GREATSOILS

June 2023

Project Report No. 91140002-06

Assessing soil ecosystem health – evaluating more detailed biological community measures and DNA-based approaches

Francois Dussart[1], Matías Fernández-Huartes[2], Christine Watson[3], Nicola Holden[3], John Elphinstone[2], Joana Vicente[2], Anne Bhogal[4], Nadia Andreani[5], Matthew Goddard[5], Matthew Shepherd[6], Paul Hargreaves[7], Bryan Griffiths[1] and Elizabeth Stockdale[8]

[1]SRUC, Peter Wilson Building, The King's Buildings, West Mains Road, Edinburgh, EH9 3JG

[2]Fera Science Ltd., Sand Hutton, York, YO41 1LZ

[3]3RUC, Ferguson Building, Craibstone Estate, Aberdeen, Bucksburn, AB21 9YA, UK

[4]ADAS Agriculture, Spring Lodge, 172 Chester Road, Helsby, WA6 0AR

[5]The School of Life Sciences, The University of Lincoln, Lincoln, LN6 7DL, UK

[6]Natural England, Sterling House, Dix's Field, Exeter, EX1 1QA

[7]SRUC, Dairy Research Centre, Crichton Royal Farm, Dumfries, DG1 4TT

[8]NIAB, 93 Lawrence Weaver Road, Cambridge, CB3 0LE

This is the final report of a 57 month project (Project 6 of the Soil Biology and Soil Health Partnership) which started in April 2017. The work was funded by AHDB and BBRO £175,030.

Contents

1. Abstract

As part of the work within the Soil Biology and Soil Health Research and Knowledge Exchange (SBSH) Partnership, this project aimed to examine soil ecosystem-focussed approaches to study soil health and to evaluate DNA-based approaches as a part of soil health measurements.

Soil biological communities are considered to be sensitive to management change. Together with the soil microbial community, soil mesofauna such as nematodes, mites (Acari) and springtails (Collembola) play a vital role in organic matter turnover, nutrient cycling and soil structure stabilisation. However, current methods to study soil biological communities are time-consuming and need expert skills; they are consequently slow and costly. For soil or water, DNA extraction methods usually recover the microorganisms (bacteria, fungi, archaea, viruses etc), alive, dead or decaying, as well as residual DNA left from higher organisms, such as trails of mucus and sloughed off cells; together this extracted DNA is termed 'environmental DNA' (eDNA). Extraction of eDNA followed by metabarcoding has enabled the identification of animal organisms from complex populations and has the potential to reduce the cost of analysis.

In the project, three different soil DNA extraction methods with different intensity and sample size were compared to assess the impact on the DNA sequence determination. A method that aimed to extract only extracellular DNA was not effective at recovering DNA from soil; no extracts composed of only extracellular eDNA could be obtained. For soil-based work, the project has concluded that the distinction between eDNA and total DNA extraction is confusing and semantic. Effective extraction of DNA from soils requires methods that also lead to cell lysis and hence the intra-cellular DNA of soil microorganisms is measured along with extracellular DNA.

DNA extraction method was shown to have a larger impact on the measured biological community than a range of long-term organic amendment treatments in one long-term trial. When the data were pooled across extraction methods, the total number of species recorded was significantly higher than separate extractions. Therefore, it is likely that the different extraction approaches target different parts of the community. However, the pooled data also showed no effect of organic amendment treatment on the total number of species or community composition; organic amendment had a small effect on relative species richness. Standardisation of DNA extraction methods was confirmed to be important when comparing samples collected from different locations and at different times.

Measurements of soil biological communities were made in a number of long-term trials where the drivers of soil biological function (pH, fertiliser, organic amendment addition, tillage) were also assessed using the soil health scorecard approach. The value of the soil ecosystem-based approaches was validated by examining soil treatment differences. For nematode community structures, site factors such as soil type and land use were the main influencing factors. Management factors appeared to have a smaller effect, with some variations arising from nutrient loading and pH.

