magnitude difference in the number of ASVs that it is not likely to influence diversity metrics. This may be especially so as 20 of the 52 contaminant ASVs are found in only that negative sample and are not present in the analysed sample data. It is likely that in analysed samples, which have the actual purified sample DNA included, contaminants will be rare, and will not greatly influence the outcomes of the analysis. Importantly, there is no reason to believe that, for either loci, meaningful contamination occurred between wells during PCR or setup that cannot be mitigated for.

### *5.4.2 Description of final dataset*

The following describe the read abundances and ASVs present in samples after removing from the dataset the contaminated sample repeats, the positives and the negatives; selecting the high read depth sample of repeated samples; removing ASVs with fewer than 10 reads; and removing 16S ASVs that have not been assigned a Phylum.

In the ITS data, I detected 2,425 ASVs across the 808 samples, with a total of 81,298,248 reads. Samples had a median read depth of 80,992 reads following this filtering and contained between 31 and 334 ASVs with a median of 153.5. All ASVs belonged to the kingdom of fungi.

In the 16S data, I detected 10,603 ASVs across 809 samples, with a total of 90,149,971 reads. Samples had a median read depth of 105,035 reads following filtering and contained between 42 and 1901 ASVs per sample, with a median of 705 ASVs in a sample. The majority of reads belonged to bacteria (10593 of the ASVs) and 10 Archaea ASVs were included. Eukaryotic reads were detected, but these were not included in the above figures and were removed from the dataset. As the majority of the 16S reads belonged to bacteria, and all the ITS1 reads belonged to fungi, I refer to these as the bacterial and fungal or 16S and ITS1 datasets hereafter.

After selecting between repeat samples, 88 samples were included from the Illumina MiSeq run and 721 from the Illumina Hiseq run for bacterial data, and the MiSeq samples were sequenced at approximately half the sequencing depth. Only 84 of the MiSeq samples were included in from the fungal data, and 724 from the HiSeq run, and the MiSeq samples were sequenced at approximately a third of the sequencing depth. In ordinations, MiSeq and HiSeq datasets do not differentiate from one another.

### *5.4.3 Shared ASVs*

Most 16S and ITS1 ASVs were present on at least a single sample in all categories of woodland, on all litter types and in both seasons (Figure 5.1). But these results are highly sensitive to changes in the inclusion threshold, e.g. required occurrence in 10% of samples in a category. Many bacterial and fungal ASVs are strictly confined to a single species, season, or litter type.

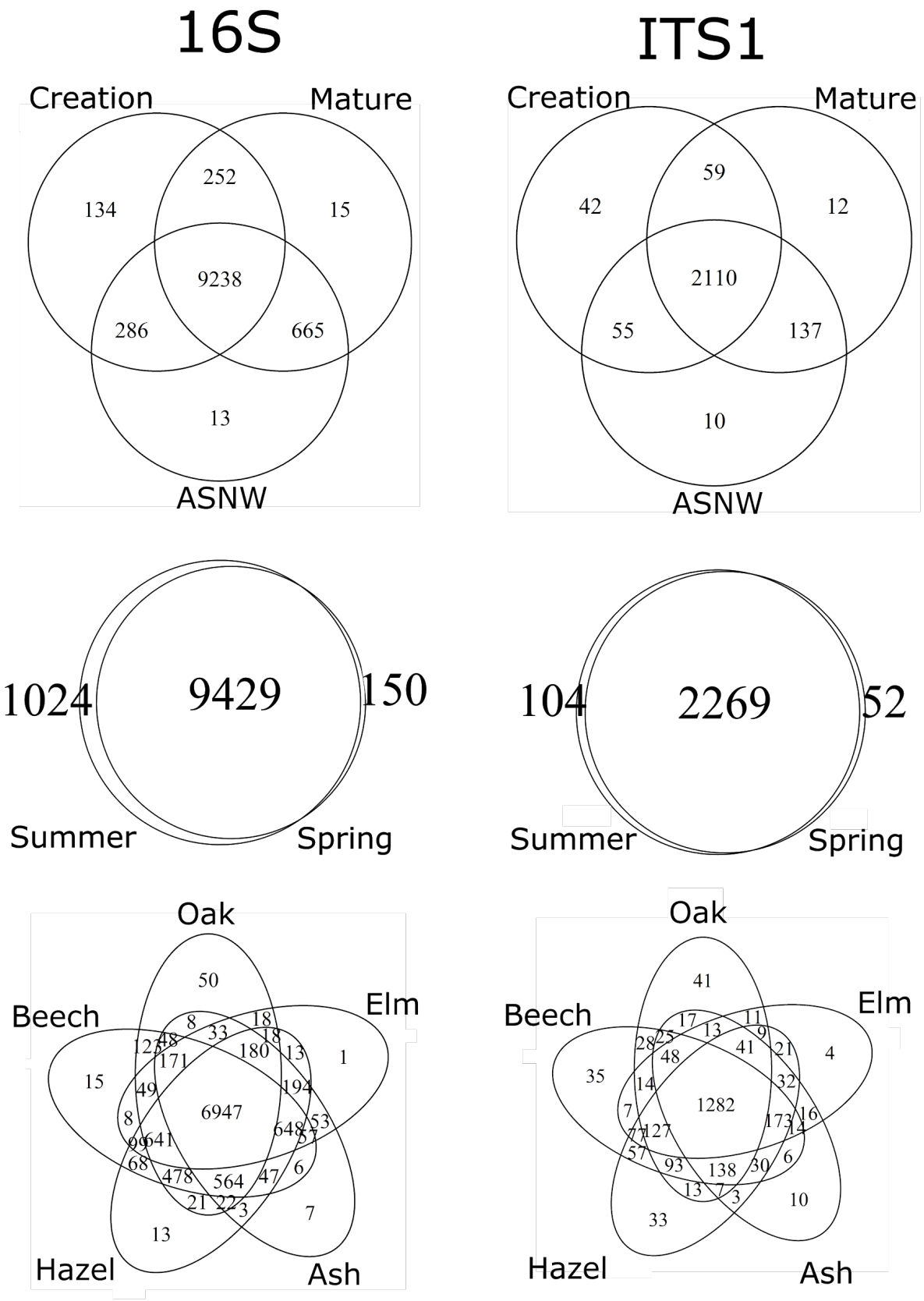


Figure 5.1 Venn diagrams displaying the number of Amplicon Sequence Variants (ASVs) present or shared across either age categories, seasons, or the litter's species. Only a single litter bag (<1%) need contain a microbial ASV for it to fit into a section of the plot.

### *5.4.4 H1 diversity differences in bacteria and fungi by litter type, season, and woodland age*

The final model explaining differences in alpha diversity for both loci included woodland age, tree species of the litter, sampling season, and an interaction between sampling season and tree species. Neither detected significant differences between the Shannon diversity of decomposing litter in woodlands of different ages (16S: $F_{(2, 16)} = 0.165$, $p = 0.85$, ITS1: ($F_{(2, 16)} = 0.819$, $p = 0.46$). Significant differences in Shannon diversity were observed between litter of different tree species for both loci (16S: $F_{(4,722)} = 95.18$, $p < 0.0001$, ITS1: ($F_{(4, 722)} = 68.58$, $p < 0.0001$). Seasonal differences were detected in bacterial Shannon diversity ($F_{(1, 722)} = 1013$, $p <.0001$), but not fungal Shannon diversity ($F_{(1, 722)} = 0.001$ $p = 0.97$), but both detected significant interactions between species and seasons (16S: $F_{(4,722)} = 30.43$, $p < 0.0001$, ITS1: ($F_{(4, 722)} = 3.247$, $p = 0.012$).

In bacterial 16S samples alpha diversity was significantly higher in summer for all species' litter, and higher in beech and hazel during spring, and higher in oak than ash in the spring (Figure 5.2A). In the summer beech was only significantly higher than oak and no other species were significantly different as assessed by 95% confidence intervals. For fungi, Shannon diversity was higher for beech and hazel than ash, oak or elm in all seasons (Figure 5.2B). Bacterial diversity responses to seasonal change were all positive, with samples increasing in diversity in the summer, but varied in magnitude by the tree species of the litter, with ash litter diversity increasing significantly more than all other tree species' litter except for elm litter. Elm increased significantly more than hazel litter bacterial diversity (Figure 5.2A).

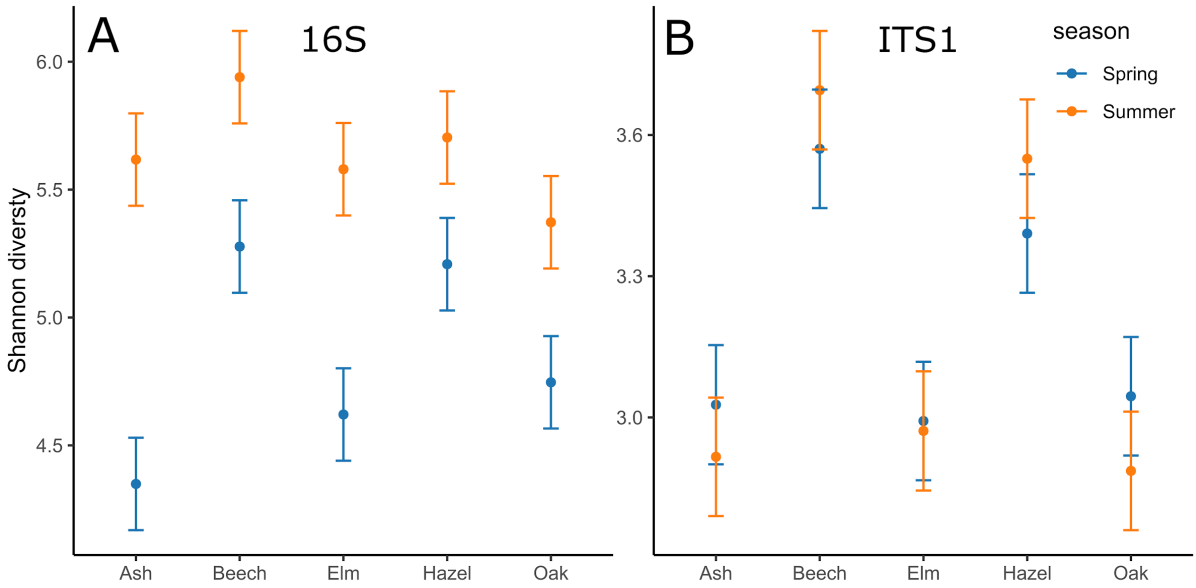


Figure 5.2 Seasonal differences in Shannon diversity of decomposing leaf litter buried in woodlands in the northeast of England. Diversity of ASVs for bacteria from the 16S locus (A) and for fungi at the ITS1 locus (B), with 95% confidence intervals around the predicted effects.

Regarding fungal diversity, oak litter had a significantly different, negative response in summer compared to their spring diversity, than beech and hazel litter's positive response to seasonal change. Additionally, ash litter's negative diversity change from spring to summer was significantly different to the positive response of hazel litter fungal diversity (Figure 5.2B). However, in no tree species did the summer fungal diversity significantly differ from its spring diversity.

Table 5.1A CCA Model variable significance on bacterial 16S community data extracted from decomposing leaf litter in spring and summer of 2018.

| Bacterial community composition | | | |
|---|---|---|---|
| Variable | DF | F | P |
| Season | 1, 249 | 14.0 | **<0.005** |
| Species | 4, 249 | 5.56 | **<0.005** |
| Species:Season | 4, 249 | 2.85 | **<0.005** |
| VegCCA1 | 1, 249 | 5.80 | **<0.005** |
| VegCA1 | 1, 249 | 3.14 | **<0.005** |
| TreeCA2 | 1, 249 | 4.10 | **<0.005** |

Table 5.1B. CCA Model variable significance" on fungal ITS1 community data extracted from decomposing leaf litter in spring and summer of 2018.

| Fungal community composition | | | |
|---|---|---|---|
| Variable | DF | F | P |
| Species | 4, 245 | 9.14 | **<0.005** |
| Season | 1, 245 | 8.82 | **<0.005** |
| Species:Season | 4, 245 | 2.84 | **<0.005** |
| VegCCA1 | 1, 245 | 5.37 | **<0.005** |
| TreeCA2 | 1, 245 | 3.59 | **<0.005** |
| VegCA1 | 1, 245 | 2.63 | **<0.005** |
| Woodland age | 2, 245 | 2.45 | **<0.005** |
| Season:Woodland age | 2, 245 | 1.45 | **<0.005** |

### *5.4.5 H2 and H3 differences in beta-diversity by sample and site factors*

### *5.4.6 Coefficient and axis significance*

The final model explaining differences in community composition between samples sequenced for Bacterial models included the tree species of the litter, season, an interaction between the two variables, the first constrained axis of the ground vegetation CCA (describing differences between woodland creation ground vegetation and the ground cover found in mature woodland and ASNW), the first unconstrained axis of the ground vegetation CCA, and the second unconstrained axis of the tree vegetation importance CCA. All of these variables significantly affected community composition (Table 5.1A), producing 12 significant bacterial community CCA constrained axes (Table 5.2). The model also contained a conditioning variable of the triplicate the sampling site belonged to. Final fungal models included the same variables and structure with the addition of the woodland's age category. All variables were significant (Table 5.1B). This CCA produced 16 axes, ten of which were significant, axes 1 to 9, and axis 14 (Table 5.3).

### *5.4.7 Detailed axis description*

Before describing the results of the CCA, I acknowledge that these outputs are complex and challenging to interpret, so provide more context. Each axis is a significant unimodal gradient, along which individual litter bag communities are described. Those at one end of the axis are more likely to contain ASVs of a specific set of fungal or bacterial taxa or contain more reads of those ASVs. As there are five tree species' litter examined, if each differed significantly across some aspect of community composition, which describes thousands of ASVs, one could make up ten contrasting gradients (ash-elm, ash-beech, ash-hazel, ash-oak, elm-beech, elm-hazel, etc.), with interactions with seasonal effects this produces a large number of possible gradients, which may also be collinear with effects of woodland age, or community composition. In other words, species that prefer the leaf litter of a certain tree species may also be more common in woods with certain vegetation types or may become more populous in the leaves of other tree species as decomposition successional stages develop.

Table 5.2 Axis significance for bacteria, with centroid scores for levels of categorical variables and axis correlations for continuous variables. Cells are coloured by the Direction of the score or correlation, green for positive, red for negative with the alpha controlled by the magnitude of the effect.

| Bacteria | | 16S | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Axis | df | f | p | Ash | Beech | Elm | Hazel | Oak | Spring | Summer | vegCCA1 | vegCA1 | treeCA2 |
| CCA1 | 1, 249 | 15.4 | 0.005 | -0.14 | 0.07 | -0.21 | -0.05 | 0.25 | -0.59 | 0.31 | 0.07 | 0.03 | -0.1 |
| CCA2 | 1, 249 | 12.2 | 0.005 | 0.49 | -0.17 | 0.46 | 0.05 | -0.58 | -0.24 | 0.13 | -0.07 | -0.06 | 0.09 |
| CCA3 | 1, 249 | 6.87 | 0.005 | 0.05 | -0.14 | 0 | -0.16 | 0.27 | -0.03 | 0.02 | -0.52 | -0.17 | 0.4 |
| CCA4 | 1, 249 | 5.49 | 0.005 | -0.26 | 0.27 | -0.14 | 0.27 | -0.23 | 0 | 0 | -0.29 | -0.05 | 0.1 |
| CCA5 | 1, 249 | 4.51 | 0.005 | 0.01 | 0.06 | -0.01 | -0.01 | -0.05 | -0.09 | 0.05 | -0.13 | 0.02 | 0 |
| CCA6 | 1, 249 | 3.19 | 0.005 | 0.02 | 0.25 | -0.09 | -0.14 | -0.06 | -0.01 | 0.01 | 0.29 | -0.15 | 0.37 |
| CCA7 | 1, 249 | 3.02 | 0.005 | 0.06 | 0.17 | 0.03 | -0.19 | -0.06 | 0.04 | -0.02 | -0.34 | 0.14 | -0.32 |
| CCA8 | 1, 249 | 2.49 | 0.01 | 0.03 | -0.05 | -0.16 | 0.19 | -0.02 | 0 | 0 | -0.06 | -0.15 | -0.08 |
| CCA9 | 1, 249 | 2.40 | 0.005 | -0.06 | 0.04 | 0.05 | -0.08 | 0.03 | -0.01 | 0.01 | -0.03 | -0.32 | -0.23 |
| CCA10 | 1, 249 | 2.16 | 0.005 | -0.28 | -0.01 | 0.25 | 0.01 | 0 | 0 | 0 | 0.01 | 0 | 0.04 |
| CCA11 | 1, 249 | 1.51 | 0.005 | 0.01 | 0.07 | 0.04 | -0.1 | -0.02 | -0.01 | 0.01 | 0.01 | -0.04 | 0.04 |
| CCA12 | 1, 249 | 1.41 | 0.005 | -0.12 | -0.01 | 0.1 | 0.02 | 0 | 0 | 0 | 0.02 | 0 | 0.03 |

Table 5.3 Axis significance for fungi, with centroid scores for levels of categorical variables and axis correlations for continuous variables. Cells are coloured by the Direction of the score or correlation, green for positive, red for negative with the alpha controlled by the magnitude of the effect.

| Fungi | | ITS1 | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Axis | df | f | p | Ash | Beech | Elm | Hazel | Oak | Spring | Summer | vegCCA1 | vegCA1 | treeCA2 | Creation | Mature | ASNW |
| CCA1 | 1, 245 | 14.6 | 0.005 | -0.18 | -0.05 | -0.41 | -0.40 | 0.96 | -0.11 | 0.08 | 0.00 | -0.01 | -0.01 | 0.00 | 0.02 | -0.01 |
| CCA2 | 1, 245 | 10.3 | 0.005 | -0.71 | 0.38 | -0.42 | 0.44 | 0.00 | 0.05 | -0.03 | -0.03 | 0.00 | -0.02 | -0.03 | 0.01 | 0.02 |
| CCA3 | 1, 245 | 9.09 | 0.005 | 0.08 | 0.11 | -0.09 | 0.07 | -0.19 | -0.46 | 0.33 | 0.18 | 0.04 | -0.14 | 0.09 | -0.02 | -0.08 |
| CCA4 | 1, 245 | 8.01 | 0.005 | 0.04 | 0.64 | 0.00 | -0.47 | -0.17 | 0.05 | -0.03 | -0.01 | 0.00 | 0.01 | -0.01 | 0.01 | 0.00 |
| CCA5 | 1, 245 | 5.92 | 0.005 | -0.04 | -0.01 | 0.02 | -0.01 | 0.03 | 0.09 | -0.07 | 0.75 | 0.11 | -0.30 | 0.38 | -0.19 | -0.21 |
| CCA6 | 1, 245 | 4.63 | 0.005 | 0.43 | 0.03 | -0.55 | 0.10 | -0.05 | 0.06 | -0.04 | 0.02 | -0.01 | -0.01 | 0.00 | 0.01 | -0.02 |
| CCA7 | 1, 245 | 4.16 | 0.005 | -0.04 | 0.00 | 0.03 | 0.02 | -0.01 | -0.02 | 0.01 | 0.08 | -0.26 | 0.45 | 0.12 | -0.02 | -0.11 |
| CCA8 | 1, 245 | 3.33 | 0.005 | 0.01 | 0.04 | -0.10 | 0.00 | 0.04 | -0.03 | 0.02 | 0.05 | -0.09 | 0.16 | 0.05 | 0.05 | -0.10 |
| CCA9 | 1, 245 | 2.97 | 0.02 | 0.06 | -0.01 | 0.02 | -0.06 | 0.03 | 0.00 | 0.00 | 0.01 | -0.05 | -0.14 | -0.02 | 0.04 | -0.02 |
| CCA14 | 1, 245 | 1.78 | 0.005 | 0.06 | -0.01 | -0.04 | 0.00 | 0.00 | -0.01 | 0.00 | -0.07 | -0.03 | 0.00 | -0.02 | 0.00 | 0.02 |

The constrained axes of models often had multiple contributing variables, which I have summarised in Table 5.2 and Table 5.3, displaying the centroid scores for levels of each categorical variable and the correlations of continuous variables with each axis. I have also represented these axes in Figures 5.3 for fungi, which describe species, season, and woodland age differences, and Figure 5.6, which describe species and season differences for bacteria. As necessary I have provided biplots of axes that are correlated with the vegetation CCA axes for ground vegetation that is associated with woodland age category- VegCCA1, ground vegetation unassociate with woodland age category - VegCA1, and tree community composition and importance unassociated with woodland age category - TreeCA2.

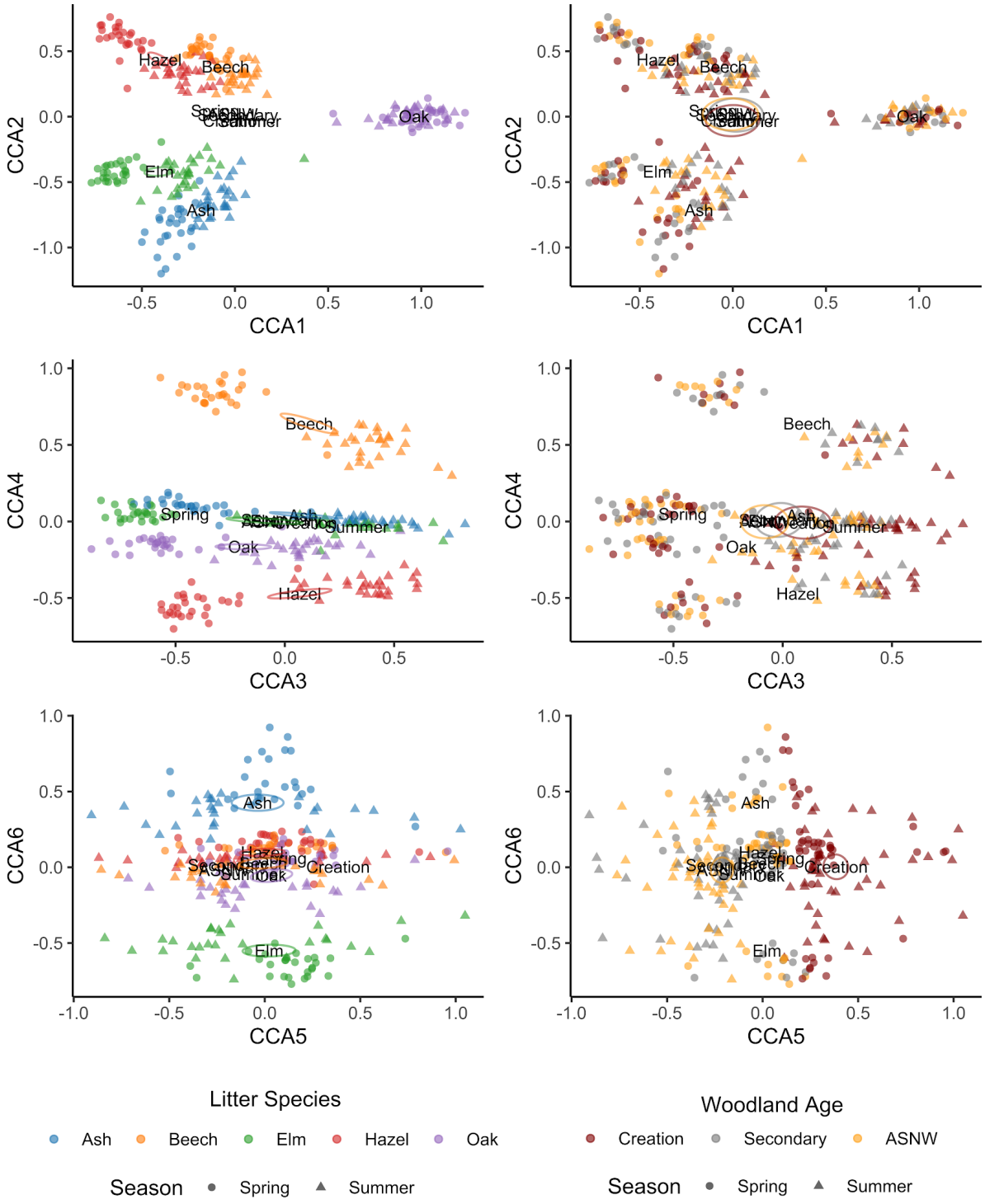


Figure 5.3 Fungal ITS1 CCA axes of individual litter samples. Point shape indicates sampling season, identical plots displayed on the left and right, but coloured by tree species on the left and woodland age on the right. All significant axes are displayed. Ellipses display 95% CI on the centroid.

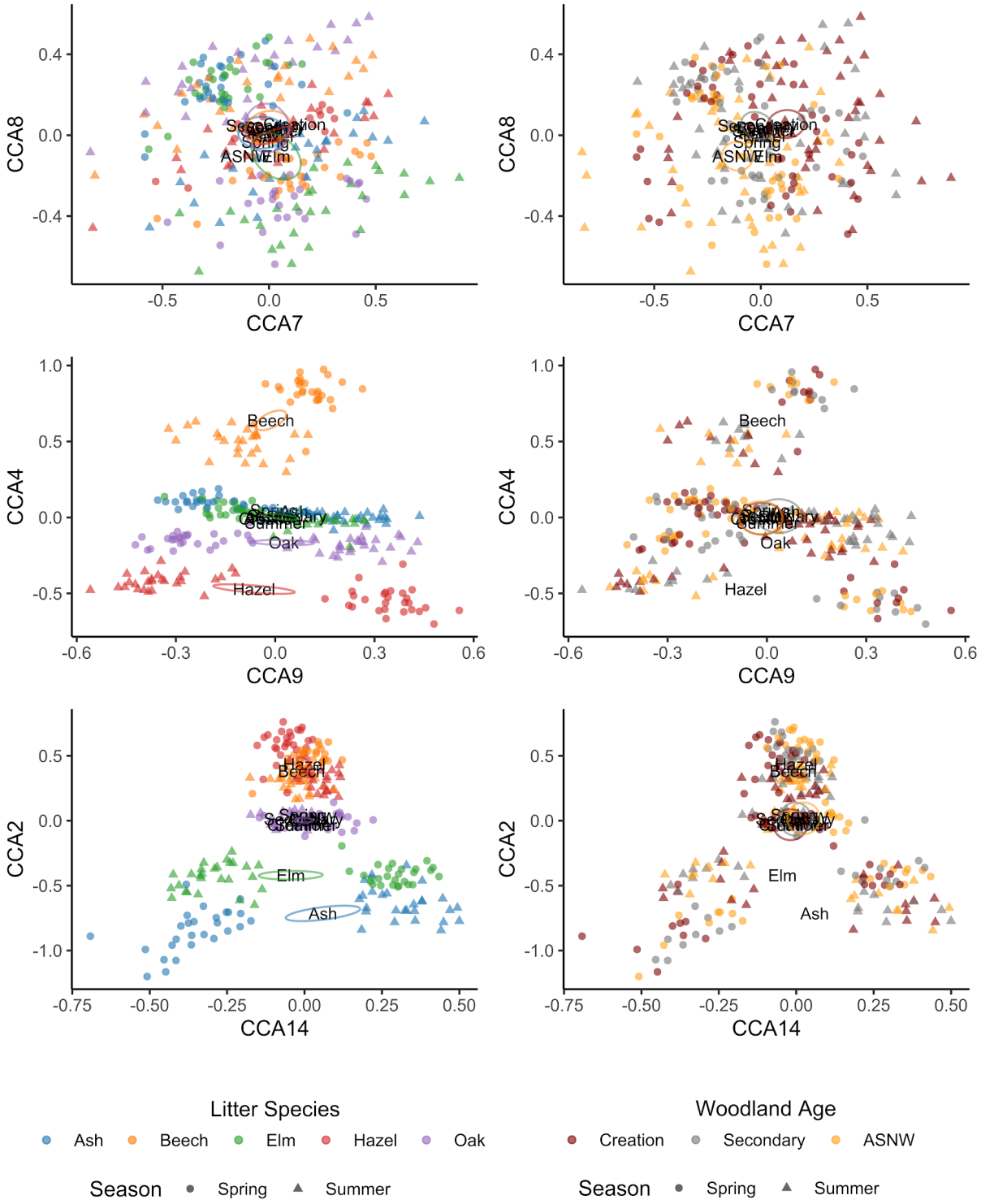


Figure 5.3 continued. Fungal ITS1 CCA axes of individual litter samples. Point shape indicates sampling season, identical plots displayed on the left and right, but coloured by tree species on the left and woodland age on the right. All significant axes are displayed. Ellipses display 95% CI on the centroid.

### *5.4.8 Fungal CCA*

The fungal CCAs identified species effects in the data controlling large proportions of data inertia, the primary axis separating the communities of oak litter samples from those of other species, particularly elm and hazel. The second axis separates aspects of community composition of fungi more commonly found on decomposing beech or hazel leaves than ash or elm leaves but neither prefer nor avoid oak.

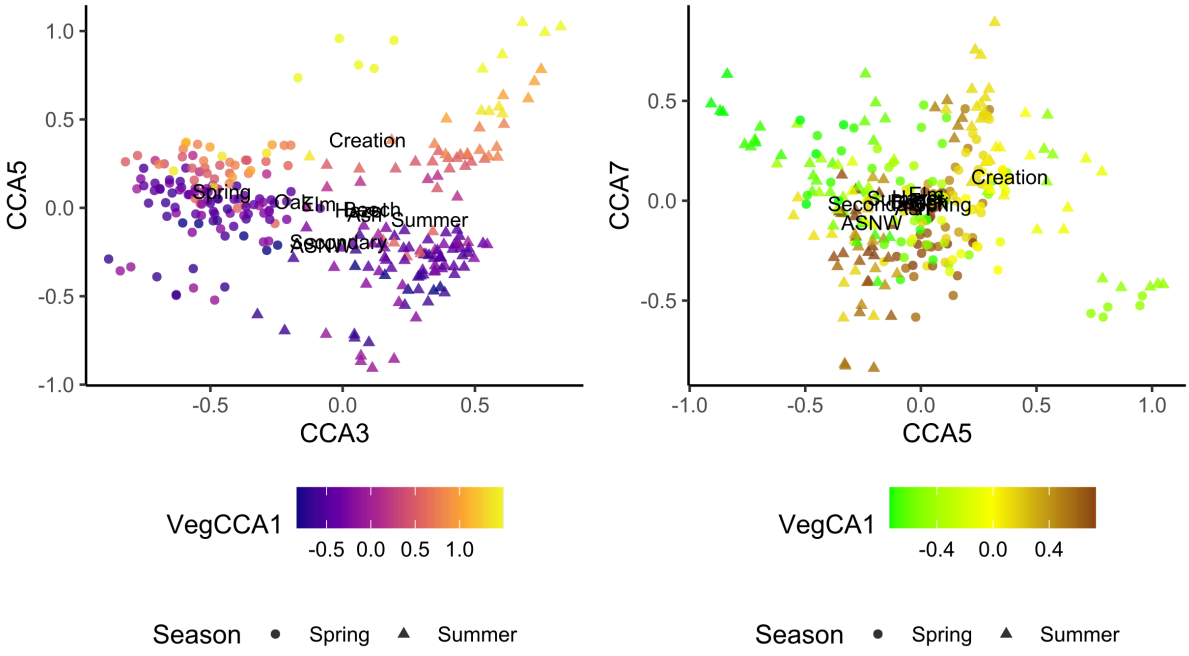


Figure 5.4 The fungal community composition axes most strongly correlated with ground vegetation axes VegCCA1 (A), constrained by woodland age, and the unconstrained axis VegCA1 (B). Points indicate the position on each axis of each sample, point shape corresponds to sampling season. Colour indicates vegetation type, on the left blue corresponds to ground cover associated with older woodlands, yellow with woodland creation. On the right, green indicates general herbaceous ground cover, brown indicates litter, or fern covered woodland.

The third axis describes significant differences in community composition caused by seasonal change from spring to summer but is also somewhat associated with aspects of ground vegetation relating to woodland age category (Figure5.4A) and tree community composition and importance (Figure 5.5) though these effects appear to be more pronounced in the summer season (Figure 5.4A) and in older woodlands (Figure 5.5). The fourth axis describes species differences, exemplified by contrasts in community composition between beech litter and that of oak or hazel leaves. The fifth axis describes differences relating to woodland age,

with young woodlands significantly different from mature and ancient woodlands, this axis is well explained by the first constrained ground vegetation CCA axis VegCCA1 (r = 0.75, Figure 5.4A), but also contributed influenced by variation in tree community composition and importance along the second unconstrained axis of the tree community and importance CCA TreeCA2 (r = -0.3, Figure 5.5) which is independent of woodland age category. Axis six described differentiation between ash and elm communities. Axis seven described other vegetation effects on decomposer communities associated with the first unconstrained ground vegetation axis VegCA1 (r = -0.26, Figure 5.4B) and TreeCA2 (r = 0.45, Figure 5.5), as well as slight but significant differentiation between ASNW and woodland creation centroids. Axis 8 also shows slight but significant shifts in composition between litter from ASNW and litter from woodland creation or mature woodland, as well as slight associations with VegCA1 (r = -0.09) and TreeCA2 (r = 0.16, Figure 5.5). VegCCA1 mostly captures between woodland creation and older woodlands, as such it does not strongly associate with the effects of age described with the previous two axes. The axes seven onwards all also capture additional interactions between species and season, or seasons and age category (Figure 5.3).

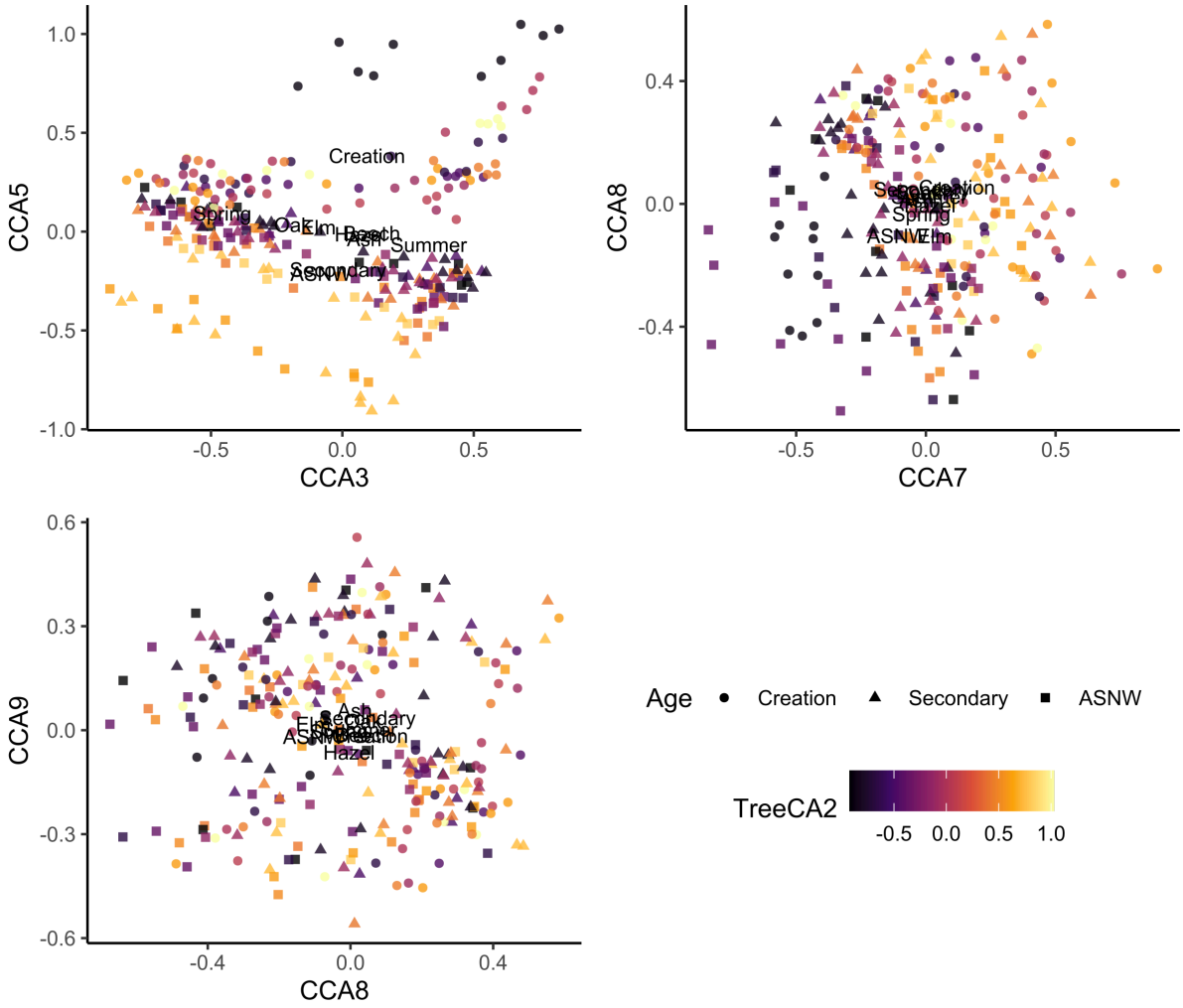


Figure 5.5. The fungal community composition axes most strongly correlated with tree composition and importance, axis TreeCA2, unconstrained by woodland age. Points indicate the position on each axis of each sample, point shape corresponds to sampling season. Colour corresponds to tree communities differences, low values in blue, among which are trees with recalcitrant leaf litter, such as oak, beech, and holly, high values in yellow correspond to woods with easily decomposable leaves, such as ash, and elm among other trends.

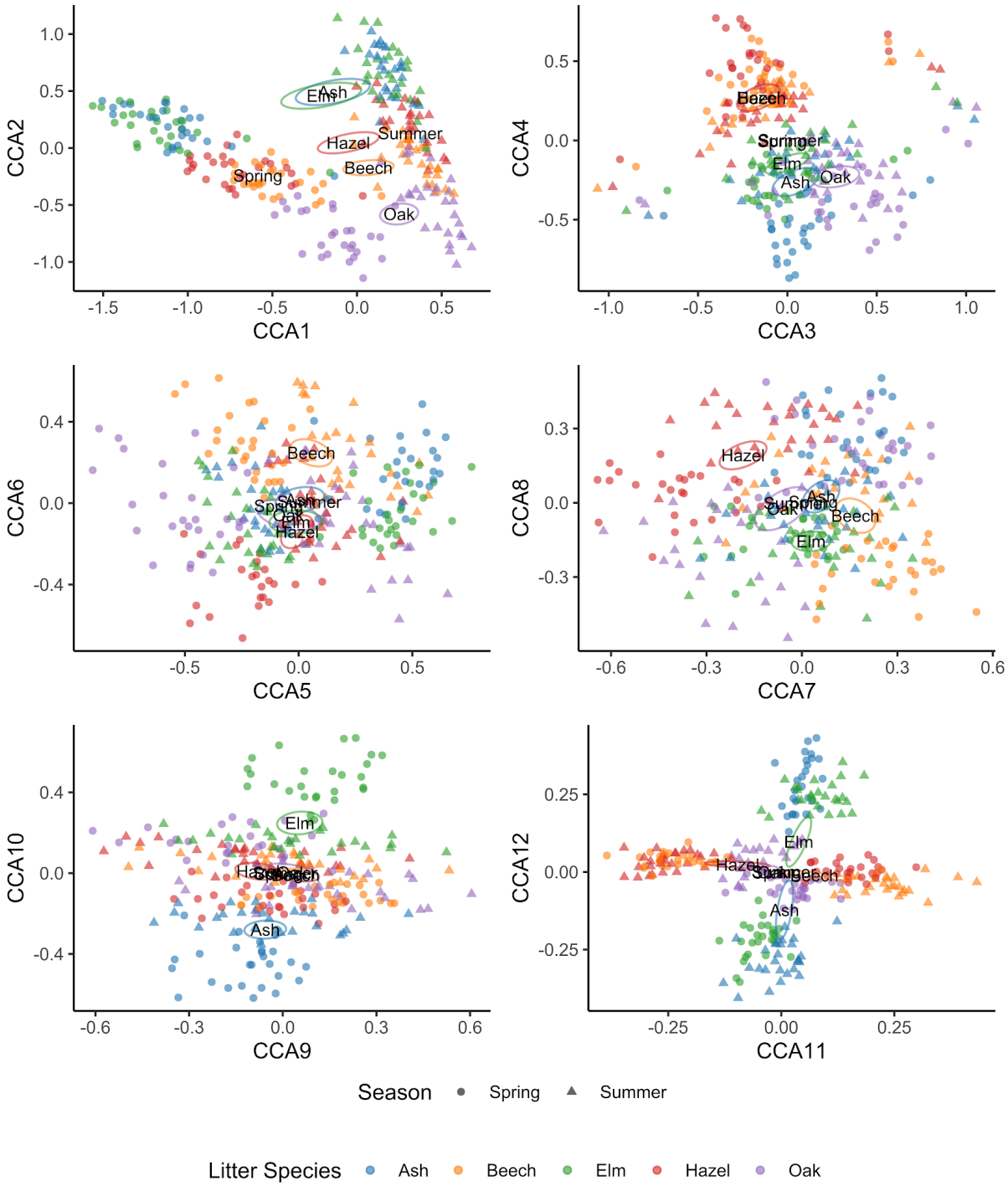


Figure 5.6 Bacterial 16S CCA axes of individual litter samples. Point shape indicates sampling season, colour indicates tree species. All significant axes are displayed. Ellipses display 95% CI on the centroid.

### *5.4.8 Bacterial CCA*

In the bacterial CCA, the strongest effect, captured by the first axis was a seasonal effect, strongly differentiating bacterial community composition in spring and summer (Figure 5.6), with a species effect and interaction in which litter's origin species position along axis one varied but the difference between spring and summer differed for species, in accordance to their positioning along CCA axis two, which differentiated communities by leaf litter's origin species from Elm and ash at high values, through hazel then beech both nearly neutral in regard to axis two, to oak communities at the negative end of the axis. Axis three was strongly associated with VegCCA1 (Figure 5.7). Woodland age category did not appear in this model, but high values of VegCCA1 (associated with ground vegetation typical of a young woodland e.g. grass) were associated with low values of axis three, and vice versa (r = -0.52). This axis was also positively associated and slightly with VegCA1 (r = -0.17, Figure 5.8), and more strongly with TreeCA2 (r = 0.4, Figure 5.9). Along this axis oak communities also significantly differentiated from those of beech and hazel litter. Axis four described additional differences between beech and hazel litter communities and those of oak ash and elm and negative effects (r = -0.29 Figure 5.7) of VegCCA1 - ground vegetation associated with woodland age. Axis five describes a slight influence of VegCCA1 on litter community composition (r = -0.13, Figure 5.7) and species-season interactions. Axis six differentiates species found on hazel litter from ash with an interaction with season that move the communities of the species in opposite directions in the summer, becoming more similar. It also distinguishes beech communities from those of other tree species' litter, and other species have other seasonal interactions. Axis seven distinguished the centroid of beech, at the top of this axis, from that of elm, oak, and hazel. The centroid of hazel communities, at the bottom of the axis, also significantly differs from those of ash and elm communities along this axis. All vegetation variables influence positions along this axis with indirect age effects via VegCCA1 (r = -0.34, Figure 5.7), other ground vegetation effects of VegCA1 (r = 0.14, Figure 5.8), and tree community composition and importance effects of TreeCA2 (r = -0.32, Figure 5.9).

Along axis 8 the centroids of elm, ash and hazel all significantly differ, as well as that of oak and beech from hazel, with different species displaying different interactions with seasonal effect. Axis 9 described community variations associated with vegetation unrelated to woodland age for VegCA1 (r = - 0.32, Figure 5.8) and TreeCA2 (r = -0.23, Figure 5.9).

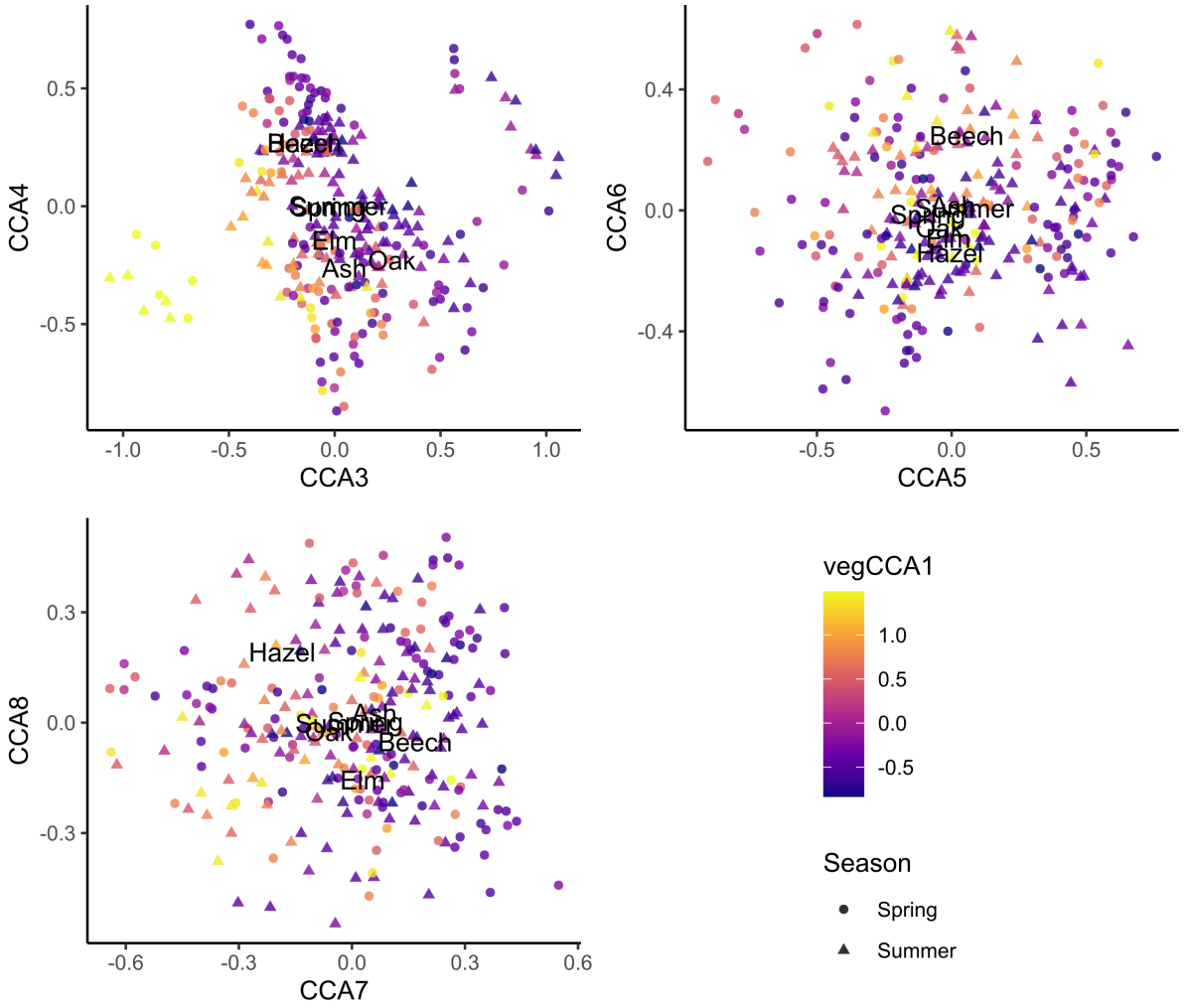


Figure 5.7 CCA axes of the bacterial community composition most strongly correlated with ground vegetation axes VegCCA1 - constrained by woodland age (i.e. axes 3, 4, 5, 6, 7, and 8). Points indicate the position on each axis of each sample, point shape corresponds to sampling season. Colour indicates vegetation type, blue corresponds to ground cover associated with older woodlands, yellow with woodland creation.

Axis 10 describes differences between elm and ash litter communities, with a seasonal interaction in which community composition becomes more similar in the summer. Axis 11 and 12 both describe litter's tree species and season interactions, with species associated with hazel and beech in spring switching association to the other in the summer along axis 11, and the same kind of interaction for elm and ash along axis 12.

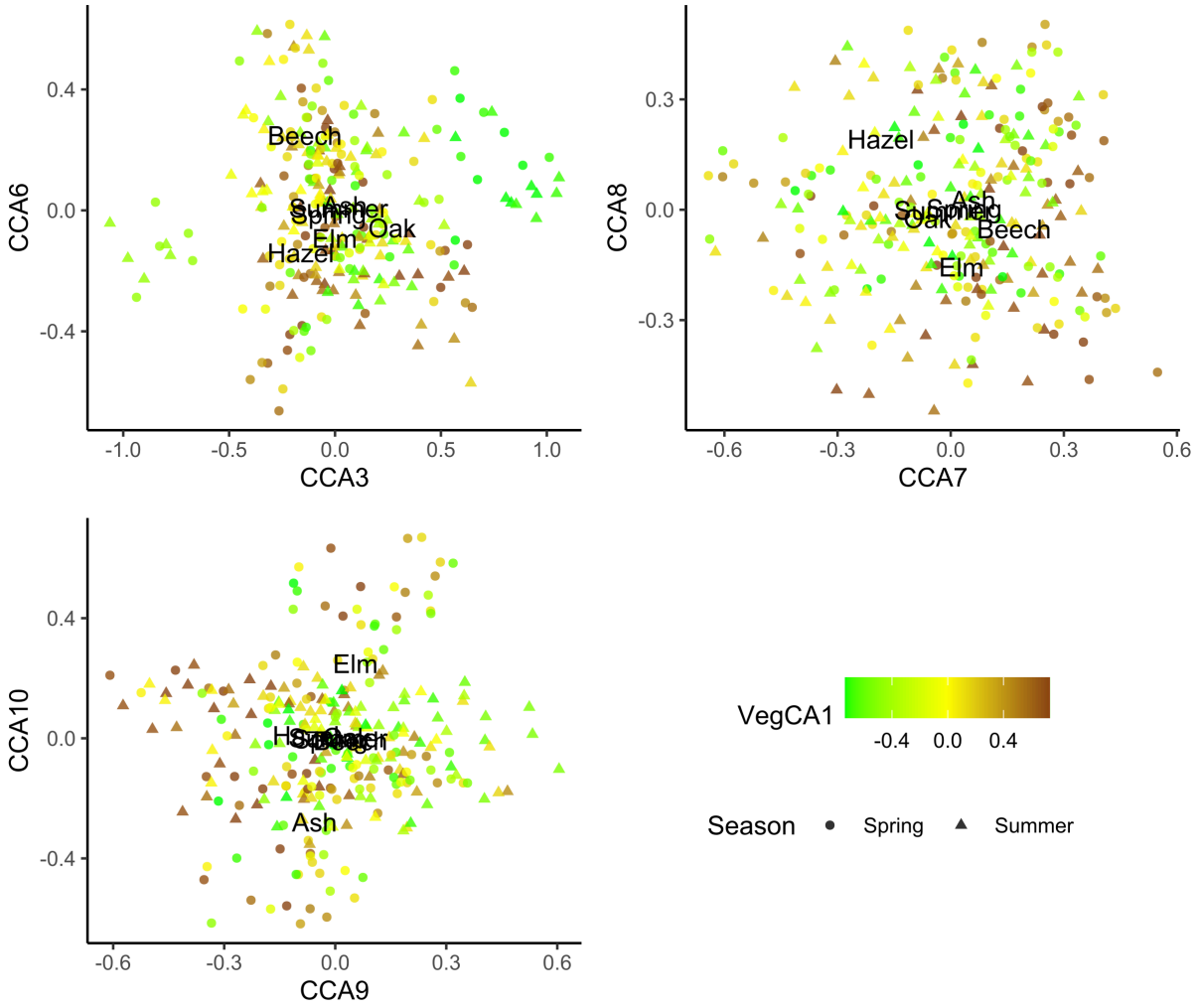


Figure 5.8 The bacterial community composition axes most strongly correlated with unconstrained ground vegetation axis VegCA1. Points indicate the position on each axis of each sample, point shape corresponds to sampling season. Colour indicates vegetation type, green generally indicates herbaceous ground cover, brown indicates litter, or fern ground cover.