This was also the case for the mesofauna assessments. Current nematode and mesofauna community analysis approaches aren't suited to routine commercial application as the methods required are laborious requiring specialist identification skills. A relatively simple characteristic of the mesofauna community, such as the total density of springtails, or the ratio of mites and springtails (as %Collembola), may provide a possible simple indicator linked to soil organic matter cycling, and properties likely to enhance soil function in an agricultural context. DNA-based methodologies may also offer a more practical and affordable solution for these analyses in the future.

For the soil microbial community, strong divergent trends were found for multiple taxonomic diversity indices in response to pH, supporting previous findings that small changes in soil pH can produce changes in community composition. In contrast, a smaller impact on community composition occurred in response to other management change, including long-term fertiliser treatments, organic amendment or tillage. As seen in other studies, geographical location significantly affected microbial community composition, here three different experimental sites with similar organic amendment treatments yielded different results. The findings suggest that low-clay soils may be more responsive to organic amendment treatments. Where strong significant divergence in microbial communities occurred in response to pH, it was possible to identify specific taxa and to infer functional divergence between microbial communities. This represents one of the first attempts to assess the feasibility of functional analysis based on DNA, alongside the broader assessment of soil health.

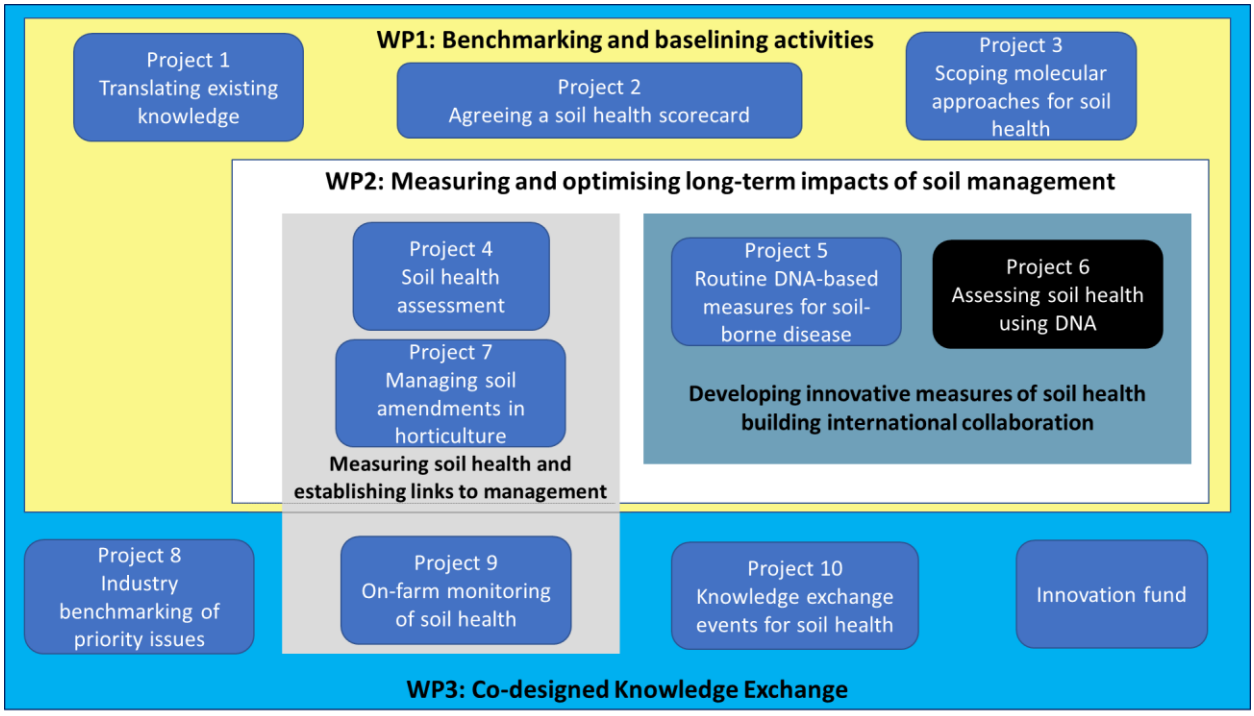
Although individual studies on soil biology have reported the effect of agricultural management practices, the implementation of diversity indexes has differed between studies. The project has developed a standardised approach to DNA analysis, building on the use of the same methods across multiple studies with high variation, to gain a broader understanding of management impacts on soil biology. The project has demonstrated that this approach is able to test the extent to which bacterial and fungal diversities change in response to management and the need to assess the magnitude of those changes across multiple datasets. However, as the costs of measurement are still very high for each sample and no UK-wide benchmarking framework can currently be established, the project does not currently recommend the use of soil ecosystem-focussed indicators for routine soil health monitoring on farm.