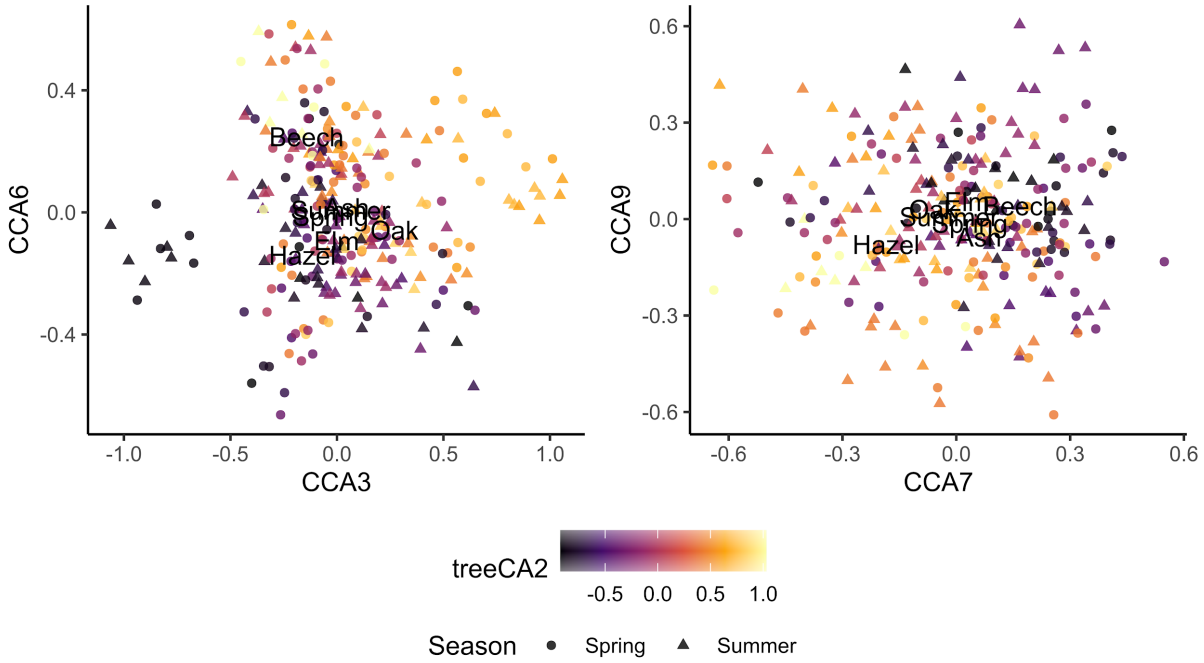


Figure 5.9 CCA axes of the bacterial community composition gradients most strongly correlated with tree composition and importance – axis TreeCA2, unconstrained by woodland age (i.e. axes 3, 6, 7, and 9). Points indicate the position on each axis of each sample, point shape corresponds to sampling season. Colour corresponds to tree communities' differences, low values in blue, among which are trees with recalcitrant leaf litter, such as oak, beech, and holly, high values in yellow correspond to woods with easily decomposable leaves, such as ash, and elm among other trends.

### *5.4.9 Fungal phyla associations*

The CCA scores of the fungal phylum centroids of Ascomycota, Basidiomycota, and Mortierellomycota are displayed in Table 5.4. Generally, these phyla are not strongly associated with one axis or another as their taxa have diverse responses to seasons and litter changes. However, Mortierellomycota, with only 115 ASVs in this CCA analysis, showed strong associations with 3, 8 and 5 (Table 5.4). This indicates an increase in abundance or increased likelihood of occurrence in the summer sampling season, in older woodlands, and a strong preference for woodlands with a low TreeCA2 score which have more, larger trees with recalcitrant leaf litter. They also have a slight association with axis 6, indicating a preference for elm leaves to ash leaves. The ASVs with the highest and lowest median scores are displayed for each axis (Table 5.4 and Table 5.5), excluding ASVs with fewer than three ASVs.

Table 5.4 The median centroid scores of the three fungal (A) or bacterial (B) taxonomic classes (with >2 ASVs) most and least associated with each axis, and the median centroid scores for each axis of the three most common fungal phyla: Ascomycota, Basidiomycota and Mortierellomycota.

| Fungi – ITS1 CCA Axis | Median Scores: Most Negative taxonomic Class | Most Positive taxonomic Class | Asco. | Basidio. | Mortierello. |
|---|---|---|---|---|---|
| 1 | Lecanoromycetes, -0.93; Agaricostilbomycetes, -0.90; Cystobasidiomycetes, -0.90 | Laboulbeniomycetes, 0.24; Taphrinomycetes, 0.53; Microbotryomycetes, 1.39 | -0.2 | -0.1 | -0.1 |
| 2 | Laboulbeniomycetes, -1.10; Agaricostilbomycetes, -0.52; Pezizomycetes, -0.48 | Dothideomycetes, 0.47; Taphrinomycetes, 1.00; Lecanoromycetes, 1.21 | 0.2 | 0.0 | -0.2 |
| 3 | Agaricostilbomycetes, -1.54; Lecanoromycetes, -1.54; Taphrinomycetes, -1.49 | Pezizomycetes, 1.13; Olpidiomycetes, 2.08; Archaeorhizomycetes, 2.12 | 0.2 | 0.0 | 0.9 |
| 4 | Lecanoromycetes, -0.66; Orbiliomycetes, -0.54; Eurotiomycetes, -0.34 | Agaricostilbomycetes, 0.18; Cystobasidiomycetes, 0.37; Taphrinomycetes, 1.17 | -0.1 | 0.0 | 0.1 |
| 5 | Laboulbeniomycetes, -1.18; Orbiliomycetes, -0.81; Mortierellomycetes, -0.47 | Agaricostilbomycetes, 0.43; Olpidiomycetes, 4.13; Archaeorhizomycetes, 5.16 | 0.0 | 0.1 | -0.4 |
| 6 | Laboulbeniomycetes, -0.72; Mortierellomycetes, -0.25; Cystobasidiomycetes, -0.18 | Microbotryomycetes, 0.24; Taphrinomycetes, 0.58; Lecanoromycetes, 0.63 | 0.1 | 0.0 | -0.3 |
| 7 | Archaeorhizomycetes, -3.51; Olpidiomycetes, -0.73; Agaricostilbomycetes, -0.30 | Pezizomycetes, 0.48; Taphrinomycetes, 0.61; Laboulbeniomycetes, 1.60 | 0.1 | 0.2 | 0.1 |
| 8 | Archaeorhizomycetes, -1.99; Taphrinomycetes, -1.45; Olpidiomycetes, -0.94 | Agaricostilbomycetes, 0.28; Lecanoromycetes, 0.36; Cystobasidiomycetes, 0.39 | 0.0 | -0.1 | -0.6 |
| 9 | Microbotryomycetes, -0.39; Agaricostilbomycetes, -0.33; Leotiomycetes, -0.11 | Archaeorhizomycetes, 0.82; Lecanoromycetes, 1.15; Olpidiomycetes, 1.34 | 0.1 | 0.0 | 0.2 |
| 14 | Archaeorhizomycetes, -1.00; Laboulbeniomycetes, -0.45; Lecanoromycetes, -0.33 | Orbiliomycetes, 0.32; Olpidiomycetes, 0.38; Agaricostilbomycetes, 0.43 | 0.0 | 0.0 | 0.0 |

Table 5.5 The median centroid scores of the three bacterial (B) taxonomic classes (with >2 ASVs) most and least associated with each axis.

| Bacteria | | 16S | | Median Scores | |
|---|---|---|---|---|---|
| Axis | df | f | p | Most Negative taxonomic Class | Most Positive taxonomic Class |
| 1 | 1, 249 | 15.44 | 0.005 | Gracilibacteria, -1.83; Sericytochromatia, -1.37; Clostridia, -1.10 | Pla4_lineage, 1.22; BD7-11, 1.23; Spirochaetia, 1.24 |
| 2 | 1, 249 | 12.2 | 0.005 | AD3, -1.44; Acidobacteriia, -0.85; S0134_terrestrial_group, -0.82 | Gracilibacteria, 0.58; Anaerolineae, 0.77; Hydrogenedentia, 1.41 |
| 3 | 1, 249 | 6.87 | 0.005 | BD7-11, -2.39; Pla4_lineage, -2.23; Mollicutes, -1.94 | Nitrososphaeria, 1.07; OLB14, 1.36; Entotheonellia, 1.37 |
| 4 | 1, 249 | 5.49 | 0.005 | BD7-11, -1.82; Mollicutes, -1.70; Sericytochromatia, -1.10 | OLB14, 1.32; AD3, 1.40; Parcubacteria, 2.08 |
| 5 | 1, 249 | 4.51 | 0.005 | Oxyphotobacteria, -1.31; Nitrososphaeria, -0.84; NC10, -0.74 | Spirochaetia, 1.10; Melainabacteria, 1.11; Gracilibacteria, 2.27 |
| 6 | 1, 249 | 3.19 | 0.005 | BD7-11, -1.71; Oxyphotobacteria, -1.25; Gracilibacteria, -0.80 | Bacilli, 0.84; Holophagae, 0.87; Nitrososphaeria, 1.19 |
| 7 | 1, 249 | 3.02 | 0.005 | 4-29-1, -2.13; Oxyphotobacteria, -0.61; Spirochaetia, -0.57 | Hydrogenedentia, 1.10; Latescibacteria, 1.13; Subgroup_25, 1.34 |
| 8 | 1, 249 | 2.49 | 0.01 | AD3, -1.22; Gracilibacteria, -0.57; Saccharimonadia, -0.49 | Hydrogenedentia, 1.34; Latescibacteria, 1.55; Parcubacteria, 2.49 |
| 9 | 1, 249 | 2.4 | 0.005 | 4-29-1, -2.92; AD3, -1.01; Oxyphotobacteria, -0.59 | Lineage_IIa, 1.04; Subgroup_25, 1.34; Entotheonellia, 1.40 |
| 10 | 1, 249 | 2.16 | 0.005 | Oxyphotobacteria, -0.90; Saccharimonadia, -0.78; 4-29-1, -0.69 | OLB14, 0.62; Latescibacteria, 0.72; Gracilibacteria, 2.53 |
| 11 | 1, 249 | 1.51 | 0.005 | Parcubacteria, -1.84; BD7-11, -1.53; Fibrobacteria, -0.99 | Subgroup_25, 1.14; Lineage_IIa, 1.22; JG30-KF-CM66, 1.25 |
| 12 | 1, 249 | 1.41 | 0.005 | Gracilibacteria, -1.97; Hydrogenedentia, -1.06; Entotheonellia, -0.87 | Gitt-GS-136, 0.69; BD2-11_terrestrial_group, 0.71; BD7-11, 1.82 |

## 5.5 Discussion

### *5.5.1 Key findings*

*H1 - Alpha Diversity*

I observed significant differences in microbial alpha diversity and community diversity depending on site and sample characteristics. I hypothesised (H1) that alpha diversity would significantly vary by the litter's species of origin, season, and woodland age, which was confirmed for both litter species and sampling season but not woodland age. Bacterial Shannon diversity increased in summer compared to spring for all litter types, differences in Shannon diversity were more pronounced in the spring with significantly higher bacterial diversity in hazel and beech litter than other litter types, and significantly higher in oak than in ash. By summer only beech had significantly higher bacterial diversity than another tree species' litter – which was oak. The size of the increase in diversity varied by litter type, with oak leaves increasing very little in Shannon diversity and ash increasing much more. Perhaps ash leaf communities changes more because the faster litter decomposition increased the niche breadth of the environment more than the increase in oak – which decomposed little over the period. Previous work has identified significant seasonal change in litter bacterial alpha diversity from *Quercus petraea* dominated woodlands (López-Mondéjar *et al.*, 2015), albeit peaking in the spring rather than summer in opposition to our results. Oak in my study (*Q. robur*) litter changes the least seasonally in my results, indicating that this work of López-Mondéjar *et al.* may represent a conservative measure of seasonal change when compared to other woodland types.

Fungal diversity did not significantly differ by season, but in both seasons beech and hazel samples had higher diversity than other tree species' litter. Despite non-significant seasonal change within species, the response to seasonal change significantly varied between species, with some tree species' litter increasing and others decreasing in diversity. This indicates that different seasonal processes occur across tree species' litter which will make predicting community response to seasonal change from a limited study difficult. Notably there appears to be a slightly positive relationship between leaf recalcitrance and Shannon diversity. Beech and hazel leaf litter had the highest diversity in both seasons, and ash and elm the lowest in the spring season. Although the apparent relationship between recalcitrance and diversity probably relates to the resource availability and microclimate of the decomposing litter, we

do not see the highest diversity on the tree species with the fastest decomposing litter, but instead on those with the slowest.

*H2 - Beta-Diversity*

My second hypothesis (H2) stated that microbial community composition would significantly vary between leaf litter or different tree species, between sampling seasons, and between woodlands of differing age. All these variables were demonstrated to significantly impact woodland microbial community structure, along with additional vegetation characteristics. In bacterial communities, season effects strongly shaped community composition with dramatic changes occurring between the spring and summer sampling seasons, these effects were modulated by the litter's species, with the largest effects occurring in ash and elm communities, and sequentially smaller effects occurring in hazel, beech, and oak leaf communities. These litter communities each differed significantly in their bacterial communities, with several CCA axes describing general gradients spanning between all leaf species or describing differences between pairs or subsets of species' litter.

In fungal communities, the tree species of litter was the most important variable in structuring decomposer communities, with oak leaf communities differing greatly from those of other tree species' litter along the primary CCA axis. Next most important were differences between either elm or ash and beech or hazel communities. Seasonal differences were the second most important variable in explaining fungal community composition, explaining large portions of the variation along CCA axis three, followed by further species variation between beech and hazel litter communities in axis four. Subsequent axes often reflected additional differences in both bacterial and fungal community composition associated with one or more species' litter. For both bacteria and fungi, species differences had significant interactions with season effects such that the magnitude or direction of seasonal change along a CCA axis varied by species. There was some evidence of certain fungal and bacterial taxa switching between litter resources between the spring and summer, providing support for the hypothesis that microbial communities take advantage of the heterogeneity of environments and resources found on diverse litter layers to maximise growth over the year. Seasonal effects on bacterial and fungal beech litter communities have been observed in previous work with additional

changes to their microbial co-occurrence networks (Purahong *et al.*, 2016) relating changes to variations in leaf properties over the course of decomposition. Other work comparing Milkvetch, sedge, and beech litter identified differentiation of bacterial communities and season-tree species interactions in bacterial community composition (Buresova *et al.*, 2019). This study did not identify differences between beech and sedge litter, but also had half the number of tree species' litter replicates and half the median sample read depth of this work. (Buresova *et al.*, 2019; López-Mondéjar *et al.*, 2015; Purahong *et al.*, 2016) *et al.* observe significant shifts in the bacterial diversity of Oak litter over seasons (2015), most different in the summer than the spring, autumn, or winter. They also note that site differences in bacterial composition were more pronounced in the soils than the leaf litter. My results agree with the existing literature that seasonal changes are widespread. Due to the highly replicated, high read depth analysis I have now demonstrated that these changes are highly variable between litter of different common tree species and, for fungal composition, appear to be further altered by the age of the woodland. Future research into the ecological impacts of site management on soil processes should anticipate seasonal effects, and decomposition experiments must be cautious in their extrapolation of decomposition data beyond the study subject species.

Woodland age category also influenced beta-diversity of both bacteria and fungi. I included three potential measures of woodland age in the maximal CCA model; a categorical variable splitting woods into either woodland creation, mature secondary woodland, or ASNW; a ground cover CCA axis reflecting aspects of ground vegetation constrained by this woodland age variable - VegCCA1, and a similar CCA axis reflecting tree community changes constrained to effects of the woodland age variable - TreeCCA1. Plant-microbe interactions may be responsible for microbial community changes, and aspects of tree or vegetation communities are determined by woodland age. By including all of these variables in the potential model structure, the best fitting, most parsimonious model could be selected and indicate whether changes in microbial diversity associated with woodland age category were only indirect effects of age caused by vegetation change. If multiple of these variables were included in the minimum adequate model it might indicate that multiple processes influenced community assembly in these systems. The final models for both bacteria and fungi included VegCCA1 as the third most influential variable in explaining microbial beta-diversity. In the model selection

process, it was the fourth variable to be added to models (after species, season and their interaction term). However, in both models, TreeCCA1, or the woodland age category variable had nearly identical model fit. In fungal models, woodland age category is later added to model structure, accounting for additional differences between that are poorly captured by ground vegetation changes. TreeCCA1 is present in none of the minimum adequate models. As a result, it is clear that woodland age category significantly influences both bacterial and fungal decomposer communities. However, in bacterial communities it is unclear whether woodland age indirectly shapes microbial decomposer communities by causing changes in the ground vegetation community, or whether ground vegetation is simply a better indicator of the age or "quality" of the woodland for woodland specialist decomposers due to its finer resolution description of age-related changes.

*H3 - Trees Influence Litter Microbes to a Greater Extent than Ground Vegetation*

Hypothesis three stated that tree community changes would be more important than ground vegetation changes in shaping the composition of microbial decomposer species on tree leaves. Although aspects of tree diversity and ground diversity unrelated to woodland age were both included in the final model and were significant, ground vegetation differences were added to the model earlier, and explain larger degrees of variation than tree community differences in both models. Although it is difficult to qualify the relative importance of these factors, it is clear that ground vegetation characteristics are unexpectedly important to structuring tree litter decomposer communities in a way that has not been documented previously. As very few studies measure ground vegetation differences between sites in decomposition studies this raises concerns about existing studies of soil diversity and decomposition in woodlands.

### *5.5.2 Limitations*

I have described the limitations of the experimental design and fieldwork portions of this project in the previous chapter (Chapter4). Below I set out challenges resulting from the molecular and bioinformatics work.

*Molecular Methods*

Critics of amplification-based assessments of diversity may disagree with using a read abundances-based measure of taxa frequency. However, as I am comparing between samples with the same approach, this analysis is valid as the taxa are over or under amplified or extracted to the same degree across the study. The read abundance within samples may not reliably be used to infer the abundance of individuals or of taxa biomass but this is not the basis of the analysis. In my previous chapters using the same approach, results generated from occurrence data closely matched those of VST adjusted read abundance data but discarded a large amount of the data.

*CCA Modelling Limitations*

Model structure and selection criteria have huge influences on the outcomes of any study. In this work, I made much use of CCAs and their constrained and unconstrained axes. I also included site triplicates as a conditioning variable and constrained permutation tests by site and by triplicate. Model exploration indicates that including triplicate was not necessary to improve model fit. Although this may be the case mathematically, I am certain that effects of pH and soil texture that had been artificially controlled in the experimental design should be included in the models due to their noted biological importance (Bach *et al.*, 2010; Seaton *et al.*, 2020). CCAs assume a unimodal relationship between model variables and the community data provided, and this may not be valid for all environmental variables, and I considered an RDA which may be more appropriate or less appropriate than a CCA for this reason. I followed suggestions provided by authors of the *vegan* package referring to established model selection procedures (Blanchet *et al.*, 2008). Model refinement is undertaken iteratively with forward addition and backward removal permutation tests. These are probabilistic tests and will give slightly different answers depending on the seed provided, as such they may fail to converge more often than other tests. I did not include soil pH in the maximal CCA model as sites' soil pH were not drawn randomly from a representative set of samples but artificially constrained in site selection and accounted for in the triplicate variable. Including pH in CCA model structure does not meaningfully change fungal models. It is of borderline significance when initially added to bacterial models but leads to a highly complex potential models that fails to converge on an optimal structure. Iterations add pH, treeCA1, woodland age category,

an age:season interaction in that order, and then perpetually add and remove C:N ratio apparently *ad infinitum*. There is no guarantee that any of these new variables, or any of the original variables would remain in the final model should it ever converge. Only the importance of the tree species of the litter, treeCCA1, and an interaction between age category and litter's species are never included in a model. Even should the structure converge, I do not prefer this model due to its increased complexity and its reliance on pH, which I have already artificially controlled. However, I have included a description of it for curious readers.

*Season or succession?*

Throughout this chapter I have described differences between leaf litter communities collected in spring and summer as seasonal effects. However, three concurrent processes occur between the two sampling seasons that are not easily separated. Firstly, the environmental conditions change, causing population changes in the microbial diversity of the soils of each site (Bossio et al. 1998; Bennett et al. 2013; López-Mondéjar et al. 2015; Zhang et al. 2020). Secondly, the litter is physically and chemically changing as it is decomposed producing different niches for microbial species to colonise (Torres et al. 2005; Voříšková and Baldrian 2013; Jackrel et al. 2019; Štursová et al. 2020). Thirdly, as time progresses so do the opportunities for colonisation events to occur, a basic principle of successional dynamics.

Here, as in many studies (e.g. Torres et al. 2005; López-Mondéjar et al. 2015), the combined effect of these processes is referred to and treated as a seasonal effect as litter commonly falls during the winter and autumn months in temperate regions and begins decomposing quickly in the spring and summer months. My study is designed to replicate this natural decomposition as closely as possible. However, another study that placed out litter bags each season and collected them in the next might isolate the effects of season alone compared to succession. However, as the amount of decomposition that occurs during a season varies throughout the year, it is very difficult to separate these effects. Reductionist verses holistic approaches to describing the natural environment both have their place in ecological research. But we should ask what purpose would separating these effects serve as it inevitably decreases the applicability of results to what actually occurs in the natural environment? Readers should consider that the effect of season is a combined effect of seasonal change and succession on decomposing litter. However, they may be comforted by the knowledge that

this more closely follows the natural course of decomposer community change than a study that sought to separate these processes.

### *5.5.3 Tree community and importance and ground cover effects on decomposer communities*

Aspects of ground vegetation composition explained by woodland age, VegCCA1, was the third non-interaction term added to CCA models in model selection for both bacteria and fungi. This was followed by two vegetation variables unassociated with woodland age VegCA1 and TreeCA2 in both models, in that order. Generally, VegCCA1 had higher scores along axes than TreeCA2, which had higher scores than VegCA1. As previously discussed, it is difficult to disentangle the impacts of ground vegetation alone, i.e. the presence or absence of ASNW indicators such as Dog's Mercury, and the age of a woodland in VegCCA1. The impacts of ground vegetation associated with woodland age may be less important than tree community composition in the final model, although the combined effect of VegCCA1 and VegCA1 or VegCCA1 alone is greater than tree community composition. It is unclear whether we can unambiguously say that ground vegetation is a greater influence on tree litter decomposer communities than site tree composition. However, it is clear that ground vegetation effects are important to woodland decomposition in a way we have not previously recognised. Previous studies based in woodlands investigating the top-down pressures of vegetation on soil diversity and litter competition often focus on woody plants alone, or plants over a set height, DBH or stem diameter. The large input of herbaceous plants into the soil in exudates and dead material and the microclimates they create in the litter layer and soil are unmeasured in these studies. Including herbaceous vegetation information in woodland decomposition experiments is not a new idea, Freschet *et al.* (2012) includes herbaceous litter quality when investigating the SMI hypothesis, but in other work understory vegetation data appears to have been gathered but not analysed (Welke & Hope, 2005). Even in studies explicitly investigating the effects of litter input on decomposition have chosen to exclude understory and herbaceous vegetation data (Xiao *et al.*, 2019). Only a small minority of woodland decomposition experiments include any ground vegetation data at all. Studies of grassland decomposition are more likely to include and analyse herbaceous vegetation data (Chuan *et al.*, 2018), but this certainly is not a regular approach. This may be due to limited site replication; if all litter bags are buried in the same area there are no vegetation differences to consider. But more often ecosystems are treated as monoliths, beech stands are beech

stands, birch stands are birch stands, and grasslands are grasslands, regardless of woodland structure or understory composition. In my previous chapters, on the WrEN project soil cores, two sites were extreme outliers in their microbial communities. One was a mature woodland that was similar to agriculture sites, and another an agricultural pasture that was similar to woodlands in its microbial composition. Those sites demonstrate that within habitat categories microbial communities and soil characteristics may be highly variable. In this study, existing sites rarely fit into neat habitat categories. This reflects the impact of anthropogenic influences on woodlands. The cultural and economic drivers of woodland management have changed over time, but their effects cannot be entirely erased from the landscape, making each site different. Differences in ground cover and between otherwise similar woodlands in this study demonstrate that woodland vegetation is unique to each site and these differences cannot be assumed to have a negligible impact on decomposition.

### *5.5.4 Future work*

*Additional Analysis of Existing Data*

Large molecular datasets permit a myriad of analytical approaches, clear directions for further analysis of this dataset would be making use of the *DESeq2* differential expression analysis to determine the ASVs that are most strongly, significantly contrasting in pairs of treatments, e.g. between oak and ash samples, or between woodland creation and ASNW. This has previously been carried out to analyse differences in a highly similar molecular assay of fungi on grapevine leaves in agricultural systems (Pauvert *et al.*, 2020). Such an analysis would be trivial given the bioinformatic work already conducted in *DESeq2* but is beyond the requirements of this chapter. The project design naturally, and intentionally, suits a bipartite network analysis approach. Such networks could be constructed relatively simply from the existing dataset, with the tree species of litter and decomposer species as categories of nodes, with ASV occurrence producing an unweighted network, or ASV transformed abundance producing weighted networks.

*Future Research Directions*

Additional research that directly ties the bacteria and fungi that I have associated with each category of woodland to changes in decomposition rate would be beneficial. These might be conducted in controlled environments, as in (Strickland *et al.*, 2009), to concretely establish

the link between slightly elevated decomposition rates in ASNW over woodland creation and the microbial diversity changes between these woodland types. However, the drawbacks of working in these controlled environments are that they may not replicate effects *in situ* well.

In the introduction, I highlighted the need to test hypotheses of soil ecology across a broad range of habitat types, across biomes, and across continents as the assumption that the microbial soil ecosystem responds identically in all environments is untested. It is important to test whether successional effects in soil microbial composition and functional diversity are present in a range of environments. Repeating a similar experiment in other environments would be useful.

Finally, I have observed very influential effects of multiple litters' origin species on microbial diversity. Generally, my litter was collected from a single site, or even a single tree. We know that litter quality varies between location and individuals, and within individual trees over their lifespan (Trap *et al.*, 2013; Savaci & Sariyildiz, 2020). A further experiment could easily be conducted at a single site, or a limited number of sites, investigating the variability of bacterial and fungal communities on several tree species' litter (perhaps *Q. robur*, *C. avellana,* and *F. excelsior*) taken from young, mature and ancient trees at a number of sites. It would be of interest to see if strong species effects on community composition dominate over regional or tree age effects.

### *5.5.5 Synthesis and applications*

Woodland planting in the UK is set to proceed at pace as national and devolved governments seek methods of mitigating polluting activities (Forest Service, 2006; DEFRA, 2018; Welsh Government, 2018; The Scottish Government, 2019), in addition many landowners are seeking to improve the conservation value of ex-plantation or plantation on ancient woodland sites (PAWS) by thinning non-native conifers and encouraging natural recovery or planting native species. Furthermore, existing high-visibility building projects such as *High-Speed Rail 2* (HS2) are seeking to mitigate the partial or complete grubbing out of ASNW sites with new planting projects or by attempting soil translocation projects on existing or new woodlands (HS2 Ltd,

2017). These projects assert that the soils of ancient woodlands have value and require conservation, both for the seedbank they contain and for their microbial diversity but are highly speculative. In this chapter I have asked, does the functional diversity of woodlands change as they age? How does it change? What factors further alter woodland soil diversity? Combined with the previous chapter on decomposition rates, it may be possible to ask: What impact do changes in diversity have on a critical ecosystem process? My results suggest that the functional diversity of woodland fungal and bacterial decomposers change as woodlands age. Ageing produces a shifting in diversity of numerous taxa, the general increase in the saprotroph containing phyla *Mortierellomycota*, and the saprophytic classes *Agaricostilbomycetes*, in the soil microfauna associated fungi classes such as *Laboulbeniomycetes*, *Orbiliomycetes*, and in ammonia oxidising archaea *Nitrososphaeria*. I have shown that vegetation in the understory and tree canopy are also good predictors of fungal and bacterial diversity, particularly woodland specialist ground cover variation associated with the woodland age category. I have also shown that the litter decomposing strongly determines decomposer community, as well as the season this is taken in. The implications of these results are that decisions regarding tree composition and understory management will have knock on effects on the active microbial diversity of the woodland floor. The age of a woodland either directly influences fungal and bacterial composition, or it encourages the establishment of ground flora that then shape the soil microbiome. The season at which samples are taken from the woodland floor impact the diversity found there. This may mean that the seasonal timing of soil translocations must be well timed to optimise diversity. More woodland soil translocation research is needed, including the importance of seasonal timings. In combination with the previous chapter, this work indicates that differences in microbial diversity between woodland creation and ASNW lead to non-significant but notable increases in decomposition rate. More work must be done to substantiate this, as suggested in the previous chapter, but the early indications are promising.

The nature of a CCA means that axes often disagree about the direction of a relationship of two variables such as tree community composition and woodland age to microbial communities as they partition aspects of the variation off. It is simplest to say that taxa associated with ancient woodlands vary in their preference for different tree compositions and ground vegetation. There is not likely to be a one-size-fits all approach to woodland

planting that maximises functional microbial diversity. However, woodlands of different ages do differ in their microbial diversity and this work indicates that manifest can continue to trust changes in ancient woodland indicator species as indicators of "old" soils. Additionally, they should expect to find different microbial communities in woodlands with an herbaceous understory than one with mostly ferns and accumulated leaf litter. This may be due to direct influences of ground cover, to the microclimate they create, or to the tree canopy changes that encourage these ground cover types. Equally, this work indicates that land managers may expect to find different microbial communities under woods dominated by tough to decompose litter (e.g. beech, oak, holly) than easy to decompose litter (e.g. ash, elm). Ascribing value to one of these habitats or the other is not the purpose of scientific enquiry, but managers seeking to maximise diversity in a large woodland might seek to produce a woodland with compartments that vary along these spectrums. Anecdotally, woodland planting projects have sometimes been undertaken by planting out mixed species and letting the best suited to the habitat survive "[*Plantabis] eos. Novit enim Dominus qui sunt eius*". If all planted woodlands in a large scale national or regional planting project are undertaken like this, it may unintendedly produce diverse individual sites but little landscape level habitat heterogeneity. Homogenous habitats form part of the heterogeneous historic landscape of Britain, and this work indicates that plant species poor habitats such as beech dominated woodlands with a thick litter layer suppressing ground cover produce different microbial habitats that have a place in a diverse landscape. In seeking to plant for an uncertain future a diverse strategy must be taken in large scale planting projects, or the woodlands of the future will be the same everywhere, above and below ground.

**5.6 References**

Aerts, R. (1997) Climate, Leaf Litter Chemistry and Leaf Litter Decomposition in Terrestrial Ecosystems: A Triangular Relationship. *Oikos* . 79 (3), 439–449.

Bach, E.M., Baer, S.G., Meyer, C.K. & Six, J. (2010) Soil texture affects soil microbial and structural recovery during grassland restoration. *Soil biology & biochemistry*. 42 (12), 2182–2191.

Baldrian, P. (2017) Forest microbiome: diversity, complexity and dynamics. *FEMS microbiology reviews*. 41 (2), 109–130.

Baldrian, P., Kolařík, M., Stursová, M., Kopecký, J., Valášková, V., Větrovský, T., Zifčáková, L., Snajdr, J., Rídl, J., Vlček, C. & Voříšková, J. (2012) Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *The ISME journal*. 6 (2), 248–258.

Bani, A., Pioli, S., Ventura, M., Panzacchi, P., Borruso, L., Tognetti, R., Tonon, G. & Brusetti, L. (2018) The role of microbial community in the decomposition of leaf litter and deadwood. *Applied soil ecology: a section of Agriculture, Ecosystems & Environment*. 12675–84.

Bayranvand, M., Akbarinia, M., Jouzani, G.S., Gharechahi, J., Kooch, Y. & Baldrian, P. (2020) Composition of soil bacterial and fungal communities in relation to vegetation composition and soil characteristics along an altitudinal gradient. *FEMS microbiology ecology*.

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., De Wit, P., Sánchez-García, M., Ebersberger, I., de Sousa, F., Amend, A.S., Jumpponen, A., Unterseher, M., Kristiansson, E., Abarenkov, K., Bertrand, Y.J.K., Sanli, K., Eriksson, K.M., Vik, U., *et al.* (2013) Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data Michael Bunce (ed.). *Methods in ecology and evolution / British Ecological Society*. 25.

Bennett, A.E., Daniell, T.J., Öpik, M., Davison, J., Moora, M., Zobel, M., Selosse, M.-A. & Evans, D. (2013) Arbuscular mycorrhizal fungal networks vary throughout the growing season and between successional stages. PloS one. 8 (12), e83241.

Berenstecher, P., Vivanco, L., Pérez, L.I., Ballaré, C.L. & Austin, A.T. (2020) Sunlight Doubles Aboveground Carbon Loss in a Seasonally Dry Woodland in Patagonia. *Current biology: CB*.

Blanchet, F.G., Legendre, P. & Borcard, D. (2008) Forward selection of explanatory variables. *Ecology*. 89 (9), 2623–2632.

Bokhorst, S. & Wardle, D.A. (2013) Microclimate within litter bags of different mesh size: Implications for the 'arthropod effect' on litter decomposition. *Soil biology & biochemistry*. 58147–152.

Bossio, D.A., Scow, K.M., Gunapala, N. & Graham, K.J. (1998) Determinants of Soil Microbial Communities: Effects of Agricultural Management, Season, and Soil Type on Phospholipid Fatty Acid Profiles. Microbial ecology. 36 (1), 1–12.

Bradford, M.A., Berg, B., Maynard, D.S., Wieder, W.R. & Wood, S.A. (2016) Understanding the dominant controls on litter decomposition. *The Journal of ecology*. 104 (1), 229–238.

Buresova, A., Kopecky, J., Hrdinkova, V., Kamenik, Z., Omelka, M. & Sagova-Mareckova, M. (2019) Succession of Microbial Decomposers Is Determined by Litter Type, but Site Conditions Drive Decomposition Rates. *Applied and environmental microbiology*. 85 (24), .

Callahan, B.J., McMurdie, P.J. & Holmes, S.P. (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal*. 11 (12), 2639–2643.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A. & Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods*. 13 (7), 581–583.

Callahan, B.J., Sankaran, K., Fukuyama, J.A., McMurdie, P.J. & Holmes, S.P. (2016) Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses. *F1000Research*. 51492.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., 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 Suppl 14516–4522.

Chuan, X., Carlyle, C.N., Bork, E.W., Chang, S.X. & Hewins, D.B. (2018) Long-Term Grazing Accelerated Litter Decomposition in Northern Temperate Grasslands. *Ecosystems* . 21 (7), 1321–1334.

Creamer, C.A., Filley, T.R., Boutton, T.W. & Rowe, H.I. (2016) Grassland to woodland transitions: Dynamic response of microbial community structure and carbon use patterns. *Journal of Geophysical Research: Biogeosciences*. 121 (6), 2016JG003347.

DEFRA (2018) *A Green Future: Our 25 Year Plan to Improve the Environment*.

Delavaux, C.S., Bever, J.D., Karppinen, E.M. & Bainard, L.D. (2020) Keeping it cool: Soil sample cold pack storage and DNA shipment up to 1 month does not impact metabarcoding results. *Ecology and evolution*. 671.

Dopheide, A., Xie, D., Buckley, T.R., Drummond, A.J. & Newcomb, R.D. (2018) Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity Michael Bunce (ed.). *Methods in ecology and evolution / British Ecological Society*. 718966.

Falconer, J.G., Wright, J.W. & Beall, H.W. (1933) The Decomposition of Certain Types of Forest Litter Under Field Conditions. *American journal of botany*. 20 (3), 196–203.

Fierer, N. & Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*. 103 (3), 626–631.

Forest Service (2006) *NI Forestry - a strategy for sustainable growth*.

Freschet, G.T., Aerts, R. & Cornelissen, J.H.C. (2012) Multiple mechanisms for trait effects on litter decomposition: moving beyond home-field advantage with a new hypothesis: Substrate-matrix quality interactions in decay. *The Journal of ecology*. 100 (3), 619–630.

Gardes, M. & Bruns, T.D. (1993) ITS primers with enhanced specificity for basidiomycetes--application to the identification of mycorrhizae and rusts. *Molecular ecology*. 2 (2), 113–118.

George, P.B.L., Lallias, D., Creer, S., Seaton, F.M., Kenny, J.G., Eccles, R.M., Griffiths, R.I., Lebron, I., Emmett, B.A., Robinson, D.A. & Jones, D.L. (2019) Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. *Nature communications*. 10 (1), 1107.

Gessner, M.O., Swan, C.M., Dang, C.K., McKie, B.G., Bardgett, R.D., Wall, D.H. & Hättenschwiler, S. (2010) Diversity meets decomposition. *Trends in ecology & evolution*. 25 (6), 372–380.

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M. & Whiteley, A.S. (2011) The bacterial biogeography of British soils. *Environmental microbiology*. 13 (6), 1642–1654.

Hackl, E., Pfeffer, M., Donat, C., Bachmann, G. & Zechmeister-Boltenstern, S. (2005) Composition of the microbial communities in the mineral soil under different types of natural forest. *Soil biology & biochemistry*. 37 (4), 661–671.

Hahn, C. & Lunt, D. (2019) *metaBEAT*. Github.

Hamonts, K., Bissett, A., Macdonald, B.C.T., Barton, P.S., Manning, A.D. & Young, A. (2017) Effects of ecological restoration on soil microbial diversity in a temperate grassy woodland. *Applied soil ecology: a section of Agriculture, Ecosystems & Environment*. 117–118117–128.

Handelsman, J. (2004) Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and molecular biology reviews: MMBR*. 68 (4), 669–685.

Hättenschwiler, S., Tiunov, A.V. & Scheu, S. (2005) Biodiversity and Litter Decomposition in Terrestrial Ecosystems. *Annual review of ecology, evolution, and systematics*. 36 (1), 191–218.

HS2 Ltd (2017) *High Speed Two Phase One: London-West Midlands Ancient Woodland Strategy*.

Jackrel, S.L., Gilbert, J.A. & Wootton, J.T. (2019) The Origin, Succession, and Predicted Metabolism of Bacterial Communities Associated with Leaf Decomposition. mBio. 10 (5)

Jacob, M., Viedenz, K., Polle, A. & Thomas, F.M. (2010) Leaf litter decomposition in temperate deciduous forest stands with a decreasing fraction of beech (Fagus sylvatica). *Oecologia*. 164 (4), 1083–1094.

Jangid, K., Williams, M.A., Franzluebbers, A.J., Schmidt, T.M., Coleman, D.C. & Whitman, W.B. (2011) Land-use history has a stronger impact on soil microbial community composition than aboveground vegetation and soil properties. *Soil biology & biochemistry*. 43 (10), 2184–2193.

Jiao, S., Chen, W., Wang, J., Du, N., Li, Q. & Wei, G. (2018) Soil microbiomes with distinct assemblies through vertical soil profiles drive the cycling of multiple nutrients in reforested ecosystems. *Microbiome*. 6 (1), 146.

Jonard, M., Andre, F. & Ponette, Q. (2008) Tree species mediated effects on leaf litter

dynamics in pure and mixed stands of oak and beech. *Canadian journal of forest research. Journal canadien de la recherche forestiere*. 38 (3), 528–538.

Kitson, J.J.N., Hahn, C., Sands, R.J., Straw, N.A., Evans, D.M. & Lunt, D.H. (2018) Detecting host-parasitoid interactions in an invasive Lepidopteran using nested tagging DNA metabarcoding. *Molecular ecology*.

Lladó, S., López-Mondéjar, R. & Baldrian, P. (2017) Forest Soil Bacteria: Diversity, Involvement in Ecosystem Processes, and Response to Global Change. *Microbiology and molecular biology reviews: MMBR*. 81 (2), e00063–16.

López-Mondéjar, R., Voříšková, J., Větrovský, T. & Baldrian, P. (2015) The bacterial community inhabiting temperate deciduous forests is vertically stratified and undergoes seasonal dynamics. *Soil biology & biochemistry*. 8743–50.

Love, M.I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*. 15 (12), 550.

Mackay, J.E., Cunningham, S.C. & Cavagnaro, T.R. (2016) Riparian reforestation: are there changes in soil carbon and soil microbial communities? *The Science of the total environment*. 566-567960–967.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 17 (1), 10–12.

Ma, S., De Frenne, P., Boon, N., Brunet, J., Cousins, S.A.O., Decocq, G., Kolb, A., Lemke, I., Liira, J., Naaf, T., Orczewska, A., Plue, J., Wulf, M. & Verheyen, K. (2019) Plant species identity and soil characteristics determine rhizosphere soil bacteria community composition in European temperate forests. *FEMS microbiology ecology*. 95 (6), .

McMurdie, P.J. & Holmes, S. (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*. 8 (4), e61217.

Nielsen, U.N., Ayres, E., Wall, D.H. & Bardgett, R.D. (2011) Soil biodiversity and carbon cycling: a review and synthesis of studies examining diversity–function relationships. *European journal of soil science*. 62 (1), 105–116.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R.,

O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. & Wagner, H. (2019) *vegan: Community Ecology Package*.

Pauvert, C., Buée, M., Laval, V., Edel-Hermann, V., Fauchery, L., Gautier, A., Lesur, I., Vallance, J. & Vacher, C. (2019) Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal ecology*. 4123–33.

Pauvert, C., Vallance, J., Delière, L., Bohan, D.A., Buée, M. & Vacher, C. (2020) Inferred microbial networks lack replicability: consequences for next-generation biomonitoring. bioRxiv p.642199.

Peay, K.G., Bidartondo, M.I. & Arnold, A.E. (2010) Not every fungus is everywhere: scaling to the biogeography of fungal–plant interactions across roots, shoots and ecosystems. *The New phytologist*. 185 (4), 878–882.

Pereira, S., Burešová, A., Kopecky, J., Mádrová, P., Aupic-Samain, A., Fernandez, C., Baldy, V. & Sagova-Mareckova, M. (2019) Litter traits and rainfall reduction alter microbial litter decomposers: the evidence from three Mediterranean forests. *FEMS microbiology ecology*. 95 (12), .

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team (2019) *nlme: Linear and Nonlinear Mixed Effects Models*.

Porre, R.J., van der Werf, W., De Deyn, G.B., Stomph, T.J. & Hoffland, E. (2020) Is litter decomposition enhanced in species mixtures? A meta-analysis. *Soil biology & biochemistry*. 145107791.

Preston, C.M., Nault, J.R., Trofymow, J.A., Smyth, C. & CIDET Working Group (2009) Chemical Changes During 6 Years of Decomposition of 11 Litters in Some Canadian Forest Sites. Part 1. Elemental Composition, Tannins, Phenolics, and Proximate Fractions. *Ecosystems* . 12 (7), 1053–1077.

Pu, G., Du, J., Ma, X., Lv, Y., Jia, Y., Jia, X. & Tian, X. (2014) Contribution of ambient atmospheric exposure to Typha angustifolia litter decomposition in aquatic environment. *Ecological engineering*. 67144–149.

Purahong, W., Kapturska, D., Pecyna, M.J., Jariyavidyanont, K., Kaunzner, J., Juncheed, K., Uengwetwanit, T., Rudloff, R., Schulz, E., Hofrichter, M., Schloter, M., Krüger, D. & Buscot, F. (2015) Effects of forest management practices in temperate beech forests on bacterial and fungal communities involved in leaf litter degradation. *Microbial ecology*. 69 (4), 905–913.

Purahong, W., Wubet, T., Lentendu, G., Schloter, M., Pecyna, M.J., Kapturska, D., Hofrichter, M., Krüger, D. & Buscot, F. (2016) Life in leaf litter: novel insights into community dynamics of bacteria and fungi during litter decomposition. *Molecular ecology*. 25 (16), 4059–4074.

van der Putten, W.H., Bardgett, R.D., Bever, J.D., Bezemer, T.M., Casper, B.B., Fukami, T., Kardol, P., Klironomos, J.N., Kulmatiski, A., Schweitzer, J.A., Suding, K.N., Van de Voorde, T.F.J. & Wardle, D.A. (2013) Plant-soil feedbacks: the past, the present and future challenges Michael Hutchings (ed.). *The Journal of ecology*. 101 (2), 265–276.

Savaci, G. & Sariyildiz, T. (2020) Effects of stand age on litter quality, decomposition rate and nutrient release of Kazdagi fir (Abies nordmanniana subsp. equi-trojani). *iForest - Biogeosciences and Forestry*. 13 (5), 396.

Schimel, J.P. & Schaeffer, S.M. (2012) Microbial control over carbon cycling in soil. *Frontiers in microbiology*. 3348.

Schneider, T., Keiblinger, K.M., Schmid, E., Sterflinger-Gleixner, K., Ellersdorfer, G., Roschitzki, B., Richter, A., Eberl, L., Zechmeister-Boltenstern, S. & Riedel, K. (2012) Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. *The ISME journal*. 6 (9), 1749–1762.

Seaton, F.M., George, P.B.L., Lebron, I., Jones, D.L., Creer, S. & Robinson, D.A. (2020) Soil textural heterogeneity impacts bacterial but not fungal diversity. *Soil biology & biochemistry*. 144107766.

Strickland, M.S., Lauber, C., Fierer, N. & Bradford, M.A. (2009) Testing the functional significance of microbial community composition. *Ecology*. 90 (2), 441–451.

Štursová, M., Šnajdr, J., Koukol, O., Tláskal, V., Cajthaml, T. & Baldrian, P. (2020) Long-term decomposition of litter in the montane forest and the definition of fungal traits in the

successional space. Fungal ecology. 46100913.

Sundsdal, A., Graae, B.J., Speed, J.D.M., Bukombe, J., Mtweve, P.J., Arneberg, M.K., Haukenes, V.L., Grevskott, R.T. & Smith, S.W. (2020) Teatime in the Serengeti: macrodetritivores sustain recalcitrant plant litter decomposition across human-modified tropical savannahs. *Plant and soil*.

Tedersoo, L., Anslan, S., Bahram, M., Põlme, S., Riit, T., Liiv, I., Kõljalg, U., Kisand, V., Henrik Nilsson, R., Hildebrand, F., Bork, P. & Abarenkov, K. (2016) *Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi*. [Online] [online]. Available from: http://mycokeys.pensoft.net/articles.php?id=4852 (Accessed 11 November 2016).

The Scottish Government (2019) *Scotland's Forestry Strategy 2019–2029*.

Torres, P.A., Abril, A.B. & Bucher, E.H. (2005) Microbial succession in litter decomposition in the semi-arid Chaco woodland. Soil biology & biochemistry. 37 (1), 49–54.

Trap, J., Hättenschwiler, S., Gattin, I. & Aubert, M. (2013) Forest ageing: An unexpected driver of beech leaf litter quality variability in European forests with strong consequences on soil processes. *Forest ecology and management*. 302338–345.

Trogisch, S., He, J.-S., Hector, A. & Scherer-Lorenzen, M. (2016) Impact of species diversity, stand age and environmental factors on leaf litter decomposition in subtropical forests in China. *Plant and soil*. 400 (1), 337–350.

Urbanová, M., Šnajdr, J. & Baldrian, P. (2015) Composition of fungal and bacterial communities in forest litter and soil is largely determined by dominant trees. *Soil biology & biochemistry*. 8453–64.

Uroz, S., Buée, M., Deveau, A., Mieszkin, S. & Martin, F. (2016) Ecology of the forest microbiome: Highlights of temperate and boreal ecosystems. *Soil biology & biochemistry*. 103471–488.

Voříšková, J. & Baldrian, P. (2013) Fungal community on decomposing leaf litter undergoes rapid successional changes. The ISME journal. 7 (3), 477–486.

Wagg, C., Schlaeppi, K., Banerjee, S., Kuramae, E.E. & van der Heijden, M.G.A. (2019) Fungal-

bacterial diversity and microbiome complexity predict ecosystem functioning. *Nature communications*. 10 (1), 1–10.

van der Wal, A., Geydan, T.D., Kuyper, T.W. & de Boer, W. (2013) A thready affair: linking fungal diversity and community dynamics to terrestrial decomposition processes. *FEMS microbiology reviews*. 37 (4), 477–494.

Wardle, D.A., Bonner, K.I. & Nicholson, K.S. (1997) Biodiversity and Plant Litter: Experimental Evidence Which Does Not Support the View That Enhanced Species Richness Improves Ecosystem Function. *Oikos* . 79 (2), 247–258.

Welke, S.E. & Hope, G.D. (2005) Influences of stand composition and age on forest floor processes and chemistry in pure and mixed stands of Douglas-fir and paper birch in interior British Columbia. *Forest ecology and management*. 219 (1), 29–42.

Welsh Government (2018) *Woodlands for Wales*.

White, T.J., Bruns, T.D., Lee, S.B. & Taylor, J.W. (1990) 'Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics', in M. A. Innis, D. H. Gelfand, J. J. Sninsky, & T. J. White (eds.) *PCR Protocols: A Guide to Methods and Applications*. [Online]. New York: Academic Press. pp. 315–322.

Wu, X., Xu, H., Tuo, D., Wang, C., Fu, B., Lv, Y. & Liu, G. (2020) Land use change and stand age regulate soil respiration by influencing soil substrate supply and microbial community. *Geoderma*. 359113991.

Xiao, W., Chen, H.Y.H., Kumar, P., Chen, C. & Guan, Q. (2019) Multiple interactions between tree composition and diversity and microbial diversity underly litter decomposition. *Geoderma*. 341161–171.

Yue, K., García-Palacios, P., Parsons, S.A., Yang, W., Peng, Y., Tan, B., Huang, C. & Wu, F. (2018) Assessing the temporal dynamics of aquatic and terrestrial litter decomposition in an alpine forest Maria Briones (ed.). *Functional ecology*. 32 (10), 2464–2475.

Zhang, K., Delgado-Baquerizo, M., Zhu, Y.-G. & Chu, H. (2020) Space Is More Important than Season when Shaping Soil Microbial Communities at a Large Spatial Scale. mSystems. 5 (3), .

Zhu, W., Cai, X., Liu, X., Wang, J., Cheng, S., Zhang, X., Li, D. & Li, M. (2010) Soil microbial population dynamics along a chronosequence of moist evergreen broad-leaved forest succession in southwestern China. *Journal of Mountain Science; Dordrecht*. 7 (4), 327–338.

# Chapter 6. Discussion

## 6.1 Legacies of Management and Societal Shifts in UK Woodland

Woodland in the UK is still in its greatest period of expansion in recorded history (Cahalan *et al.*, 2011). But as critics of these trends have noted, this does not mean that woodlands in the UK are not under threat from land-use change or to improvement in one aspect or another according to the fashions of the day (Rackham, 2012). The UK and devolved governments appear to be set on high targets for woodland creation in order to reach multiple commitments to greening the UK (Forest Service, 2006; DEFRA, 2018; Welsh Government, 2018; The Scottish Government, 2019). Woodland creation is once again a policy priority, but also of public interest with 88% of UK respondents recently surveyed agreeing or strongly agreeing that "a lot more trees should be planted" as a "response to the threat of climate change" (Forest Research, 2019). The issue has also been popular with prominent figures associated with the "Green" movement in the UK under the banner of rewilding, featuring strongly in modern popular environmental literature such as the controversial book *Feral* by George Monbiot (2014).

Existing woodland resources throughout Britain are still responding to the legacy of a changing forestry policy stretching back to the Middle Ages. Forestry legislation, due to its importance to the inhabitants of the British Isles, has been foundational to UK law, beginning with one of the longest standing English statutes *The Charter of the Forest* (*Carta de Foresta*); enacted in 1217 and active in aspects of UK law until the 1971 *Wild Creatures and Forest Laws Act (Shoenberger, 2015)*. Historic governmental actions on woodlands and forest (in its original definition) have introduced new herbivores, compartmentalised and fragmented woodland, planted and grubbed-out woodlands, converted native trees to non-native, and restocked ancient semi-natural woodlands (ASNW) with Plantations on Ancient Woodland Sites (PAWS) (Rackham, 1986, 2012). Modernisation of our industrial and forestry sectors for a thousand years has made the planting of various species, varieties, or provenances fashionable for a time and the relics of the forestry sector linger in most woodlands (Rackham, 2012). Most importantly, this combination of state involvement and changing industry has altered the culture of the forestry sector and academic forestry work. Woodland is still grubbed-out

today; and although protected woodland is not immune from these threats (Rackham, 2008; HS2 Ltd, 2017) its destruction or modification is far more likely to receive attention than similar quality mature secondary woodland.