The detailed morphological analysis and molecular methods (DNA extraction and metabarcoding) used within this project have delivered an increased understanding of soil biological health, in particular the interacting impacts of both site and soil management. Reflection on these data together with the broader soil health data sets (presented in Project 4) alongside the simple visual tool developed at the start of the Partnership has validated the embedded qualitative relationships. Therefore, the descriptive visual tool can provide useful indications for farmers and consultants considering a change to management with positives and challenges for soil biology and soil health highlighted.

2. Introduction

This project is part of a suite of integrated projects within the Soil Biology and Soil Health Research and Knowledge Exchange Partnership (see Diagram below showing how this project fits into the wider organisation of projects). This project (Project 6 of the Soil Biology and Soil Health Partnership, together with Project 5) sought to develop innovative measures of soil health, largely based on molecular approaches. This work builds on the review "Molecular approaches for routine soil-borne disease and soil health assessment – establishing the scope" presented in Project 3 (Elphinstone et al., 2017). Project 6 was conducted within Work Package 2 (WP2) of the Soil Biology and Soil Health (SBSH) Partnership which brings together Projects 4, 5, 6, and 7. The overall aim of WP2 was to improve understanding of the role of soil biology in overall soil health, with respect to the ability of soils to support and sustain healthy and productive crops. In particular, it aimed to establish any aspects of soil biology that are key to this function of soils and to explore practical and effective ways to measure and manage these key components alongside already established measurement and monitoring of soil physical and chemical properties.

Project 06 shown within the integrated project delivery of the Soil Biology and Soil Health Research and Knowledge Exchange Partnership



2.1. Aims and Objectives

This project (Project 6 of the SBSH Partnership) sought to develop methods for the analysis of soil extracted DNA that could replace the existing incubation and examination techniques for a series of metrics: microbial biomass, fungal and bacterial biomass, functional gene analysis (with the potential to replace measures of nitrification and denitrification) and nematode community analysis. It also aimed to better understand how assessment of environmental DNA (eDNA) in soil could be used as a metric for soil health, including the assessment of the abundance and diversity of mesofauna (Decaëns et al., 2013). Ultimately the aim is to enable the further development of effective on-farm decision support systems for farmers and agronomists.

These DNA based approaches were evaluated on the samples from long-term field trials collected in Project 4 and hence also provide further evaluation of the effects of soil management practices on the soil biological community and its function. To the best of our knowledge this is the first test of the eDNA technique for soil samples under UK conditions.

Given the complexities in the interactions between management practices and the physical, chemical and biological properties of soil, their interaction and consequent soil function, simple predictions of impacts are not possible. In fact, farmers themselves routinely mention that the same management practice can result in different effects in different seasons or on different soil types. In addition, Sherwood and Uphoff (2000) note that one of the challenges for improving soil management is that of supporting farmer engagement with information so that they can identify and prioritise problems and opportunities, test and evaluate innovations and become partners in sharing the information gained. Hence the SBSH Partnership has developed a visual tool (Project 1) which allows farmers to investigate the impacts of a range of management scenarios. The tool is designed to display the likely effects of the different management scenarios visually, in a way that shows both the complexities (i.e. the interconnections between physics, biology and chemistry) and the trade-offs (e.g. that with added compost you are likely to increase the positive biological attributes and yield, even though you may immobilise nutrients and need more effective weed control). Therefore, the SRUC-based team regularly integrated data from across the SBSH Partnership and the evaluation and development in the visual tool is also reported here.