The legacy of UK forestry policy, industrial history, environmental change, and the recent cultural shifts influence the interpretation and application of empirical questions in this work. It has shaped the canopy structure, management, and planting choice of all the woodlands in this study. Additionally, it has affected the fragmenting of green space, the area of woodlands and compartments of Ancient-Semi Natural Woodland (ASNW), and the shape of these woodlands, each of which has implications for the processes of microbial dispersal and ecology relevant to chapters 3 and 5 (Humphrey *et al.*, 2015). Interpretation of the overall results from this thesis on the effects of woodland age on soil biotic and abiotic factors cannot be separated from their relevance to three nationally important conversations: Why should older woodlands be conserved? Why should we plant new woodland? And how can we mitigate the loss of ASNW to land-use change? This research would not have been conducted if those questions were not of importance to the funders of my PhD - The Woodland Trust - and the researchers involved in the project, including myself. Previous generations of forestry scientists and forestry economists who did not appreciate this blind spot provided insight into forestry policy that led to the decisions that shape the landscape today (Rackham, 2012). Historically, these include the restocking of ancient woodland with conifer, the advice provided to landowners to plant for markets that would disappear by the time trees came to maturity, and a litany of steps that has distanced the public from natural spaces (Rackham, 1986, 2012), potentially influencing the British psyche (Milligan & Bingley, 2007; Maller *et al.*, 2009). These kinds of missteps are not unique to forestry research (Kuhn, 1970; Feyerabend, 1978), but the impact of landscape management has such a potential to profoundly impact the public that they are less tolerable. The lessons from the past are clear. Those who advise on topics of landscape management must do so with the understanding that their views are likely shaped by a subjective conversation about the land they work in (Mace, 2014), and the appreciation that policy decisions have impacts that alter the countryside for centuries.

Regarding this work, it means that recommendations for woodland management are provided with the understanding that they are part of a national conversation with many stakeholders.

### 6.2 Brief Summary of Findings

In this work I engaged in two classical empirical ecological experiments that have been modernised using culture independent molecular microbial techniques. In the first experiment, I analysed 210 soil cores collected as part of the Woodland creation & Ecological Networks (WrEN) project. The classical approach reanalysed multiple soil measurements with mixed-effect models to determine how agricultural sites and woodlands of differing age differed in their soil chemistry and structure. I advanced this conventional method of measuring soils (that mostly focus on abiotic factors) by seeking to identify differences in soil fungal and bacterial communities using advances in molecular biology (i.e. massively parallel sequencing technology known as next-generation sequencing, NGS). Combined, I demonstrated significant changes to soil organic carbon, total nitrogen, C:N ratio, bulk density (compaction), and ammonia resulting from afforestation and succession. I identified no significant differences between woodlands at differing successional stages, but variability in whether woodland types significantly differed from either arable or pasture sites. Generally, the results indicated that land used types differentiated in their soil chemistry along a gradient of disturbance with arable farmland at one end, then pasture, young woodland, and mature woodland in between with ASNW at the end of the lowest disturbance. These differences were mirrored in the microbial beta-diversity of these sites. After accounting for differences in soil chemistry, the land use type, or the age category of a woodland, remained an important predictor of its microbial composition. Resultantly, the abiotic influences on diversity during afforestation and succession are likely added to by concomitant biological phenomena, such as taxonomic filtering, adaptation and evolution. . Woodland microbial composition continues to differentiate from agricultural microbial compositions for hundreds of years after its creation; again, beyond the amounts predicted by their soil chemistry alone.

In the second project, I sought to relate changes in microbial diversity to key ecological processes by comparing the microbial communities (generated using NGS) associated with decomposing leaf litter over time. Using a fully-replicated, ambitious experimental design, I

first quantified the vegetation characteristics and soil properties of sites and related them to differences in leaf decomposition rate. Site characteristics, including soil C:N ratio, the dominance of trees with difficult to decompose litter, and ground cover variations relating to the degree of herbaceous or leaf litter cover, all significantly influenced the litter weight lost to decomposition. The age category of the woodland did not significantly influence decomposition rates. However, ASNW were predicted to decompose 16% more leaf material by summer than woodland creation sites. Given the marginal significance of the results, I suspect that a different approach to preparing the samples for calculating weight loss may have yielded significant results by improving precision and may change when I am to process and add the final autumn season of litter weights. Molecular analysis of the bacterial and fungal communities of these leaves revealed significant variations in alpha diversity between tree species' litter, the bacterial component of this also significantly increased in diversity between spring and summer. Analysis of the microbial community composition showed that litter's species, season, and vegetation characteristics significantly affected the fungal and bacterial communities of decomposing leaves. The age of woodlands affected both bacterial and fungal communities indirectly via changes in the ground vegetation cover of sites but was also included as a separate additional variable in the minimum adequate model of fungal community composition. This indicates that after accounting for or controlling differences in soil properties, woodland age is important in determining community structure in addition to, and independently of, vegetation changes in these systems.

## 6.3 Implications for Ecological Theory

### *6.3.1 Niche and neutral processes, and the role of time in microbial community assembly*

The relative importance of environmental constraints and stochastic events in determining the composition of a community is a fundamental question in ecology (Wennekes *et al.*, 2012). At its core, this question asks: are species where they are because of adaptation or chance? Adaptations to take advantage of certain environmental conditions are niche processes; dispersal, ecological drift, and allee effects are neutral processes (Barberán *et al.*, 2014). In terrestrial systems the lifespans of most species, the size of their ranges, and their capacity for movement present challenges to studies of niche and neutral processes. Species generally take a long time to adapt to a new environment. Additionally, the effects of a hostile environment on an individual may take a long time to kill it or demonstrate reduced fecundity.

Source-sink population relationships and ecological traps may lead to the appearance of habitat suitability when in reality it is dispersal that maintains population sizes, with implications for measuring fitness (Kozłowski, 1993) and niche-differentiation (Holt, 1996). Disentangling these may be highly complex in traditional faunal studies. Microbial populations resolve some of these issues (Nemergut *et al.*, 2013). They move comparatively slowly, often have limited short-range dispersal but are capable of long-range dispersal, they are often short-lived or at least have short generation times. Certain microbes are easily manipulable *in situ* and *ex situ* via inoculation, nutrient application or limitation, fungicides, antibiotics, heat, or radiation. As such, microbial systems *in situ* have been used to explore neutral processes in community assembly, particularly by using space to separate neutral from niche processes of dispersal and environmental constraints (Dumbrell *et al.*, 2010; Chen *et al.*, 2019). Time can also be used to separate niche and neutral processes once environmental conditions have become relatively stable (Bahram *et al.*, 2015; Rapacciuolo & Blois, 2019). Site characteristics between mature and ASNW are comparatively similar, but hundreds of years of dispersal, and drift separate them. Differences that do exist between them in their soil characteristics and vegetation can be quantified. After doing so differences that remain can be attributed to neutral and biotic processes occurring over time.

This work was the first study to compare the soil microbial communities of farmland and woodlands across a chronosequence relevant to the woodland policy and conservation framework relevant to the UK and to broadleaf woodland succession found in the UK landscape. This work separated the influences of soil and vegetation gradients within the landscape and across the chronosequence from the effects of woodland age itself. Over this time the environmental pressures present at sites change, creating new ecological niches and eliminating old ones. Neutral forces, such as dispersal, slowly alter the pool of species from which communities may assemble (Barberán *et al.*, 2014). The rapid generation times of microbial species may mean that a degree of local adaptation may occur in early arrivals to the habitat in ecological timescales (Ayres *et al.*, 2009). Outcomes that may reflect this kind of local adaption of bacterial communities to environmental input have been shown to occur in the decomposition of locally varying *Alnus rubra* leaf aromatics with remarkable spatial specificity, down to scales of less than a kilometre (Jackrel *et al.*, 2019). In my work, I have tried to separate identifiable niche variation from unexplained, possible neutral, variation

associated with time. I have done this by including land use or woodland age category as variables in models alongside the variables I have identified to change in response to succession. In this way, I have used models to ask the question: Is niche change associated with succession enough to explain variations in microbial communities or is a neutral variable - time - also required?

Time has been needed as a variable necessary to explain compositional changes of microbial diversity, but not absolute diversity changes in both projects. As woodlands establish and age, they stochastically gain microbial species from dispersal events. Early on in woodland establishment, we might expect disturbance and extreme habitat heterogeneity increase the niche diversity of young woodlands. In the youngest woodlands, much of the habitat has hardly changed from the previous land use. In my young woodland study sites grasses dominate, the woods are open and full of sunlight, and the furrows of the agricultural land are sometimes visible (indeed the features of agricultural landscapes are often identifiable even in ancient woodlands). In these environments taxa that are adapted to woodlands, grasslands, and agricultural systems might all find ideal conditions within patches of the young woodlands (Chen *et al.*, 2019). As the habitat transitions to mature woodland these niches are eroded and the taxa within them must adapt to be competitive with woodland species or, more likely, be replaced by them. In this way, the importance of competition and mutualistic processes might be thought to become more important to microbial community composition in woodlands over time.

How long does it generally take for these woodland taxa to reach woodlands? Previous work indicates that stark changes have already occurred in the microbial composition of young woodlands by the time they reach a couple of decades in age (Jiao *et al.*, 2018), but no studies have rigorously considered change over longer successional time scales using next-generation sequencing methods. My work indicates that community assembly does not finish after the first few decades of woodland establishment. Woodland microbial communities continue to differentiate away from agricultural ones for hundreds of years after their establishment. This may be related to mutualistic associations with woodland ground flora that can also take hundreds of years to arrive at a site (Hermy *et al.*, 1999), or due to changes in soil chemistry.

However, the models used in this thesis indicate that after accounting for these changes, woodland age - time - remains important for explaining community composition. If this is correct, it has implications for how we value the age of a woodland in conservation decisions.

### *6.3.2 Top-down or bottom-up control of microbial communities*

Classical systems approaches to viewing woodlands describe them as complex interchanges between parent rock materials, soils, vegetation, and climate (Trudgill, 1988). Each of these factors can influence microbial populations and traditionally trees have been thought of as the dominant ecosystem engineers of the woodland soil microbiome (Uroz *et al.*, 2016) as the largest primary producers in the system. Generally, my results agree that vegetation exerts significant pressure on the composition of soil litter layer communities. Bottom-up control of woodland microbial communities from the soil properties at the site are also identified in the work. However, this thesis is one of the first studies that has attempted to determine the effect of differences in ground cover on the litter microbiome in addition to tree community composition. In doing so, I have revealed that ground vegetation explains a greater amount of variation in the composition of the litter microbial communities. Additionally, few studies have quantified variations in vegetation community composition in the detail that I have by using multivariate community measures. In failing to do so important gradients of change in tree communities that impact the soil microbiome have been missed. Had I quantified my sites using a traditional NDVI, dominant tree, or other categorical approaches I would have been unable to detect the importance of tree composition in determining decomposition rate and the associated microbial community structure. These categories may be useful for habitat and landscape-scale decision making and planning, but this work demonstrates that they are not sufficient explanatory tools for describing the impact of vegetation on soil communities. Despite the considerable effort needed to generate this data, future work must embrace multivariate descriptions of vegetation composition and dedicate sufficient time to describing variations in ground cover to explain trends in soils.

### *6.3.3 The Home Field Advantage and the Substrate-quality Matrix-quality Interaction*

Much of the motivation for this work came from the need to better understand the Home Field Advantage (HFA) effect. This effect describes an increase in decomposition rate in

translocation experiments of leaf litter at its home site than at a different location (Ayres *et al.*, 2009). Seemingly, the cause of this effect must be related to the capacity of soil communities to specialise towards the decomposition of litter with particular properties, either by filtering communities to the most competitive and effective decomposers or by local adaptation of populations to the litter of the site (Freschet *et al.*, 2012). Better understanding the mechanisms of this effect have practical implications for carbon cycling and budgeting, but could potentially improve our fundamental understanding of microbial ecology and ecological processes by providing a clear case of either rapid environmental selection, rapid local adaptation, or an interaction between these two processes. If the HFA hypothesis is correct, it might operate at a variety of taxonomic or phenotypic scales, the local decomposer community may adapt to individual trees, to populations (Jackrel *et al.*, 2019), to species or broader taxonomic groups, or to polyphyletic groups with similar litter chemistry phenotypes. I tested whether the dominance of a tree species at a site significantly impacted the decomposition rates or decomposer diversity of litter from the same genus (which were monospecific at the sites studies, aside from oak and elm) and found no effect. This may be due to the generally low dominance of the subject species at sites. Perhaps an effect would have been found if sites were less diverse in tree species. However, an alternative hypothesis to the HFA has been proposed, the Substrate-quality Matrix-quality Interaction (SMI) which suggests that the decomposer communities of sites where litter recalcitrant to decomposition dominates will adapt to decompose recalcitrant litter types, and *vice versa* with litter labile to decomposition. This might be thought of as the HFA hypothesis acting at the level of the phenotype. I found more support for this hypothesis in my work, TreeCCA2 (Chapter4, Section 4.4) - a gradient within my tree community data that separated important tree species with labile and recalcitrant litter and was unassociated with woodland age - was important in modifying decomposition rates and microbial communities for both bacteria and fungi. I found that leaf litter bag mass was lower in sites dominated by tree species with labile litter. However, as most of the decomposition occurring in the spring and summer months was from tree species that produce labile litter (true almost by definition) this might indicate an SMI effect. I did not model for an interaction effect of TreeCCA2 with the tree species of litter that may have identified different relationships for recalcitrant and labile litter types. In the specific chapter discussion, I have outlined how careful selection of sites for those with contrasting tree species dominating between sites would allow for a more rigorous test of the HFA in future work.

### *6.3.4 The effect of diversity on ecosystem functioning in the soil*

I found little evidence to suggest that diversity alone improved decomposition rates, suggesting that decomposer communities supporting decomposition had high redundancy. Instead, the highest diversity was generally found on species with more recalcitrant leaf litter. This may indicate that leaf litter of recalcitrant species provide a great variety of niches suiting different functional groups, increasing diversity. These highly diverse beech and hazel litter microbial communities often differentiated from ash and elm litter communities along CCA axes, with oak communities differentiating from both of these groups (Chapter 5, Figure 5.3 and Figure 5.6). Generally, this is consistent with the hypothesis that soil diversity improves the multifunctionality of soils (Wagg *et al.*, 2014, 2019). These tree species differences interacted with seasonal differences, such that the impact of seasonal change on community composition varied by species. A potential interaction between the tree species of litter and woodland age category was tested for in model refinement but was non-significant and not included in models, suggesting that species that support this multifunctional litter effect may be early arrivals to woodland systems during woodland creation. However, an interaction between woodland age category and season significantly influenced fungal community composition, woodlands of different age categories differed to a greater extent in the summer and the spring, suggesting that fungi that are more common in older woodlands may tend to be slow colonisers of leaf litter or are active later in the decomposition process. The overall picture of my results indicates that the relationship between diversity and functioning in soils is not simple. One of the difficulties in detecting these relationships using alpha diversity metrics is that communities that are very different in their composition may have similar alpha diversity. Only by looking at beta diversity metrics can direct relationships be drawn between the proportions of specific compartments of diversity and ecosystem functioning. However, this means using multivariate methods that are more complex to implement and may be more difficult to interpret. Soil processes are simplified in conceptual models of ecological systems in such a way that there may appear to be no relationship between diversity and functioning. However, the huge number of separate processes that must occur to facilitate soil processes requires highly multifunctional soil communities, perhaps varying slightly for each species' leaf litter and certainly varying over the course of a year.

### *6.3.5 The value of replication for future leaf litter decomposition studies*

To accomplish the second experiment in this thesis I buried and recovered over 1200 leaf litter bags in 27 locations across 23 woodlands spread over an area of over 1,000km$^2$. To enable this, I collected over 1kg dry weight of falling leaf litter and analysed the soil of 160 potential study locations spread across 29 woodlands. This is not including the multiple woodlands I visited in which a suitable studly location - on level ground, large enough to conduct the work, and safely accessible in all seasons - could not be found. This groundwork allowed me to control for variables such as pH and soil texture that are known to alter soil communities (Bach *et al.*, 2010; Griffiths *et al.*, 2011), and to conduct a study that was highly replicated within and between sites. The high degree of replication in these experiments have enabled me to draw new conclusions about the role of the tree species producing litter and woodland age in shaping the functional diversity of woodland microbes. To calculate the decomposition rate *k* a different sampling design must be adopted in which litter bags are sampled frequently in the early stages of decomposition and more sparsely as decomposition continues. I have no criticism of this method, but when time and effort are limited resources, it reduces the number of replicates possible in the study, and therefore the statistical power, compared to less frequent sampling. To elucidate the subtle effects of environmental conditions on complex datasets high statistical power is paramount. In my work I chose to capture seasonal snapshots of natural decomposition, burying leaves in the winter and uncovering one set of replicates every 13 weeks. This reduced my ability to determine decomposition rate precisely but greatly increased statistical power.

A substantial barrier to increasing statistical power of leaf litter studies is the difficulty of recovering samples. This is a logistical problem, but a serious one. Anecdotally, the proportion of litter bags lost during experiments can approach 50% in certain designs. Methodology is key here. Standardisation of methodology has the capacity for greatly increasing the statistical power of studies and preventing the wasting of many days of work. I recovered 100% of the over 1,200 leaf litter bags I buried for this work. Of the 809 bags, a single bag was damaged in such a way that litter may have been lost from it. Had I lost all three replicates of even a single species/season/site combination I would not have been able to make use of restricted permutation testing of microbial community CCA significance that required all study design variables to be completely balanced (i.e. to have an equal number of replicates in all sites, and

plots) and the statistical power of my work would have been greatly reduced. For testing for differences in decomposition alone, general linear mixed models can be applied with imperfectly balanced studies, although this is less optimal than perfectly balanced designs. No one in the scientific community benefits from the wasted hours of researchers in their field as it slows our progress as a research community to a fuller understanding of the natural world. I suggest that future work follows my method of predetermining litter bag sampling replicates, securing them to rigid plastic or metal frames with a monofilament line (I used plastic garden fencing cut to size), and mapping and photographing sites to indicate where these sampling frames are buried or placed.

Reviewing the literature of leaf litter decomposition studies, surprisingly few studies include reasonable levels of site replication in their treatment categories. This means that treatment categories may be completely confounded with vegetation or soil differences that I have identified as important drivers of decomposition rate and soil microbial diversity. The need for within-site replication has been identified in previous work (Bradford *et al.*, 2016). Ideally, my study would have included more than three replicates of each tree species' litter, at each site, in each sampling period but this level of replication at least enabled me to estimate decomposition with higher confidence. However, as the single research working on the project with occasional intermittent volunteer assistance, higher degrees of replication were not possible. Larger, better funded studies may achieve greater success in this area. Replicates of different tree species' litter also contribute to the confidence of estimates of random site effects, while also describing decomposition capacity more broadly than is possible with a single species' litter. Given the notable differences in the composition of microbial communities on decomposing leaves, studies that only examine the decomposition of a single species are only examining a subset of the decomposer community at a site. Previous work has sometimes included litter from multiple species, including one study of beech, sedge, and milkvetch litter decomposing in either a grassland or beech woodland site (Buresova *et al.*, 2019). In Buresova *et al.*'s study, beech and sedge litter communities significantly diverged with seasonal interaction that increased differences in the summer months, similar to my results. However, they did not detect differences between milkvetch communities and those of other litter types. It is unclear if increased site replication, particularly in woodlands not dominated by beech litter, or increased sequencing depth would have altered these results. I

am aware of no studies that examine microbial diversity of decomposing litter from multiple common, broadleaf woodland tree species. Given the advantages of using multiple species in detecting multifunctionality of woodland soils, I would encourage others to adopt this approach.

To summarise, I would make the following recommendations for future litter experiments:

1. Survey potential sites for factors that have been identified as important drivers of the process you are trying to study.
2. Pair or group sites such that a site replicate of each treatment category can be compared against similar sites of other treatment categories.
3. Aim to achieve a high degree of site replication while working within these limitations.
4. Consider using multiple plant litter types that commonly occur in the area local to your sites.
5. Aim to achieve the highest degree of within-site replication possible. Adjust the mass of litter included in leaf litter bags, if necessary, to maximise the number of replicates in the study.
6. Adjust the sampling frequency of your work to balance the precision with which you can determine the initial decomposition rate (*k*) and the statistical power of a study.
7. Maximise litter bag recovery by securing litter bags to a larger, easily locatable object, such as a sampling frame. This saves time later as individual sets of replicates can be recovered simultaneously by recovering the sampling frame without extensive searching.
8. Either during the study or during the initial site selection, collect tree and ground cover vegetation data and quantify differences between sites using multivariate approaches such that the importance of these variables can be accounted for.
9. If leaf litter bag mesh size does not exclude soil mesofauna, quantify the differences in the diversity and abundance of these decomposers between sites such that the importance of these decomposers can be accounted for.
10. Employ suitable statistical approaches that can account for non-independence of data without unnecessary merging or discarding data from replicates.

### *6.3.6 Critical evaluation of next-generation sequencing assessment of the soil microbiome*

When designing this project, concerns were raised that the degree of noise in soil microbial systems would mask subtle differences in soil diversity between land use categories. Legitimate concerns were also raised around the sophistication of molecular methods and their capacity to answer complex ecological questions. The argument supporting the approach has been that molecular approaches, although imperfect, are still capable of answering certain questions in ecology that cannot be feasibly broached with other existing methodological approaches (Baird & Hajibabaei, 2012; Pompanon & Samadi, 2015). Weaknesses of molecular ecology that make it less suited to answer certain questions that traditional methods can handle easily, often due to current limitations in relating read abundance to absolute population size (Clare, 2014; Barnes & Turner, 2015). Further discussion of this issue can be found in the introduction chapter. I suggest that weaknesses in molecular biology have analogues in widely adopted, classical ecological methods. These include parallels with common challenges to ecological research: detectability, cryptic species, incomplete taxonomies, challenging taxonomic keys, sampling effort, sampling scales, temporal sampling windows, and issues of improper study design. Working within the limitations of molecular methods, I have demonstrated that they can disentangle complex scientific problems and detecting subtle environmental drivers of diversity in soil microbial systems. However, some challenges involved in the approach are clear and I will comment on these below.

In the first chapter in which I sequence soil cores for bacteria and fungi from 210 soil cores, equally divided amongst 21 sites. Soil cores were sampled from a $20m^2$ area in the centre of the site. The degree of difference between individual core richness and aggregated richness across all cores provided an indication of community homogeneity across this $20m^2$ area. Aggregated site bacterial richness was approximately three times higher than individual core richness, but samples within sites were clustered in multivariate representations of the community data. Soil core samples analysed to detect fungi were far more variable within sites. This is not surprising given the high degree of spatial variability in the soil microbiome identified in previous work (Jiao *et al.*, 2018; Chen *et al.*, 2019; Zhang *et al.*, 2020). Researchers aiming to characterise the soil microbiome may have to sample within sites more intensively to describe the range of fungal taxa present at a site than the bacterial taxa. And my work suggests that characterising either bacteria or fungi will require multiple soil samples per site.

Studies have made clear that soil scientists must think very carefully about how they sample vertical and horizontal space in these systems (Jiao *et al.*, 2018; Chen *et al.*, 2019), not to mention the temporal variations in soils (López-Mondéjar *et al.*, 2015). The existence of clear temporal variation in studies of soils indicates that concerns around the persistence of inactive DNA in soils may be overstated. Although the DNA of larger organisms, deposited in large amounts may be detectable in soils for a long time (Barnes and Turner 2015), it is not clear how long microbial DNA persists as it is present in less recalcitrant tissue and in lower quantities. This is not to say that the DNA of dead or dormant organisms will not be detected in metabarcoding studies of eDNA, but they appear to be a subordinate signal to the seasonal variations of active taxa.

Rarefaction curves not presented in this thesis (McMurdie & Holmes, 2014) indicate that samples were comfortably within the read depth needed to capture soil core diversity. From this work I would suggest that next-generation sequencing approaches that generate tens of thousands of reads per sample are required to adequately measure differences in soil bacterial and fungal diversity and composition. Previous sequencing platforms that produce only tens of thousands of reads in total may lack the read depth required to describe the bacterial or fungal diversity of a highly replicated experiment. Soils also appear to be more diverse than leaf litter lags and so these two kinds of study may vary in their read depth requirements. Insufficient read depth is known to influence measures of microbial diversity (Singer *et al.*, 2019). Arguments exist for including multiple PCR replicates of individual samples in order to capture soil diversity (Dopheide *et al.*, 2018). However, that does not appear to have been necessary for identifying ecological trends in this study. High soil heterogeneity and diversity increase the importance of ensuring equal sample concentrations are pooled to libraries and provided to the sequencer. Methodologically, the approach adopted in this work for normalising concentrations within libraries using Solid Phase Reversible Immobilization (SPRI) beads has been highly successful, both when conducted by hand pipetting, or when using an OT-2 Liquid Handling robot (Opentrons Labworks). If researchers find themselves with the capacity to replicate sample PCRs and sequence them, perhaps they would be better served by increasing the number of loci included in the study instead. Suggesting a specific sequencing platform, read depth, or a number of replicates is impossible to do without knowledge of the specific research question and numbers are certain to change rapidly with

technological and computational advances. However, to capture variation is soil diversity, especially fungal diversity studies, should aim to sample soils with high replication and produce at least tens of thousands of reads per sample.

When analysing soil diversity data, different approaches may need to be taken with different loci, especially between 16S and ITS1. These regions vary greatly in their size and variability which may result in different requirements for bioinformatics and inferring phylogeny following short read length sequencing. I have described slight differences in the bioinformatic approach used to analyse 16S and ITS1 sequences in my work. Adapting the methods to fit the requirements of the data is encouraged in future work. If possible, research should move away from descriptions of soil community composition in simplified two-dimensional space such as NMDS plots, as the sheer number of microbial taxa means that additional ecological trends are likely to be important for different groups. Caution should still be taken in the interpretation of bioinformatic results as the fundamental biology of bacterial and fungal phylogenetics is complex and evolving. Software such as *DADA2* are capable of assigning taxonomies to bacterial and fungal sequences but our confidence in these assignments is highly variable. We also do not know how to precisely relate read abundance to taxonomic frequency or biomass with high confidence (Clare, 2014). However, comparing the relative abundance of taxa between samples appears to be a fruitful avenue of research, provided that there are reasonable levels of replication and samples are treated identically. In this work, I have avoided making too many specific claims about specific taxonomic groups and their relationships to variables because individual claims are of lower certainty that aggregated trends. At this stage of soil and litter eDNA studies, identifying ecological trends may be the most productive form of analysis, hopeful specific details will become clearer as these methods mature.

## 6.4 Implications for Practitioners in British Woodlands

### *6.4.1 Clear reasons to conserve old woods and create new ones*

Practitioners are perhaps most concerned with whether the microbiome of woodland creation sites rapidly accommodates microbial species typical of established woodlands; and whether ASNW contains a unique microbiome that merits conservation, perhaps more so than that of

established mature woodland. The results of my work suggest that the answer to both questions is generally, yes. The soil communities of young woodland sites (50-60 years old) had significantly differentiated from those of arable agricultural soils and did not significantly differ from other woodland types. However, although I detected no significant differences in the decomposition rate of woodlands of different age, the decomposition measured in woodland creation sites (less than 50 years old) was generally lower than that of established mature secondary woodland, and notably lower on average than ASNWs after accounting for other influences. Beyond seasonal and leaf litter differences, bacterial and fungal community composition was most strongly influenced by woodland age, either directly or indirectly through ground vegetation changes. These young woodlands had significantly different community compositions than mature secondary woodland and ASNW. Furthermore, ASNW differed from mature woodland and woodland creation sites for certain aspects of fungal community composition. Although these differences explained a smaller amount of the variation than the seasonal and the litter's species, they were significant, and they were the most important site-specific explanatory variable.

Generally, the ground flora or ground cover most associated with ancient woodland in my study area explained variation in microbial decomposer communities better than simply categorising woodlands into broad age classes. Mature woodlands with ground cover more similar to a woodland creation site, e.g. with grasses, thistle, or nettle; had soil communities more similar to younger woodlands. Equally, woodland creation sites where the ground cover had taken on aspects of mature woodlands, most typically by increasing cover of ferns, litter, and ground ivy, were more similar to mature woodland and ASNW in their soil community. Microbial decomposer communities also differed between sites along a ground cover gradient, with litter and fern dominated woodlands at one end, and herbaceous cover at the other. This cautiously suggests that ground flora indicators including ancient woodland indicator species are associated with certain soil microbiomes communities. This should be encouraging to those in the UK forestry sector who already incorporate indicator species into their assessments of woodlands.

### *6.4.2 Planting for soil diversity*

Multiple vegetation gradients were significantly associated with soil microbial decomposer communities. These described differences in ground cover between young and old woodlands, but also differences irrespective of woodland age to do with herbaceous or litter ground cover, and between sites dominated by tree species with recalcitrant leaf litter or labile leaf litter. These variables also significantly impacted the decomposition rates of leaf litter within sites, which was higher when sites were dominated by tree species with easy to decompose leaf litter or sites with high herbaceous, mossy ground cover. These are correlative relationships, not necessarily implying a causal relationship, but they are significant. These variables are products of the planting and management decisions taken at these woodlands. Some of these decisions may have been recent, others made by previous generations.

Planting and management choices matter to soil and litter microbial diversity. I aimed to identify if management impacts diversity, but it is beyond the scope of this work, and probably not appropriate, to suggest which combination of conditions is "best" for conserving woodland diversity. Such suggestions would require further enquiry and the input of land managers. The advice that I can give from these results is that a diversity of approaches to habitat planting and management will likely produce a diversity of soil and litter microbial communities. Across a landscape, those seeking to conserve biodiversity at these scales might achieve best results by ensuring landscape habitat heterogeneity. The gradients identified as significant in this work suggests that diverse habitats might be created by a mixture of young, mature and ancient woodlands, some open, some closed, stocked with a selection of trees that vary in their composition between sites.

### *6.4.3 Translocation projects - questions raised*

Soil translocation projects have been proposed as mitigation methods for damage or grubbing out of ancient woodland. This approach is not new (Helliwell *et al.*, 1996), but at the time of writing the issue has become prominent in respect to large scale infrastructure projects, specifically the HS2 rail line (HS2 Ltd, 2017, 2020). Currently the HS2 project expects to translocate up to 9.6 ha of ancient woodland soils to alternative sites, which they propose to undertake in the autumn months (HS2 Ltd, 2020). The basis of these translocations is the

recognition that woodland specialist plant species are slow dispersers that should be protected and conserved, and that the soils of ancient woodlands differ in their properties and probably their diversity due to their land-use history. However, studies that compare the diversity of ancient woodland soils to those of younger woodlands are not common, even for larger soil fauna (Ashwood *et al.*, 2019). Analysis of soil biodiversity that can touch upon the 95% of soil bacteria that are not culturable with traditional techniques have only recently become available to researchers (Handelsman, 2004). The capacity for these methods to identify differences in taxonomic diversity has been highly variable, and generally only possible to a low resolution. Molecular sequencing-based approaches, bolstered by next-generation technology and software has greatly increased our capacity to generate high-resolution taxonomic inventories of woodland microbes (Baird & Hajibabaei, 2012; Pawlowski *et al.*, 2014). As such the evidence base upon which translocation projects are based upon is currently small. However, the theoretical framework is sensible. Key questions required further investigation: across a typical UK landscape are ASNW soil microbial communities substantially different from woodland creation or mature woodland sites? When should translocation experiments be timed to maximise microbial diversity and abundance? What factors might influence the success of translocation projects? My research indicates that there are aspects of the soil microbial community that significantly prefer ASNW to other woodland types, much broader aspects of microbial diversity prefer either mature or ASNW sites over woodland creation sites. This indicates that the intentions behind translocation projects are merited. However, multiple studies have found microbial soil communities to vary in diversity throughout the year, differing in the autumn and winter from that of the spring and summer (Davey *et al.*, 2012; López-Mondéjar *et al.*, 2015), being more diverse in the summer than the winter (Zhang *et al.*, 2020). I found higher bacterial decomposer diversity on litter in the summer months and seasonal differences in decomposer fungal and bacterial compositions in all species. It is unclear how the timing of relocation projects will affect their success, but there are clear grounds for further investigation. Finally, the factors that appear to influence the microbial composition of woodlands include soil properties, tree cover, understory vegetation, and land use history. If translocation projects are to be a success soil microbes associated with ASNW need to be not only introduced to a new woodland, but surrounding areas of the woodland likely need to be managed to more closely replicate the properties of the donor woodland. If site owners are preparing for autumnal translocations, they should bear in mind that soil microbes are not just for Christmas. A lot of work will need to go into

ensuring that the translocated soil fauna establish and propagate through the rest of the woodland in the years following the translocation event, rather than the soil fauna of the translocated soil being swamped by the dominant fauna of the receptor site.

### *6.4.4 Engaging the public with soil and woodland ecology*

Though it is not reflected in my thesis, I have devoted considerable time to engaging with members of the public and the membership of The Woodland Trust, my funders, about woodland science as part of my PhD work. This has included talks to the Natural History Society of Northumbria, The British Ecological Society, student researchers at The Woodland Trust, students at Newcastle University, to readers of the Newcastle university science magazine *{REACT}*, to academics at various conferences, and with talks that I had hoped to give in 2020 to attendants of the *art.Earth* conference *Borrowed Time* and of the *treescapes2020* conference. I have given virtual classes on woodland systems to primary school children in the USA during the coronavirus pandemic. I hope to be able to continue to assist and inform The Woodland Trust in the future as this work is published and as further reports for The Woodland Trust are prepared. Communicating the scientific facts regarding woodlands to as many people as possible and as diverse communities of people as possible is essential to equipping decision makers with the information they need to decide the future of these landscapes and is part of my development as a researcher. The facts I have communicated relating to decomposition rates, or to aspects of microbial diversity can only inform readers. Judgements regarding what to do with this land, how to manage it, and what it should be managed for are beyond the scope of this work and must be made by the people who live in and around these landscapes. Woodland research is increasingly identifying multiple public goods that might be produced by wooded landscapes relating to health (Maas *et al.*, 2009), carbon sequestration (Cannell, 1999), and ecosystem services (Burton *et al.*, 2018). However, that does not mean that there is sufficient financial support, woodland workers, or public enthusiasm necessary to put the suggestions into action. In many cases public goods are mutually exclusive. When decisions must be made, those with the capacity to enact changes should put the question to an informed public to decide which public good to attempt to achieve (Oxman *et al.*, 2010). Otherwise, decision makers must bear the entire responsibility for the successes and failure of their work. This process requires a sufficiently informed public. Scientific researchers have a responsibility to be one of the knowledge

brokers working to achieve this (Tabbush, 2004). Researchers who disagree may find that all decisions regarding the implementation of their work are conducted by policy makers and commercial foresters. This is no bad thing, but those researchers may find themselves targeting their work to these groups if they wish to make an impact.

My suggestions to those seeking to communicate the beauty and richness of these landscapes is to not overly simplify their complexity and to acknowledge our knowledge limitations. Our current understanding of the ecology of soils, particularly its microbiota, is lacking in many areas due to the complex nature and structure of the ecosystem (Mocali & Benedetti, 2010; Fierer, 2017). I suggest that embracing this complexity and our primitive understanding is the honest approach to communicating the wonder of soils. I believe that anthropomorphising soils; suggesting that compartments of it talk or facilitate "chatter" between trees; or overstating simplifications of the system such as "the wood-wide-web" will ultimately lead to negative consequences when public opinion meets policy. This is something that applies equally to researchers as to other knowledge brokers. The real recalcitrance of soils to interrogation and characterisation communicates the challenges soil scientists face to the public and may inspire some to become soil scientists themselves. Communication issues are also a challenge in forestry. Woodlands in the UK are not primary forests. Much of the difficulty in planning for a sustainable future is a resistance to change an environment that is perceived as fixed and natural (Hubbard, 1993). It is simply the case that no affordable management regime will preserve the British countryside in stasis or revert it to any form that might be found there historically. There is no standing still or returning to the past in UK landscape management. New woodlands are set to be created at an increasing rate within the UK over the coming decades (DEFRA, 2018). Change is coming and an informed public has a place in deciding where public money is spent (Tabbush, 2004).

Science that concerns the landscapes of the British Isles has the capacity to enthuse, surprise, challenge, and amaze its residents. In my work I have found that complex approaches and results are perfectly communicable to all audiences, including children. These audiences, that mostly lack a financial stake in these environments, do not need to be tailored to. They tend to have an appetite for understanding science that allows researchers to express their

positions unreservedly. As such they are probably the audience that researchers can achieve the greatest impact by reaching. If researchers are unconcerned about the implications and impact of their work, then engaging with the public is unnecessary. However, if researchers working in the UK discover that they grow to care about the landscapes they study, I suspect they will find the most like-minded audience for their work to be the British public.

### 6.5 References


Ashwood, F., Watts, K., Park, K., Fuentes-Montemayor, E., Benham, S. & Vanguelova, E.I. (2019) Woodland restoration on agricultural land: long-term impacts on soil quality. *Restoration Ecology*. 42182.

Ayres, E., Steltzer, H., Simmons, B.L., Simpson, R.T., Steinweg, J.M., Wallenstein, M.D., Mellor, N., Parton, W.J., Moore, J.C. & Wall, D.H. (2009) Home-field advantage accelerates leaf litter decomposition in forests. *Soil biology & biochemistry*. 41 (3), 606–610.

Bach, E.M., Baer, S.G., Meyer, C.K. & Six, J. (2010) Soil texture affects soil microbial and structural recovery during grassland restoration. *Soil biology & biochemistry*. 42 (12), 2182–2191.

Bahram, M., Peay, K.G. & Tedersoo, L. (2015) Local-scale biogeography and spatiotemporal variability in communities of mycorrhizal fungi. *The New phytologist*. 205 (4), 1454–1463.

Baird, D.J. & Hajibabaei, M. (2012) Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular ecology*. 21 (8), 2039–2044.

Barberán, A., Casamayor, E.O. & Fierer, N. (2014) The microbial contribution to macroecology. *Frontiers in microbiology*. 5203.

Barnes, M.A. & Turner, C.R. (2015) The ecology of environmental DNA and implications for conservation genetics. *Conservation genetics* . 17 (1), 1–17.

Bradford, M.A., Berg, B., Maynard, D.S., Wieder, W.R. & Wood, S.A. (2016) Understanding

the dominant controls on litter decomposition. *The Journal of ecology*. 104 (1), 229–238.

Buresova, A., Kopecky, J., Hrdinkova, V., Kamenik, Z., Omelka, M. & Sagova-Mareckova, M. (2019) Succession of Microbial Decomposers Is Determined by Litter Type, but Site Conditions Drive Decomposition Rates. *Applied and environmental microbiology*. 85 (24), .

Burton, V., Moseley, D., Brown, C., Metzger, M.J. & Bellamy, P. (2018) Reviewing the evidence base for the effects of woodland expansion on biodiversity and ecosystem services in the United Kingdom. *Forest ecology and management*. 430366–379.

Cahalan, C., Hester, A., Humphrey, J., Kirby, K., Moffat, A. & Valatin, G. (2011) 'Woodlands', in National Ecosystem Assessment (ed.) *The UK National Ecosystem Assessment Technical Report*. [Online]. Cambridge: UNEP-WCMC. pp. 241–294.

Cannell, M.G.R. (1999) Growing trees to sequester carbon in the UK: answers to some common questions. *Forestry*. 72 (3), 237–247.

Chen, W., Jiao, S., Li, Q. & Du, N. (2019) Dispersal limitation relative to environmental filtering governs the vertical small-scale assembly of soil microbiomes during restoration Gaowen Yang (ed.). *The Journal of applied ecology*. 355.

Clare, E.L. (2014) Molecular detection of trophic interactions: emerging trends, distinct advantages, significant considerations and conservation applications. *Evolutionary applications*. 7 (9), 1144–1157.

Davey, M.L., Heegaard, E., Halvorsen, R., Ohlson, M. & Kauserud, H. (2012) Seasonal trends in the biomass and structure of bryophyte-associated fungal communities explored by 454 pyrosequencing. *The New phytologist*. 195 (4), 844–856.

DEFRA (2018) *A Green Future: Our 25 Year Plan to Improve the Environment*.

Dopheide, A., Xie, D., Buckley, T.R., Drummond, A.J. & Newcomb, R.D. (2018) Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity Michael Bunce (ed.). *Methods in ecology and evolution / British Ecological Society*. 718966.

Dumbrell, A.J., Nelson, M., Helgason, T., Dytham, C. & Fitter, A.H. (2010) Relative roles of niche and neutral processes in structuring a soil microbial community. *The ISME journal*.

4 (3), 337–345.

Feyerabend, P.K. (1978) *Against method: Outline of an anarchist theory of knowledge*. NLB.

Fierer, N. (2017) Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature reviews. Microbiology*. 15 (10), 579–590.

Forest Research (2019) *Public Opinion of Forestry 2019 UK and England*.

Forest Service (2006) *NI Forestry - a strategy for sustainable growth*.

Freschet, G.T., Aerts, R. & Cornelissen, J.H.C. (2012) Multiple mechanisms for trait effects on litter decomposition: moving beyond home-field advantage with a new hypothesis: Substrate-matrix quality interactions in decay. *The Journal of ecology*. 100 (3), 619–630.

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M. & Whiteley, A.S. (2011) The bacterial biogeography of British soils. *Environmental microbiology*. 13 (6), 1642–1654.

Handelsman, J. (2004) Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and molecular biology reviews: MMBR*. 68 (4), 669–685.

Helliwell, D.R., Buckley, G.P., Fordham, S.J. & Paul, T.A. (1996) Vegetation succession on a relocated ancient woodland soil. *Forestry* . 69 (1), 57–74.

Hermy, M., Honnay, O., Firbank, L., Grashof-Bokdam, C. & Lawesson, J.E. (1999) An ecological comparison between ancient and other forest plant species of Europe, and the implications for forest conservation. *Biological conservation*. 91 (1), 9–22.

Holt, R.D. (1996) Adaptive Evolution in Source-Sink Environments: Direct and Indirect Effects of Density-Dependence on Niche Evolution. *Oikos* . 75 (2), 182–192.

HS2 Ltd (2020) *High Speed Two Phase 2a: Ancient Woodland Strategy*.

HS2 Ltd (2017) *High Speed Two Phase One: London-West Midlands Ancient Woodland Strategy*.

Hubbard, P. (1993) The Value of Conservation: A Critical Review of Behavioural Research. *The Town planning review*. 64 (4), 359–373.

Humphrey, J.W., Watts, K., Fuentes-Montemayor, E., Macgregor, N.A., Peace, A.J. & Park,

K.J. (2015) What can studies of woodland fragmentation and creation tell us about ecological networks? A literature review and synthesis. *Landscape ecology*. 30 (1), 21–50.

Jackrel, S.L., Gilbert, J.A. & Wootton, J.T. (2019) The Origin, Succession, and Predicted Metabolism of Bacterial Communities Associated with Leaf Decomposition. *mBio*. 10 (5), .

Jiao, S., Chen, W., Wang, J., Du, N., Li, Q. & Wei, G. (2018) Soil microbiomes with distinct assemblies through vertical soil profiles drive the cycling of multiple nutrients in reforested ecosystems. *Microbiome*. 6 (1), 146.

Kozłowski, J. (1993) Measuring fitness in life-history studies. *Trends in ecology & evolution*. 8 (3), 84–85.

Kuhn, T.S. (1970) *The structure of scientific revolutions*.

López-Mondéjar, R., Voříšková, J., Větrovský, T. & Baldrian, P. (2015) The bacterial community inhabiting temperate deciduous forests is vertically stratified and undergoes seasonal dynamics. *Soil biology & biochemistry*. 8743–50.

Maas, J., Verheij, R.A., de Vries, S., Spreeuwenberg, P., Schellevis, F.G. & Groenewegen, P.P. (2009) Morbidity is related to a green living environment. *Journal of epidemiology and community health*. 63 (12), 967–973.

Mace, G.M. (2014) Whose conservation? *Science*. 345 (6204), 1558–1560.

Maller, C., Townsend, M., St Leger, L., Henderson-Wilson, C., Pryor, A., Prosser, L. & Moore, M. (2009) Healthy Parks, Healthy People: The Health Benefits of Contact with Nature in a Park Context. *The George Wright forum*. 26 (2), 51–83.

McMurdie, P.J. & Holmes, S. (2014) Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS computational biology*. 10 (4), e1003531.

Milligan, C. & Bingley, A. (2007) Restorative places or scary spaces? The impact of woodland on the mental well-being of young adults. *Health & place*. 13 (4), 799–811.

Mocali, S. & Benedetti, A. (2010) Exploring research frontiers in microbiology: the challenge

of metagenomics in soil microbiology. *Research in microbiology*. 161 (6), 497–505.

Monbiot, G. (2014) *Feral: Rewilding the Land, the Sea, and Human Life*. University of Chicago Press.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., Knelman, J.E., Darcy, J.L., Lynch, R.C., Wickey, P. & Ferrenberg, S. (2013) Patterns and processes of microbial community assembly. *Microbiology and molecular biology reviews: MMBR*. 77 (3), 342–356.

Oxman, A.D., Bjørndal, A., Becerra-Posada, F., Gibson, M., Block, M.A.G., Haines, A., Hamid, M., Odom, C.H., Lei, H., Levin, B., Lipsey, M.W., Littell, J.H., Mshinda, H., Ongolo-Zogo, P., Pang, T., Sewankambo, N., Songane, F., Soydan, H., Torgerson, C., *et al.* (2010) A framework for mandatory impact evaluation to ensure well informed public policy decisions. *The Lancet*. 375 (9712), 427–431.

Pawlowski, J., Esling, P., Lejzerowicz, F., Cedhagen, T. & Wilding, T.A. (2014) Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. *Molecular ecology resources*. 14 (6), 1129–1140.

Pompanon, F. & Samadi, S. (2015) Next generation sequencing for characterizing biodiversity: promises and challenges. *Genetica*. 143 (2), 133–138.

Rackham, O. (2008) Ancient woodlands: modern threats. *The New phytologist*. 180 (3), 571–586.

Rackham, O. (1986) *The History of the Countryside*. London: J.M. Dent.

Rackham, O. (2012) *Woodlands*. London: William Collins.

Rapacciuolo, G. & Blois, J.L. (2019) Understanding ecological change across large spatial, temporal and taxonomic scales: integrating data and methods in light of theory. *Ecography*. 71.

Shoenberger, A. (2015) Magna Carta, the Charter of the Forest, and the Origin of the Jury System Case Notes and Commentary. *Nottingham Law Journal*. 24156–162.

Singer, G.A.C., Fahner, N.A., Barnes, J.G., McCarthy, A. & Hajibabaei, M. (2019) Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. *Scientific reports*. 9 (1), 5991.

Tabbush, P. (2004) Public money for public good? Public participation in forest planning. *Forestry* . 77 (2), 145–156.

The Scottish Government (2019) *Scotland's Forestry Strategy 2019–2029*.

Trudgill, S.T. (1988) *Soil and Vegetation Systems*. Oxford: Oxford University Press.

Uroz, S., Buée, M., Deveau, A., Mieszkin, S. & Martin, F. (2016) Ecology of the forest microbiome: Highlights of temperate and boreal ecosystems. *Soil biology & biochemistry*. 103471–488.

Wagg, C., Bender, S.F., Widmer, F. & van der Heijden, M.G.A. (2014) Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (14), 5266–5270.

Wagg, C., Schlaeppi, K., Banerjee, S., Kuramae, E.E. & van der Heijden, M.G.A. (2019) Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. *Nature communications*. 10 (1), 1–10.

Welsh Government (2018) *Woodlands for Wales*.

Wennekes, P.L., Rosindell, J. & Etienne, R.S. (2012) The neutral-niche debate: a philosophical perspective. *Acta biotheoretica*. 60 (3), 257–271.

Zhang, K., Delgado-Baquerizo, M., Zhu, Y.-G. & Chu, H. (2020) Space Is More Important than Season when Shaping Soil Microbial Communities at a Large Spatial Scale. *mSystems*. 5 (3), .