2.1.1. Specific objectives of Project 6

1. Short-list targets for analysis for DNA-based approaches from: molecular biomass; total bacteria and fungi; microbial community structure; functional genes; nematodes; microarthropods; earthworms. (This objective was covered in detail within Project 3 and the outputs of that review were re-evaluated and updated during this project).
2. Optimise the extraction of environmental DNA (eDNA) from soil and compare with other approaches to soil DNA extraction to describe the soil microbial community.

3. Compare results from soil extracted DNA with current methods for assessing soil health used on the experimental sites in Project 4 and further evaluate the effects of soil management practices (rotations, amendments and cover crops) on the soil biological community and its function.
4. Evaluate whether more detailed information of the soil biological community should become part of the soil health scorecard to provide information whether through microscopy or molecular testing.
5. Update the visual tool showing likely impacts of management developed in Project 1, using appropriate data from across the SBSH Partnership.

Work for Objectives 1 to 4 linked closely with that taking place within Project 5 and was carried out in partnership between SRUC, Fera and the University of Lincoln.

2.2. Background

Soil microbial communities have an important role in regulating nutrient and carbon cycling in soil. Below-ground microbial communities interact directly with plant roots and may be a key factor affecting plant species diversity and productivity. The composition and diversity of the soil microbial community are also widely recognised as key factors in enabling land-based ecological functions.

Soil mesofauna such as nematodes, mites (Acari) and springtails (Collembola) also play a vital role in organic matter turnover, nutrient cycling and the development of soil structure, and are widely regarded as a soil component that is vulnerable to management changes (Blair et al., 1997, Stork and Eggleton, 1992); they also include important plant pathogens. Whilst plant parasitic nematode populations are routinely determined in soils growing sensitive crops (e.g. root crop and horticultural rotations), assessment of the wider (and larger) mesofauna population is often overlooked in soil health monitoring studies (George et al., 2017). This is not only due to practical problems associated with sampling and identification, but also a lack of understanding of the links between mesofauna populations and soil processes, the metrics to use and their interpretation (Stork and Eggleton, 1992).

Advances in DNA sequencing technologies (as reviewed for the SBSH Partnership in Project 3) have led to unprecedented changes in our understanding of soil microbial communities. However, this is still an emerging area raising important questions. Therefore, it is important to understand the constraints and influences in any real-world application so that the evidence provided is robust and trustworthy.

2.2.1. Nematodes and mesofauna communities as soil health indicators

Nematodes (commonly known as round worms) are a diverse group of soil organisms most recognized for their potential for crop damage, either via direct feeding damage or through the transmission of viruses. Indeed, it is this group of nematodes for which commercial laboratories regularly test. However, these are just one component of the nematode community within soils, which is characterised by a wide range of organisms that can be grouped according to how they feed, and which are distributed throughout the soil food web: bacterial feeders, fungal feeders, omnivores, predators and plant feeders (Liu et al., 2020). This diversity makes them ideally suited to indicate the biological health of the soil, with various ecological indices developed to classify nematode community structure (Ferris et al., 2001, Ritz and Trudgill, 1999, Biswal, 2022).

Project 2 of the SBSH partnership proposed a way of interpreting two such indices, the enrichment and structure index (which are given a scored between 0-100 based on the work of Ferris et al., 2001), in order to position the soil nematode community in relation to an idealised optimum (Figure 1). The enrichment index reflects the presence of those nematodes that reproduce rapidly in response to high nutrient levels (generally bacterial feeders), whereas the structure index reflects the stability of the food web characterised by relatively long-lived predatory and omnivorous species. The 'best' soil biological health is indicated by a balanced community with a lower enrichment index (scoring 25-50) than the structure index (scoring 50-75). Very low values for the structure index or very high values for the enrichment index indicate a very unbalanced nematode community which is very likely to result from an 'unhealthy' soil biological community. Griffiths *et al*. (2018), turned the position of the nematode community measured in a soil sample using the graph shown in Figure 1 into a score from 1 (good biological health) to 5 (poor biological health). These were then grouped into the three traffic lights used by the scorecard in a similar manner to the Visual Evaluation of Soil Structure (VESS) scoring system, whereby a score of 1 and 2 would be green, 3 amber and 4 and 5, red. Such an analysis would be in addition to the determination of specific plant parasitic nematodes, depending on the crops grown in the rotation.