# Supplementary Data

## Chapter 2: Supplementary Data

***S2.a. file: ByrneData.csv***

***S2.b. ByrneCh.2Script.R***

```
#########################
####                 ####
#### Byrne Protocol  ####
####                 ####
#########################


#### Info ####
#R version 3.6.0 (2019-04-26)


#### Packages ####
library(ggplot2)
library(ggfortify)
library(grid)
library(gridExtra)
library(gridBase)
library(ggpubr)
library(ggsci)
library(patchwork)


#### Read in soil characteristic data ####


soil.dat <- read.csv("ByrneData.csv")
soil.dat$Site_Age = factor(soil.dat$Site_Age,
                           levels(soil.dat$Site_Age)[c(1,4,5,3,2)])


#soil.dat$NO3_mg_per_kg[141] <- NA


Site_Means<- aggregate(soil.dat, by = list(soil.dat$Site), FUN =
mean, na.rm=T)
Site_Means2<- aggregate(soil.dat, by = list(soil.dat$Site), FUN =
head, 1)
Site_Means[5:11] <- Site_Means2[5:11]
rownames(Site_Means)<- Site_Means$Site
rm(Site_Means2)


coordinates(Site_Means) <- ~Easting+Northing
proj4string(Site_Means) <- CRS("+proj=tmerc +lat_0=49 +lon_0=-2
+k=0.9996012717 +x_0=400000 +y_0=-100000 +ellps=airy +datum=OSGB36
+units=m +no_defs")
```

```
##### Site Eigen values ####


MEMEig <- dbmem(cbind(Site_Means$Easting, Site_Means$Northing))


forward.sel(Site_Means$pH_H2O, MEMEig)          #1 Significant
#forward.sel(Site_Means$Total_C_PerCent, MEMEig) #1 Significant
forward.sel(Site_Means$TOC_PerCent, MEMEig)     #1 Significant
#forward.sel(Site_Means$TIC_PerCent, MEMEig)     #1 Significant
forward.sel(Site_Means$Total_N_PerCent, MEMEig) #1 Significant
forward.sel(Site_Means$NH4_mg_per_kg, MEMEig)   #None Significant
forward.sel(Site_Means$NO3_mg_per_kg, MEMEig)   ##1 Significant or
NOT DEPENDS on inclusion of data point
forward.sel(Site_Means$TotalCtoN, MEMEig)            #None
Significant
#forward.sel(Site_Means$OrganicCtoN, MEMEig)            #None
Significant
forward.sel(Site_Means$Bulk_Dens_g_cm3, MEMEig) #None Significant


Site_Means$SpEig <- MEMEig[,1]


for (i in levels(soil.dat$Site)){
  soil.dat$SpEig[soil.dat$Site == i] <-
Site_Means$SpEig[Site_Means$Site == i]
}


reponseVariables<- soil.dat[,c("pH_H2O", "Total_N_PerCent",
"TOC_PerCent",
                      "TotalCtoN", "NH4_mg_per_kg", "NO3_mg_per_kg"
)]


#### PCA for whole dataset ####
cor(reponseVariables)        #Some values are strongly correlated


diag(cov(reponseVariables))  # However the variances of the data
varly a lot, but we
# can scale the data in the PCA


PRCOMP1 <- prcomp(reponseVariables[-141,], scale = T)


summary(PRCOMP1)


# The explanatory power of the PCs begins to decline steeply after
PC5,
```

```
# however the eighenvalues of PC4 and PC5 are both below 1.


plot(PRCOMP1, type = "lines")
abline(h = 1, col = "red", lty = 4)


rownames(PRCOMP1$rotation)<- c("pH", "Total N", "Organic Carbon",
"C:N", "NH4", "NO3")


PC1_2 <- autoplot(PRCOMP1, data = soil.dat[-144,], colour =
'Site_Age',# shape = F,
                  loadings = TRUE, loadings.colour = 'black',
                  loadings.label = TRUE, loadings.label.size = 4,
                  loadings.label.colour = "black",
                  loadings.label.repel = FALSE,
                  loadings.label.fontface = "bold")+
  theme_classic() +
  theme(legend.position = "none") +
  scale_color_brewer(palette = "Set1")


PC2_3 <- autoplot(PRCOMP1, x=2, y=3, data = soil.dat[-144,], colour
= 'Site_Age',# shape = F,
                  loadings = TRUE, loadings.colour = 'black',
                  loadings.label = TRUE, loadings.label.size = 4,
                  loadings.label.colour = "black",
                  loadings.label.repel = FALSE,
                  loadings.label.fontface = "bold") +
  theme_classic() +
  theme(legend.position = "none") +
  scale_color_brewer(palette = "Set1")


PC1_3 <- autoplot(PRCOMP1, x=3, y=1, data = soil.dat[-144,], colour
= 'Site_Age',# shape = F,
                  loadings = TRUE, loadings.colour = 'black',
                  loadings.label = TRUE, loadings.label.size = 4,
                  loadings.label.colour = "black",
                  loadings.label.repel = FALSE,
                  loadings.label.fontface = "bold") +
  theme_classic() +
  theme(legend.position = "none") +
  scale_color_brewer(palette = "Set1")


PCLegend <-get_legend(autoplot(PRCOMP1, x=3, y=1, data = soil.dat[-
144,], colour = 'Site_Age',# shape = F,
                               loadings = TRUE, loadings.colour =
'blue',
                               loadings.label = TRUE,
loadings.label.size = 3,
                               loadings.label.colour =
"black")+theme_classic()+
```

```
                        guides(col = guide_legend(title = "Land
use",
                                                  title.position =
"top"))+
                        scale_color_brewer(palette = "Set1")
)



#svg(file = "PCA_WrEN.svg", width = 8, height = 7)
grid.arrange(PC1_2,PCLegend, PC2_3, PC1_3, nrow=2, ncol=2)
#dev.off()


#### Model Plotting objects ####


plotFont <- "sans"


#### pH ####


pH_model <- lme(pH_H2O ~ Site_Age+SpEig, random = ~ 1|Site,
                  data = soil.dat, method = "ML",)


summary(pH_model)


plot(pH_model)
plot(pH_model,pH_H2O~fitted(.))
qqnorm(pH_model,~ resid(.)|Site)


anova(pH_model)   # Non-significnat
glhtSummary <- glht(pH_model, linfct=mcp(Site_Age="Tukey"))
summary(glhtSummary) # No significant - matches 95% confidences
intervals


pdat <- expand.grid(Site_Age=levels(soil.dat$Site_Age),
                    SpEig=0)


pred <- predict(pH_model, newdata = pdat,
                level = 0, se.fit = TRUE,
                na.action = "na.exclude", type = "response")


predframe <- data.frame(pdat, pred = pred)
```

```
pH_Plot <- qplot(x = Site_Age ,
                 y = pred.fit,
                 data = predframe,
                 xlab = "", ylab = "pH") +
  geom_errorbar(aes(
    ymin  = (pred.fit-(pred.se.fit*1.96)),
    ymax  = (pred.fit+(pred.se.fit*1.96)),
    width = 0.15))+theme_classic()+
  theme(axis.text.x=element_blank())#,axis.text.x=element_text(angle
=40, hjust=1))


pH_Plot


#### Total C ####


T_C_model <- lme(Total_C_PerCent ~ Site_Age+SpEig, random = ~
1|Site,
                   data = soil.dat, method = "ML")
plot(T_C_model)
summary(T_C_model)


r.squaredGLMM(T_C_model)


plot(T_C_model)
plot(T_C_model,Total_C_PerCent~fitted(.))
qqnorm(T_C_model,~ resid(.)|Site)


anova(T_C_model)    # Significant
glhtSummary <- glht(T_C_model, linfct=mcp(Site_Age="Tukey"))
summary(glhtSummary) # Matches 95% confidence intervals


pdat <- expand.grid(Site_Age=levels(soil.dat$Site_Age),
                    SpEig=0)


pred <- predict(T_C_model, newdata = pdat,
                level = 0, se.fit = TRUE,
                na.action = "na.exclude", type = "response")


predframe <- data.frame(pdat, pred = pred)


T_C_plot <- qplot(x = Site_Age,
```

```
                  y = pred.fit,
                  data = predframe,
                  xlab = "", ylab = "Total C %") +
  geom_errorbar(aes(
    ymin  = (pred.fit-(pred.se.fit*1.96)),
    ymax  = (pred.fit+(pred.se.fit*1.96)),
    width = 0.15))+theme_classic() +
  geom_text(aes(label=c("a", "ab", "b", "ab","b"), family =
plotFont,  y=6.6),
            colour="black", vjust=0, size=4,
position=position_dodge(0))+
  coord_cartesian(ylim = c(0, 9))


T_C_plot



#### TOC ####


TOC_model <- lme(TOC_PerCent ~ Site_Age+SpEig, random = ~ 1|Site,
                   data = soil.dat, method = "ML")
plot(TOC_model)
summary(TOC_model)


r.squaredGLMM(TOC_model)


anova(TOC_model) #Signifcant
glhtSummary <- glht(TOC_model, linfct=mcp(Site_Age="Tukey"))
summary(glhtSummary) # Matches 95% confidence intervals



plot(TOC_model)
plot(TOC_model,TOC_PerCent~fitted(.))
qqnorm(TOC_model,~ resid(.)|Site)


pdat <- expand.grid(Site_Age=levels(soil.dat$Site_Age),
                    SpEig=0)


pred <- predict(TOC_model, newdata = pdat,
                level = 0, se.fit = TRUE,
                na.action = "na.exclude", type = "response")


predframe <- data.frame(pdat, pred = pred)


TOC_plot <- qplot(x = Site_Age,
                  y = pred.fit,
```

```
                  data = predframe,
                  xlab = "", ylab = "Organic carbon %") +
  geom_errorbar(aes(
    ymin  = (pred.fit-(pred.se.fit*1.96)),
    ymax  = (pred.fit+(pred.se.fit*1.96)),
    width = 0.15))+theme_classic() +
  geom_text(aes(label=c("a", "ab", "b", "ab","b"), family =
plotFont,  y=6.7), colour="black", vjust=0, size=4,
position=position_dodge(0))+
  coord_cartesian(ylim = c(1.5, 6.9)) +
  theme(axis.text.x=element_text(angle=90, hjust=1, vjust = 0.3))


TOC_plot


#### Total N ####


T_N_model <- lme(Total_N_PerCent ~ Site_Age+SpEig, random = ~
1|Site,
                   data = soil.dat, method = "ML")
plot(T_N_model)
summary(T_N_model)


r.squaredGLMM(T_N_model)


anova(T_N_model) # Sig
glhtSummary <- glht(T_N_model, linfct=mcp(Site_Age="Tukey"))
summary(glhtSummary) # Matches 95% confidence intervals



plot(T_N_model)
plot(T_N_model,Total_N_PerCent~fitted(.))
qqnorm(T_N_model,~ resid(.)|Site)


pdat <- expand.grid(Site_Age=levels(soil.dat$Site_Age),
                    SpEig=0)


pred <- predict(T_N_model, newdata = pdat,
                level = 0, se.fit = TRUE,
                na.action = "na.exclude", type = "response")


predframe <- data.frame(pdat, pred = pred)


T_N_plot <- qplot(x = Site_Age,
                  y = pred.fit,
```

```
                  data = predframe,
                  xlab = "", ylab = "Total nitrogen %") +
  geom_errorbar(aes(
    ymin  = (pred.fit-(pred.se.fit*1.96)),
    ymax  = (pred.fit+(pred.se.fit*1.96)),
    width = 0.15))+theme_classic()+
  geom_text(aes(label=c("a", "ab", "b", "ab","b"), family =
plotFont,  y=.58), colour="black", vjust=0, size=4,
position=position_dodge(0))+
  coord_cartesian(ylim = c(0.15, 0.61)) +
  theme(axis.text.x=element_text(angle=90, hjust=1, vjust = 0.3))


T_N_plot


#### NH4 ####


NH_model <- lme(NH4_mg_per_kg ~ Site_Age+SpEig, random = ~ 1|Site,
                  data = soil.dat, method = "ML")
plot(NH_model)
summary(NH_model)


r.squaredGLMM(NH_model)


anova(NH_model) #SIG
glhtSummary <- glht(NH_model, linfct=mcp(Site_Age="Tukey"))
summary(glhtSummary) # Matches 95% CIs


plot(NH_model)
plot(NH_model,NH4_mg_per_kg~fitted(.))
qqnorm(NH_model,~ resid(.)|Site)


pdat <- expand.grid(Site_Age=levels(soil.dat$Site_Age),
                    SpEig=0)


pred <- predict(NH_model, newdata = pdat,
                level = 0, se.fit = TRUE,
                na.action = "na.exclude", type = "response")


predframe <- data.frame(pdat, pred = pred)


NH_plot <- qplot(x = Site_Age,
                 y = pred.fit,
                 data = predframe,
```

```
                 xlab = "", ylab = expression("NH"[4]*""^"+"*" mg
Kg"^"-1") )+
  geom_errorbar(aes(
    ymin  = (pred.fit-(pred.se.fit*1.96)),
    ymax  = (pred.fit+(pred.se.fit*1.96)),
    width = 0.15))+theme_classic() +
  geom_text(aes(label=c("a", "b", "ac", "ac","c"), family =
plotFont,  y=17), colour="black", vjust=0, size=4,
position=position_dodge(0))+
  coord_cartesian(ylim = c(0, 17.5)) +
  theme(axis.text.x=element_blank())#,axis.text.x=element_text(angle
=40, hjust=1))


NH_plot


#### NO3 ####
NO_model3b <- lme(NO3_mg_per_kg ~ Site_Age+SpEig, random = ~ 1|Site,
                  data = soil.dat, method = "ML", na.action =
"na.omit")
plot(NO_model3b)
summary(NO_model3b)


r.squaredGLMM(NO_model3b)


plot(NO_model3b)
plot(NO_model3b,NO3_mg_per_kg~fitted(.))
qqnorm(NO_model3b,~ resid(.)|Site)


anova(NO_model3b) #SIG
glhtSummary <- glht(NO_model3b, linfct=mcp(Site_Age="Tukey"))
summary(glhtSummary) # Matches 95% confidence intervals


pdat <- expand.grid(Site_Age=levels(soil.dat$Site_Age),
                    SpEig=0)


pred <- predict(NO_model3b, newdata = pdat,
                level = 0, se.fit = TRUE,
                na.action = "na.exclude", type = "response")


predframe <- data.frame(pdat, pred = pred)


NO_plot <- qplot(x = Site_Age,
                 y = pred.fit,
                 data = predframe,
```

```
                  xlab = "", ylab = expression("NO"[3]*""^"-"*" mg
Kg"^"-1")) +
  geom_errorbar(aes(
    ymin  = (pred.fit-(pred.se.fit*1.96)),
    ymax  = (pred.fit+(pred.se.fit*1.96)),
    width = 0.15))+theme_classic() +
  geom_text(aes(label=c("a", "b", "ab", "ab","ab"), family =
plotFont,  y=30), colour="black", vjust=0, size=4,
position=position_dodge(0))+
  coord_cartesian(ylim = c(-4, 31))+
  theme(axis.text.x=element_blank())


NO_plot


#### CtoN ####
C_N_model <- lme(TotalCtoN ~ Site_Age+SpEig, random = ~ 1|Site,
                   data = soil.dat, method = "ML")
plot(C_N_model)
summary(C_N_model)


r.squaredGLMM(C_N_model)


plot(C_N_model)
plot(C_N_model,TotalCtoN~fitted(.))
qqnorm(C_N_model,~ resid(.)|Site)


anova(C_N_model) #Sig
glhtSummary <- glht(C_N_model, linfct=mcp(Site_Age="Tukey"))
summary(glhtSummary) # Matches 95% confidence intervals


pdat <- expand.grid(Site_Age=levels(soil.dat$Site_Age),
                    SpEig=0)


pred <- predict(C_N_model, newdata = pdat,
                level = 0, se.fit = TRUE,
                na.action = "na.exclude", type = "response")


predframe <- data.frame(pdat, pred = pred)


C_N_plot <- qplot(x = Site_Age,
                  y = pred.fit,
                  data = predframe,
                  xlab = "", ylab = "C:N") +
  geom_errorbar(aes(
```

```
    ymin  = (pred.fit-(pred.se.fit*1.96)),
    ymax  = (pred.fit+(pred.se.fit*1.96)),
    width = 0.15))+theme_classic() +
  geom_text(aes(label=c("a", "a", "b", "b","b"), family =
plotFont,  y=13.5), colour="black", vjust=0, size=4,
position=position_dodge(0))+
  coord_cartesian(ylim = c(9.5, 14))


C_N_plot


#### Bulk Density ####
dens_model <- lme(Bulk_Dens_g_cm3 ~ Site_Age+SpEig, random = ~
1|Site,
                  data = soil.dat, method = "ML", na.action =
"na.omit")


plot(dens_model)
summary(dens_model)


r.squaredGLMM(dens_model)


anova(dens_model) #SIG
glhtSummary <- glht(dens_model, linfct=mcp(Site_Age="Tukey"))
summary(glhtSummary) # Matches the 95% intervals



plot(dens_model)
plot(dens_model,Bulk_Dens_kg_per_ha~fitted(.))
qqnorm(dens_model,~ resid(.)|Site)


pdat <- expand.grid(Site_Age=levels(soil.dat$Site_Age),
                    SpEig=0)


pred <- predict(dens_model, newdata = pdat,
                level = 0, se.fit = TRUE,
                na.action = "na.exclude", type = "response")


predframe <- data.frame(pdat, pred = pred, grp = c("ab", "a",
"b","ab","ab"))


dens_plot <- qplot(x = Site_Age,
                   y = pred.fit,
                   data = predframe,
```

```
                    xlab = "", ylab = expression("Bulk Density g
cm"^"-3")) +
                    #aes(x=, y=pred.fit, group = grp)) +
  geom_errorbar(aes(
    ymin  = (pred.fit-(pred.se.fit*1.96)),
    ymax  = (pred.fit+(pred.se.fit*1.96)),
    width = 0.15))+ theme_classic()+
  scale_y_continuous(labels = function(x) format(x, scientific = F))
+
  geom_text(aes(label=c("a", "ab", "b", "ab","ab"),
                family = plotFont,  y=1.31),
            colour="black", vjust=0, size=4,
position=position_dodge(0))+
  coord_cartesian(ylim = c(.78, 1.33))+
  theme(axis.text.x=element_blank())


dens_plot


#### Assembling plots for publication ####


layout <- c("
AABB
CCDD
EEFF
GGGG


")


#svg(file= "Combo WrEN plots.svg", width = 6, height = 9)


pH_Plot +  dens_plot + NO_plot + NH_plot +
  TOC_plot + T_N_plot + C_N_plot + plot_layout(design=layout)
#dev.off()




#save(soil.dat, file = "Live files/ByrneSoilDat.RData")


#### Bibliography ####
# H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-
Verlag New York, 2016.
# Yuan Tang, Masaaki Horikoshi, and Wenxuan Li. "ggfortify: Unified
Interface to Visualize Statistical Result of Popular R Packages."
The R Journal 8.2 (2016): 478-489.
```

# R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
# Baptiste Auguie (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra
# Paul Murrell (2014). gridBase: Integration of base and grid graphics. R package version 0.4-7. https://CRAN.R-project.org/package=gridBase
# Alboukadel Kassambara (2019). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.2.4. https://CRAN.R-project.org/package=ggpubr
# Nan Xiao (2018). ggsci: Scientific Journal and Sci-Fi Themed Color Palettes for 'ggplot2'. R package version 2.9. https://CRAN.R-project.org/package=ggsci
# Thomas Lin Pedersen (2019). patchwork: The Composer of Plots. R package version 1.0.0. https://CRAN.R-project.org/package=patchwork

### *S2.c PERMANOVA results*

In order to test for a statistical difference between groups of multivariate data I employed a PERMANOVA, which indicates differences in the spread or centroid of groups. In order to determine differences between groupings, the assumption of homogeneity of multivariate dispersions must be tested (Table S2.c.1). In all PERMANOVA tests the permutations were constrained by site, to reflect the nested structure of the experiment, with cores located inside sites. The multivariate data analysed were the same variables used to generate Figure 2.1, i.e. pH, % total N, % total organic C, C:N, $NH_4$ (mg $kg^{-1}$), and $NO_3$ (mg $kg^{-1}$). I scaled these variables before analysis and calculated a euclidean distance matrix between each soil core for these six variables.

Land use groups had significantly different dispersions (Pseudo-$F(4,204) = 19.82$, $p < 0.001$), and post hoc permutation tests of the homogeneity of the groups' variances indicated that the assumption was violated for some, but not all groups (Table S2.c.1), comparisons between the following groupings may be valid: arable and pasture, arable and mature woodland, pasture and mature woodland, and young woodland and Ancient Semi-Natural Woodland (ASNW). All of these groupings were statistically significant (Table S2.c.1) but p-values were not corrected for multiple testing. Interpreting the importance of these significant values must be undertaken with caution, as many tests have been carried out and the assumptions of many of these tests have not been met.

Although some pairs did not differ in distribution and might be fairly compared using pairwise PERMANOVAs. The interpretation of these results must be undertaken with caution; differences in dispersion, sample numbers, and level of significance between groups add to uncertainty that is brought about by multiple statistical testing. I have chosen not to include them in the main text of chapter 2, as they do not add clarity to understanding of the main results, and the included analysis achieve the same goals as this analysis very well. I encourage the reader to draw their own conclusions about the importance of these results with the aid of the PCoA provided (Figure S2.c.2). In the context of the results reported in chapter 2, these results are somewhat surprising as there were few significant differences between woodland land uses, or between woodland and pasture in the GLMM based analyses. Many of these contrasts differ significantly when examined with a PERMANOA. However, the qualitative trends indicated in the PCoA appear to agree with the understanding of the soils from my GLMM analysis. Arable sites differ most strongly with young woodland and ASNW; and pasture sites sit intermediately between other land use types.

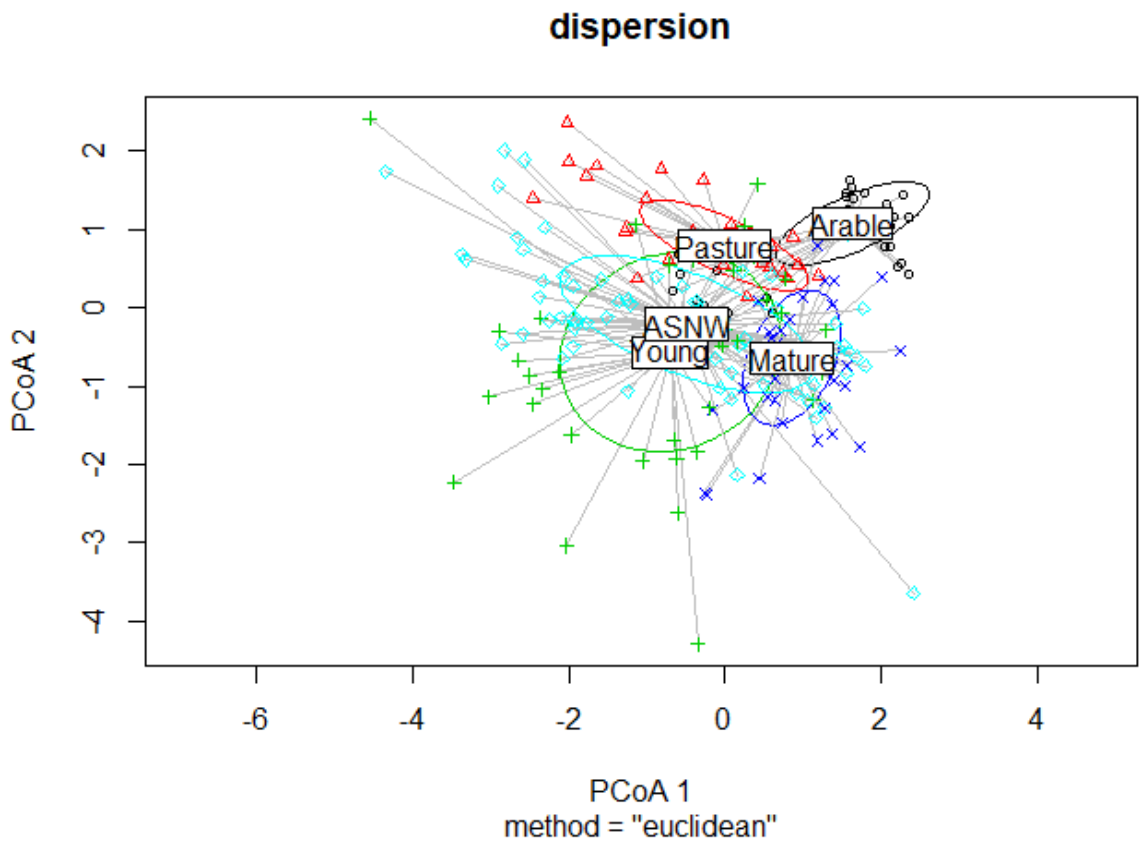


Figure S2.c.1 PCoA indicating land use differences in multivariate data. Ellipses indicate coverage of data within one standard deviation around the centroid.

Table S2.c.1 PERMANOVA and the following post-hoc, pairwise PERMANOVAs contrasting levels of land use. All models were fitted using 999 permutations. The p-value of a permutation test for homogeneity of multivariate dispersions is included before the results to indicate whether the assumptions of the PERMANOVA are met (i.e. a non-significant test result). No p-values are corrected for multiple testing.

| | Overdispersion p-value | d.f. | SS | R2 | Pseudo-F | P-value |
|---|---|---|---|---|---|---|
| *Overall* | | | | | | |
| Land use | 0.001 *** | 4 | 349.27 | 0.28 | 19.82 | 0.001 *** |
| Residuals | | 204 | 898.73 | 0.72 | | |
| Total | | 208 | 1248 | 1 | | |
| *Arable - Pasture* | | | | | | |
| Land use | 0.999 | 1 | 114.09 | 0.419 | 48.39 | 0.001 *** |
| Residuals | | 67 | 157.97 | 0.581 | | |
| Total | | 68 | 272.06 | 1 | | |
| *Arable - Young Woodland* | | | | | | |
| Landuse | <0.001*** | 1 | 155.73 | 0.31 | 34.96 | 0.001 *** |
| Residuals | | 78 | 347.42 | 0.69 | | |
| Total | | 79 | 503.14 | 1 | | |
| *Arable - Mature Woodland* | | | | | | |
| Landuse | 0.999 | 1 | 87.89 | 0.375 | 40.721 | 0.001 *** |
| Residuals | | 68 | 146.77 | 0.625 | | |
| Total | | 69 | 234.66 | 1 | | |
| *Arable - ASNW* | | | | | | |
| Landuse | <0.001*** | 1 | 158.19 | 0.233 | 32.861 | 0.001 *** |
| Residuals | | 108 | 519.89 | 0.767 | | |
| Total | | 109 | 678.07 | 1 | | |
| *Pasture - Young Woodland* | | | | | | |
| Landuse | <0.001*** | 1 | 71.56 | 0.181 | 14.835 | 0.001 *** |
| Residuals | | 67 | 323.18 | 0.819 | | |
| Total | | 68 | 394.74 | 1 | | |
| *Pasture - Mature Woodland* | | | | | | |
| Landuse | 0.999 | 1 | 86.5 | 0.413 | 40.24 | 0.001 *** |
| Residuals | | 57 | 122.53 | 0.586 | | |
| Total | | 58 | 209.04 | 1 | | |
| *Pasture - ASNW* | | | | | | |
| Landuse | 0.001*** | 1 | 67.47 | 0.12 | 13.21 | 0.001 *** |
| Residuals | | 97 | 495.65 | 0.88 | | |
| Total | | 98 | 563.13 | 1 | | |
| *Young Woodland - Mature Woodland* | | | | | | |
| Landuse | <0.001*** | 1 | 67.24 | 0.178 | 14.66 | 0.001 *** |
| Residuals | | 68 | 311.98 | 0.823 | | |
| Total | | 69 | 379.22 | 1 | | |
| *Young Woodland - ASNW* | | | | | | |
| Landuse | 0.974 | 1 | 17.91 | 0.025 | 2.8238 | 0.026* |
| Residuals | | 108 | 685.1 | 0.974 | | |
| Total | | 109 | 703.01 | 1 | | |
| *Mature Woodland - ASNW* | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Landuse | <0.001*** | 1 | 57.15 | 0.106 | 11.56 | 0.001 *** |
| Residuals | | 98 | 484.45 | 0.894 | | |
| Total | | 99 | 541.6 | 1 | | |

Table S2.c.2 Permutation test for homogeneity of multivariate dispersions.

| | d.f. | SS | MS | Pseudo-F | Permutations | P-value |
|---|---|---|---|---|---|---|
| Landuse | 4 | 44.98 | 11.246 | 10.85 | 999 | 0.001 *** |
| Residuals | 204 | 211.42 | 1.036 | | | |

***S2.d. file: ByrnePERMANOVA.RData***

***S2.e. ByrneS2PermanovaSscript.R***

```
###################################
####                           ####
#### ByrneS2PermanovaScript.R ####
####                           ####
###################################


#save(soil.dat, Site_Means, file = "PERMANOVAdat.Rdata")


load("PERMANOVAdat.Rdata")


library(vegan)


set.seed(12) #reproducible results


# Use this section to test on the whole dataset


permaMat <- scale(soil.dat[-141,c(
  "pH_H2O",
  "Total_N_PerCent",
  "TOC_PerCent",
  "TotalCtoN",
  "NH4_mg_per_kg",
  "NO3_mg_per_kg"
)])


soil.dist<-vegdist(permaMat, method='euclidean')
```

```
permaDat <- soil.dat[-141,c(
  "Site",
  "Site_Age",
  "pH_H2O",
  "Total_N_PerCent",
  "TOC_PerCent",
  "TotalCtoN",
  "NH4_mg_per_kg",
  "NO3_mg_per_kg"
)]


#### Use this section of the script to recalculate based on site
means



# permaMat <- scale(Site_Means[, c(
#   "pH_H2O",
#   "Total_N_PerCent",
#   "TOC_PerCent",
#   "TotalCtoN",
#   "NH4_mg_per_kg",
#   "NO3_mg_per_kg"
# )])
#
#
# #soil.mat<-sqrt(permaMat)#square root transform
#
# soil.dist<-vegdist(permaMat, method='euclidean')
# #soil.dist<-vegdist(soil.mat, method='bray')
#
#
# permaDat <- Site_Means[,c(
#   "Site",
#   "Site_Age",
#   "pH_H2O",
#   "Total_N_PerCent",
#   "TOC_PerCent",
#   "TotalCtoN",
#   "NH4_mg_per_kg",
#   "NO3_mg_per_kg"
# )]



# #PERMANOVA without nested study design incorporated
# soil.div <- adonis2(soil.dist~Site_Age, data=permaDat,
permutations = 999, method="euclidean")
# soil.div



# PERMANOVA design including site as a random effect
Soil.div2 <- adonis2(soil.dist~Site_Age, strata = Site, data =
permaDat, permutations = 999, method = "euclidean")
soil.div2
```

```
#An assumption of the test is that the data is not overdispersed.
This is a test that assumption


dispersion <- betadisper(soil.dist, group = permaDat$Site_Age)
permutest(dispersion) #significant differences in overdispersion
permutest(dispersion, strata = permaDat$Site)# - can be done but
doesn't seem to change results


plot(dispersion)
boxplot(dispersion)
mod.HSD <- TukeyHSD(dispersion)
mod.HSD
plot(mod.HSD)


plot(dispersion, hull=FALSE, ellipse=TRUE) ##sd ellipse


#Pairwise comparison, no p-value adjustment


combos <- combn(unique(levels(permaDat$Site_Age)),2)


results <-list()
sigs<- vector(length = 10)


for (i in 1:ncol(combos)){
  combo          <-  c(combos[1,i], combos[2,i])
  contrast       <- permaDat[permaDat$Site_Age %in% combo,]
  tempMat        <- permaMat[permaDat$Site_Age %in% combo,]
  Soil.dist.temp <- vegdist(tempMat, method='euclidean')
  Soil.div       <- adonis2(soil.dist.temp~Site_Age, strata = Site,
data=contrast, permutations = 999, method="euclidean")
  sigs[i]        <- soil.div$`Pr(>F)`[1]
  results[[i]]   <- soil.div
}


results
resTable <- rbind(combos,sigs)
resTable #results of all pairwise comparisons
```

***Table S2.e. Site Summary Data***

| Site | Type | Landuse | Soil Type | Texture | SpEig | pH | Total N % | TOC % | C:N | $NH_4$ ($mg^{-1}$ kg) | $NO_3$ ($mg^{-1}$ kg) | Bulk Density ($g^{-1}$ cm) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Alpha Wood | Woodland | Young | Deep Clay | Silty Clay Loam | -1.04 | 5.28 | 0.67 | 7.67 | 11.43 | 8.12 | 10.68 | 0.67 |
| Alpha Farm | Agricultural | Arable | Deep Clay | Silty Clay | -1.04 | 5.81 | 0.40 | 4.34 | 10.78 | 4.75 | 32.24 | 0.98 |
| Bravo Wood | Woodland | ASNW | Deep Clay | Clay Loam | -1.35 | 5.55 | 0.41 | 4.57 | 11.29 | 5.67 | 3.39 | 1.15 |
| Bravo Farm | Agricultural | Arable | Deep Clay | Clay Loam | -1.35 | 6.95 | 0.24 | 2.33 | 9.90 | 1.66 | 18.62 | 1.27 |
| Charlie Wood | Woodland | ASNW | Deep Clay | Clay Loam | -0.79 | 7.38 | 0.60 | 7.10 | 12.21 | 7.07 | 17.45 | 1.05 |
| Delta Wood | Woodland | ASNW | Deep Clay | Clay Loam | -0.99 | 7.46 | 0.68 | 7.06 | 10.96 | 12.20 | 36.21 | 1.02 |
| Echo Wood | Woodland | Young | Seasonally wet deep clay | Clay Loam | -1.17 | 4.25 | 0.44 | 5.65 | 12.66 | 2.73 | 32.18 | 0.84 |
| Echo Farm | Agricultural | Pasture | Deep clay | Clay Loam | -1.16 | 6.01 | 0.34 | 3.56 | 10.58 | 10.62 | 5.59 | 1.01 |
| Foxtrot Wood | Woodland | ASNW | Deep clay | Sandy Clay Loam | 0.85 | 6.64 | 0.21 | 2.64 | 13.67 | 4.17 | 0.62 | 1.17 |
| Foxtrot Farm | Agricultural | Arable | Deep clay | Sandy Clay Loam | 0.85 | 6.21 | 0.17 | 1.68 | 10.06 | 1.33 | 24.41 | 1.18 |
| Golf Wood | Woodland | ASNW | Seasonally wet deep red clay | Clay Loam | 0.59 | 4.91 | 0.29 | 3.35 | 11.76 | 4.39 | 11.46 | 0.91 |
| Golf Farm | Agricultural | Pasture | Seasonally wet deep red clay | Clay Loam | 0.59 | 6.68 | 0.48 | 4.79 | 10.18 | 19.87 | 4.60 | 0.96 |
| Hotel Wood | Woodland | Mature | Seasonally wet red loam to clay | Clay Loam | 1.31 | 5.45 | 0.25 | 3.15 | 12.29 | 2.97 | 2.07 | 1.06 |
| India Wood | Woodland | Mature | Deep clay | Clay Loam | 0.67 | 6.14 | 0.32 | 4.38 | 13.79 | 4.70 | 3.92 | 1.24 |
| India Farm | Agricultural | Pasture | Deep clay | Silty Clay Loam | 0.47 | 5.91 | 0.38 | 3.88 | 10.09 | 10.49 | 2.32 | 1.08 |
| Juliet Wood | Woodland | ASNW | Seasonally wet loam to clayey over red shale | Clay Loam | 0.78 | 6.54 | 0.46 | 5.51 | 12.25 | 10.04 | 3.51 | 0.94 |
| Kilo Wood | Woodland | Young | Seasonally wet deep clay | Clay Loam | 1.16 | 7.46 | 0.35 | 3.97 | 12.10 | 8.68 | 6.09 | 1.12 |
| Kilo Farm | Agricultural | Arable | Seasonally wet deep clay | Sandy Clay Loam | 1.16 | 7.38 | 0.18 | 1.88 | 10.66 | 2.24 | 9.52 | 1.30 |
| Lima Wood | Woodland | Mature | Seasonally wet deep loam to clay | Sandy Clay Loam | 1.03 | 6.47 | 0.30 | 3.47 | 11.62 | 4.61 | 2.50 | 1.20 |
| Mike Wood | Woodland | Young | Seasonally wet deep clay | Silty Clay | 0.67 | 5.76 | 0.46 | 5.04 | 10.95 | 7.70 | 8.04 | 0.89 |
| November Wood | Woodland | ASNW | Deep clay | Clay Loam | -1.23 | 6.35 | 0.52 | 5.87 | 11.20 | 5.68 | 4.85 | 0.84 |

**Chapter 3: Supplementary Data**

***S3.1 NMDS plots***

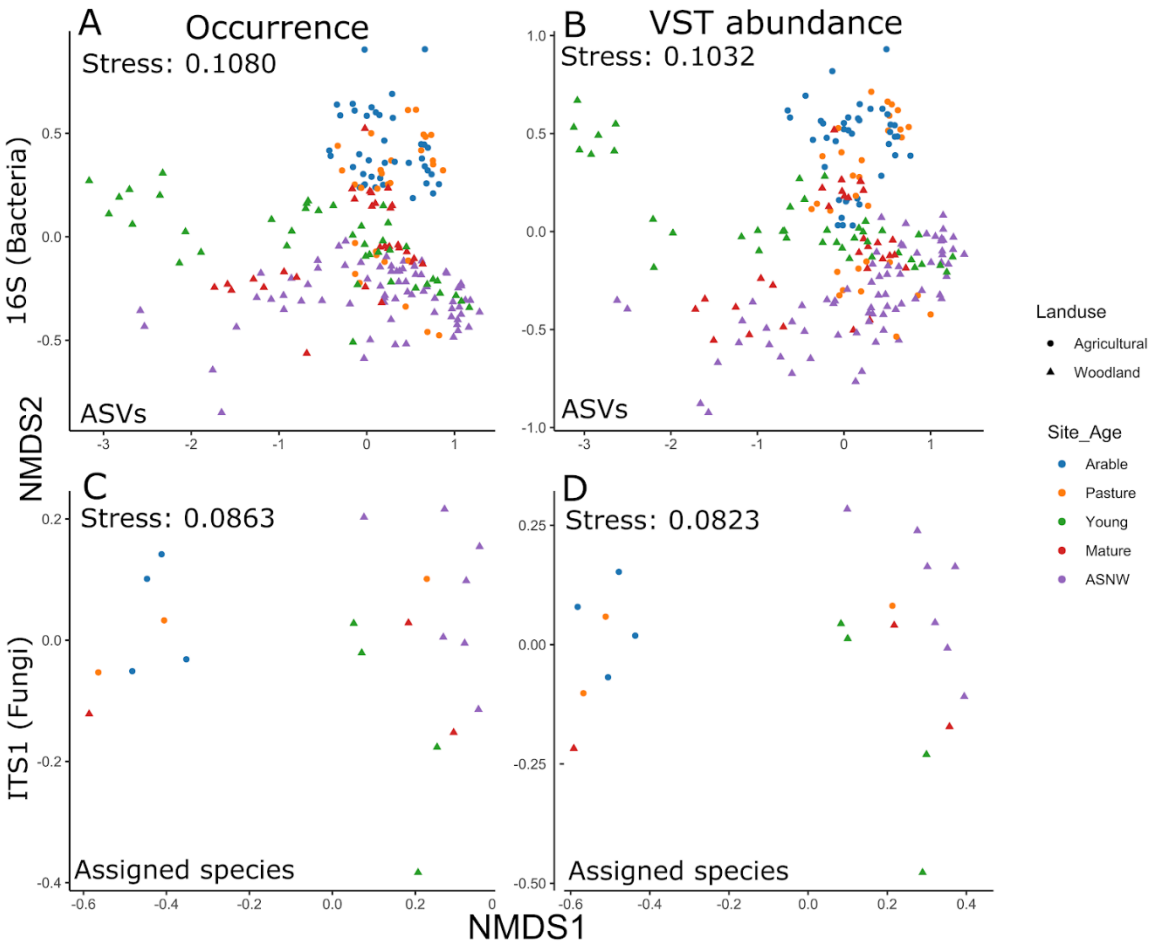


Figure S3.1 Bray-Curtis community dissimilarity NMDS ordinations of ASV occurrence (A), transformed ASV abundance (B), species occurrence (C), and transformed species abundance (D). Bacterial samples (A and B) are ordinated individually, but fungal samples are pooled by site (C and D). Land use is indicated by point colour, with circles representing all agricultural land use types, and triangles representing all woodland land use types. The direction and units of NMDS axes are arbitrary.

### *S3.2 Solutions and protocol for DNA*

*Buffer components*

| Component | Contents | Required | Chemical molarity | per L | per 100 mL | pH |
|---|---|---|---|---|---|---|
| **Lysis solution 1** | guanidine thiocyanate | 147 mM | 118.16 | 17.33 g | 1.73 g | 9.0 |
| | trisodium phosphate | 228 mM | 380.13 | 86.67 g | 8.67 g | |
| | sodium chloride | 26 mM | 58.44 | 1.5 g | 0.15 g | |
| | 1 M Tris HCl | 67 mM | - | 67 mL | 6.7 mL | |
| | 0.5 M EDTA | 27 mM | - | 53 mL | 5.3 mL | |
| **Lysis solution 2** | aluminium ammonium sulphate* | 90 mM | 453.33 | 40.8 g | 4.08 g | - |
| | SDS | 1.25 % | - | 12.5 g | 1.25 g | |
| **Protein flocculant** | ammonium acetate | 5 M | 77.0825 | 385.41 g | 38.54 g | - |
| **Inhibitor flocculant 1** | aluminium ammonium sulphate* | 180 mM | 453.33 | 81.6 g | 8.16 g | - |
| **Inhibitor flocculant 2** | calcium chloride dihydrate | 204 mM | 147.01 | 30 g | 3 g | - |
| **Binding solution** | guanidine HCl | 5.5 M | 95.53 | 525.42 g | 52.54 g | - |
| **Wash solution** | EtOH | 80 % | - | - | - | - |
| **Elution buffer** | Pure $H_2O$ or 1mM Tris | - | - | - | - | 8.0 |

*aluminium ammonium sulphate = aluminium ammonium sulphate dodecahydrate (CAS 7784-26-1), if using anhydrous powder, adjust the calculation above.

Adjust the pH of ***Lysis solution 1*** to 9.0 with 5M HCl and bring to volume with $ddH_2O$. This will probably require much less than 20 ml of HCl and should be mixed in the fume hood. All other components apart from ***Elution buffer*** are used at the pH of the mixture without modification. Sterilise all solutions in a suitable manner (autoclave or filter).

***Protocol***

*Sample lysis*

1. Add 2 grams of 1 - 1.4 mm diameter sterile garnet beads to a 5ml eppendorf, screw cap tube
2. Add 2200 µL of ***Lysis solution 1*** and vortex briefly.
3. Add .25 grams sample to tube, shake briefly by hand to mix contents.
4. Place in Geno/Grinder 2010 with appropriate adapters @ ~ 1750 RPM for 2 mins
5. Wait 30 seconds
6. Grind again for an additional 2 mins at 1750 RPM
7. Add 800 µL of ***Lysis solution 2*** per gram of sample.
8. Centrifuge @ 4,000xg for 1 min @ room temperature.
9. Transfer supernatant to a fresh 1.5 ml tube. (Transfer 1ml and save 500µl)
10. Centrifuge @ 10,000xg for 1 min @ room temperature .
11. Transfer 500 µl of supernatant to fresh tube 1.5 ml tube.
    *Note: Remaining lysate can now be stored @ -20°C for future work*

*DNA purification*

12. Add 200 µl volume of ***Protein flocculant***, vortex briefly and incubate on ice for a minimum of 10 mins.
13. Centrifuge @ 10,000xg for 1 min @ room temperature.
14. Transfer supernatant to fresh tube 1.5ml tube.
15. Make a mastermix of:

a. 1.1*n samples*100 µl of ***Inhibitor flocculant 1.***

b. 1.1*n samples*100 µl of ***Inhibitor flocculant 2.***

16. Add 200 µl of inhibitor flocculant mastermix.
17. Centrifuge @ 10,000xg for 1 min @ room temperature.
18. Transfer supernatant to fresh 5ml tube.
19. Add 1568 µl of ***Binding solution***, invert several times to mix.
20. Fill silica spin column to capacity with the above mixture, centrifuge @ 10,000xg for 1 min @ room temperature, discard flow-through and repeat until all mixture has passed through the spin column.
21. Add 392 µl of **Wash solution**, centrifuge @ 10,000xg for 1 min @ room temperature, discard flow-through.
22. Centrifuge @ 10,000xg for 1 min @ room temperature, replace collection tube with a fresh 1.5ml tube.
23. Add 313 µl of ***Elution buffer*** **heated to 70 °C** directly to the silica filter membrane. Leave for 2 minutes at room temperature.
24. Centrifuge @ 10,000xg for 1 min @ room temperature.
25. DNA is now in the collection tube.

**Chapter 4: Supplementary Data**

***S4.1 Site Data***

| Site | Age | triplicate | Lat | Long | texture | Dominant Genus | pH | organic C (%) | nitrogen (%) | C:N | available P (mg kg-1) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Paddock Wood | Woodland Creation | A | 55.1968 | -1.654 | silt loam | Quercus | 5.9 | 2.3 | 0.13 | 19.01 | 5.64 |
| Nunsbrough Wood | Mature Secondary | A | 54.9296 | -2.079 | silt loam | Ulmus | 5.5 | 3.97 | 0.263 | 16.51 | 5.35 |
| Nunsbrough Wood | ASNW | A | 54.9298 | -2.0791 | sandy silt loam | Malus | 5.8 | 4.24 | 0.27 | 15.38 | 5.58 |
| Castle Hill | Woodland Creation | B | 54.8369 | -1.8794 | sandy silt loam | Quercus | 4.9 | 6.21 | 0.362 | 20.28 | 3.93 |
| Congburn | Mature Secondary | B | 54.8525 | -1.6032 | silt loam | Fagus | 4.5 | 3.94 | 0.11 | 30.18 | 7.48 |
| Congburn | ASNW | B | 54.853 | -1.6019 | silt loam | Quercus | 4.4 | 3.78 | 0.123 | 23.75 | 4.18 |
| Railway Wood | Woodland Creation | C | 54.8415 | -1.5091 | silt loam | Populus | 6.1 | 4.73 | 0.274 | 20.39 | 5.58 |
| Wallington | Mature Secondary | C | 55.1485 | -1.9499 | clay loam | Fagus | 6.1 | 3.26 | 0.157 | 19.16 | 7.61 |
| Wallington | ASNW | C | 55.147 | -1.9589 | silty clay loam | Acer | 6.5 | 4.46 | 0.287 | 15.87 | 5.69 |
| Stanley Burn | ASNW | D | 54.9589 | -1.8241 | loamy sand | Fraxinus | 5.2 | 3.04 | 0.123 | 29.63 | 5.73 |
| Pontburn | Woodland Creation | D | 54.9012 | -1.7724 | loamy sand | Quercus | 5.4 | 3.27 | 0.215 | 18.67 | 7.8 |
| Vigo Wood | Mature Secondary | D | 54.8785 | -1.5577 | sandy loam | Fagus | 5 | 5.65 | 0.549 | 33.41 | 7.67 |
| Dora's wood | Woodland Creation | E | 54.8169 | -1.7383 | sandy silt loam | Quercus | 5.3 | 2.03 | 0.147 | 28.31 | 4.42 |
| Hamsterly | Mature Secondary | E | 54.9041 | -1.8119 | sandy silt loam | Fraxinus | 5.3 | 4.79 | 0.27 | 19.33 | 3.77 |
| Allensford | ASNW | E | 54.8484 | -1.8717 | sandy silt loam | Quercus | 5.2 | 3.67 | 0.212 | 18.95 | 5.65 |
| Hedley Hall | Woodland Creation | F | 54.9021 | -1.6613 | sandy silt loam | Fraxinus | 5.6 | 2.76 | 0.132 | 27.09 | 7.4 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Harraton Generals Wood | Mature Secondary | F | 54.8774 | -1.5386 | sandy silt loam | Fagus | 5.6 | 7.94 | 0.34 | 26.64 | 7.59 |
| Sacriston | ASNW | F | 54.8258 | -1.6358 | sandy loam | Betula | 5.3 | 13.27 | 0.432 | 45.61 | 3.64 |
| | | | | | | | | | | | |
| Newfield | Woodland Creation | G | 54.8585 | -1.6193 | sandy loam | Fraxinus | 6.9 | 4.48 | 0.244 | 28.78 | 5.87 |
| Newfield | Mature Secondary | G | 54.8596 | -1.6248 | sandy silt loam | Alnus | 7.1 | 3.73 | 0.16 | 44.64 | 7.61 |
| Whittle Dean | ASNW | G | 54.9832 | -1.8858 | sandy loam | Acer | 6.7 | 3.45 | 0.231 | 18.96 | 5.76 |
| | | | | | | | | | | | |
| Bright Lea Wood | Woodland Creation | H | 54.8924 | -1.5906 | sandy loam | Quercus | 6.1 | 2.78 | 0.156 | 30.18 | 7.78 |
| Morton Wood | Mature Secondary | H | 54.8421 | -1.5063 | sandy loam | Quercus | 5.8 | 33.2 | 1.251 | 35.39 | 7.34 |
| Letah Wood | ASNW | H | 54.9373 | -2.0955 | loamy sand | Quercus | 5.8 | 1.63 | 0.112 | 17.55 | 5.79 |
| | | | | | | | | | | | |
| Deerness | Woodland Creation | I | 54.7777 | -1.6587 | sandy loam | Quercus | 6.8 | 5.85 | 0.226 | 32.89 | 6.19 |
| Elemore Woods | Mature Secondary | I | 54.7914 | -1.445 | sandy silt loam | Acer | 6.7 | 3.2 | 0.246 | 19.88 | 7.62 |
| Hesleden Dene | ASNW | I | 54.7336 | -1.317 | sandy loam | Fraxinus | 7.1 | 13.17 | 0.801 | 29.86 | NA |

**Bioinformatics Supplementary Data**

***S5.1 Read data outputs from DADA2 prior to ASV filtering.***

The total read numbers input to the DADA2 pipeline (input), passed quality filtering (filtered), passed forward and revers denoising (denoisedF and denoisedR), passed pair-ends merging (merged), and passes chimera removal (non chimera). Non-chimera reads were passed on to further filtering for low abundance ASVs and other checks described in the relevant chapter.

| Run informaton | Experiment | Loci | Run # | input | filtered | denoisedF | denoisedR | merged | non chimera |
|---|---|---|---|---|---|---|---|---|---|
| Illumina MiSeq v3 (2x 300 bp) | Chapter 3 | ITS1 | 1 | 14,464,606.00 | 10,067,365.00 | 9,997,922.00 | 9,985,969.00 | 9,851,486.00 | 9,744,230.00 |
| Illumina MiSeq v3 (2x 300 bp) | Chapter 3 | 16S | 2 | 14,819,312.00 | 12,491,308.00 | 11,969,691.00 | 11,976,797.00 | 10,630,231.00 | 10,595,797.00 |
| Illumina MiSeq v3 (2x 300 bp) | Chapter 5 | ITS1 | 3 | 3,540,284 | 3,071,082 | 3,047,943 | 3,045,210 | 3,010,918 | 3,001,315 |
| Illumina MiSeq v3 (2x 300 bp) | Chapter 5 | 16S | 3 | 6,730,437 | 6,338,138 | 6,233,789 | 6,252,527 | 5,912,814 | 5,874,458 |
| Illumina HiSeq 2500 (2x250bb) | Chapter 5 | ITS1 | 4 | 95,179,580 | 90,408,976 | 90,014,706 | 89,979,690 | 87,962,740 | 87,313,475 |
| Illumina HiSeq 2500 (2x250bb) | Chapter 5 | 16S | 4 | 103,272,513 | 98,901,472 | 96,545,294 | 96,909,604 | 90,381,256 | 90,034,764 |

***S5.2 Negative read data outputs from DADA2 prior to ASV filtering.***

The total read numbers of each negative input to the DADA2 pipeline (input), passed quality filtering (filtered), passed forward and revers denoising (denoisF and denoisR), passed pair-ends merging (merged), and passes chimera removal (nonchim).

| Run Details | Run # | Chapter | Locus | Plate | Negative | input | Filtered | denoisF | denoisR | merged | nonchim |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 1 | N1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch.3 | ITS1 | 1 | N2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 2 | N1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 2 | N2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 3 | N1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 3 | N2 | 2 | 1 | 1 | 1 | 1 | 1 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 4 | N1 | 58 | 13 | 13 | 13 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 4 | N2 | 964 | 167 | 166 | 166 | 166 | 166 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 5 | N1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 5 | N2 | 1 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 6 | N1 | 15 | 3 | 3 | 3 | 3 | 3 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 6 | N2 | 80 | 55 | 55 | 55 | 55 | 55 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 7 | N1 | 104 | 40 | 40 | 40 | 40 | 40 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 7 | N2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 8 | N1 | 52 | 25 | 12 | 22 | 12 | 12 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 8 | N2 | 2 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 9 | N2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 1 | Ch. 3 | ITS1 | 10 | N2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 11 | N1 | 16 | 9 | 2 | 1 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 11 | N2 | 1354 | 999 | 983 | 981 | 963 | 963 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 12 | N1 | 15 | 8 | 8 | 8 | 8 | 8 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 12 | N2 | 787 | 542 | 524 | 516 | 500 | 500 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 13 | N1 | 8 | 1 | 1 | 1 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 13 | N2 | 1129 | 852 | 846 | 851 | 845 | 845 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 14 | N1 | 15 | 8 | 1 | 1 | 1 | 1 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 14 | N2 | 1253 | 1011 | 1011 | 1011 | 1011 | 1011 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 15 | N1 | 18 | 5 | 5 | 5 | 5 | 5 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 15 | N2 | 573 | 428 | 425 | 424 | 422 | 422 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 16 | N1 | 4 | 3 | 1 | 1 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 16 | N2 | 49 | 39 | 35 | 35 | 35 | 35 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 17 | N1 | 6 | 2 | 1 | 1 | 1 | 1 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 17 | N2 | 166 | 125 | 121 | 121 | 96 | 96 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 18 | N1 | 45 | 15 | 1 | 2 | 0 | 0 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 18 | N2 | 278 | 221 | 193 | 203 | 189 | 189 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 19 | N2 | 359 | 286 | 273 | 265 | 261 | 261 |
| Illumina MiSeq v3 (2x 300 bp) | 2 | Ch. 3 | 16S | 20 | N2 | 208 | 160 | 139 | 143 | 118 | 118 |
| Illumina MiSeq v3 (2x 300 bp) | 3 | Ch. 5 | ITS1 | 1 | N1 | 526 | 460 | 358 | 361 | 341 | 341 |
| Illumina MiSeq v3 (2x 300 bp) | 3 | Ch. 5 | ITS1 | 1 | N2 | 472 | 408 | 356 | 342 | 341 | 341 |
| Illumina MiSeq v3 (2x 300 bp) | 3 | Ch. 5 | 16S | 2 | N1 | 2 | 2 | 1 | 1 | 1 | 1 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 4 | N1 | 8 | 2 | 2 | 2 | 2 | 2 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 4 | N2 | 107 | 99 | 92 | 87 | 75 | 75 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 6 | N1 | 48 | 11 | 3 | 3 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 6 | N2 | 5 | 5 | 3 | 3 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 8 | N1 | 13 | 5 | 2 | 2 | 1 | 1 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 8 | N2 | 281 | 259 | 215 | 225 | 195 | 195 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 10 | N1 | 61 | 9 | 6 | 2 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 10 | N2 | 11 | 10 | 8 | 8 | 8 | 8 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 12 | N1 | 141 | 34 | 27 | 27 | 27 | 27 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 12 | N2 | 34 | 31 | 21 | 22 | 14 | 14 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 14 | N1 | 71 | 11 | 6 | 6 | 6 | 6 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 14 | N2 | 30 | 29 | 16 | 20 | 16 | 16 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 16 | N1 | 1664 | 323 | 271 | 286 | 265 | 265 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 16 | N2 | 10 | 6 | 2 | 2 | 2 | 2 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 18 | N1 | 14 | 3 | 3 | 3 | 3 | 3 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 18 | N2 | 22 | 20 | 20 | 20 | 13 | 13 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 20 | N1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | 16S | 20 | N2 | 200 | 180 | 167 | 164 | 158 | 158 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 3 | N1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 3 | N2 | 24 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 5 | N1 | 133 | 98 | 84 | 88 | 76 | 76 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 5 | N2 | 1 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 7 | N1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 7 | N2 | 1 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 9 | N1 | 89 | 62 | 62 | 62 | 62 | 62 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 9 | N2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 11 | N1 | 84 | 78 | 78 | 78 | 78 | 78 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 11 | N2 | 9690 | 8964 | 8964 | 8936 | 8929 | 8929 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 13 | N1 | 2101 | 1544 | 1544 | 1544 | 1544 | 1544 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 13 | N2 | 2727 | 2511 | 2511 | 2511 | 2511 | 2511 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 15 | N1 | 4 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 15 | N2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 17 | N1 | 2 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 17 | N2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 19 | N1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Illumina HiSeq 2500 (2x250bb) | 4 | Ch. 5 | ITS1 | 19 | N2 | 0 | 0 | 0 | 0 | 0 | 0 |