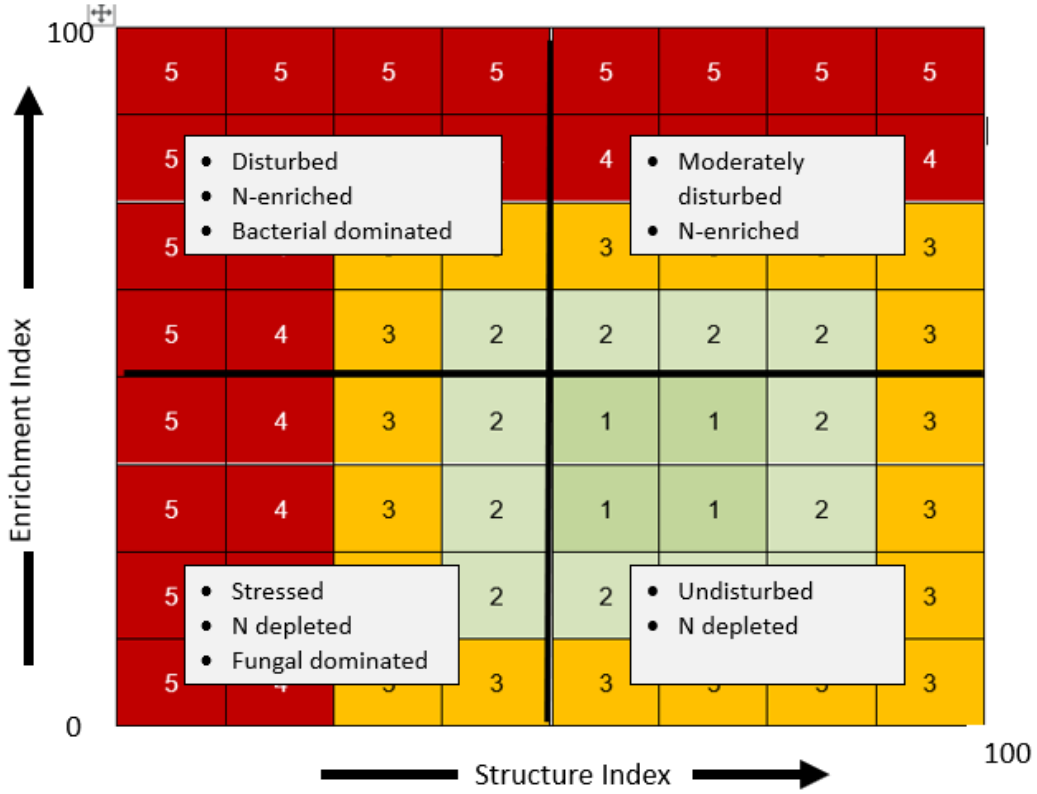


Figure 1: Outline of the use of ecological index values (enrichment and structure) for nematode community analysis. Scores (0-100) for the enrichment and structure index are allocated using the Ninja software and then a further overall index (1-5) is assigned to indicate overall soil biological health. The allocation of the index is based on the two-way plot shown and these index values are colour-coded: red – investigate, potentially poor biological health; amber – review; green – good biological health, continue rotational monitoring (from Griffiths et al. 2018).

Mesofauna (including microarthropods) represent a major component of the soil food web involved in the decomposition of organic matter, nutrient cycling and the development of good soil structure (Barrios, 2007). Acari (mites) and Collembola (springtails) are the most abundant groups of mesofauna and their inclusion within national soil health monitoring programmes has been encouraged (George et al., 2017). However, whilst clear differences in populations have been identified across broad habitat types (e.g. arable, woodland, moorland and heathland), differences between agricultural management practices or similar habitat types (e.g. managed vs. semi natural grasslands) have been difficult to detect, particularly at high taxonomic levels (George et al., 2017).

Know Your Mesofauna

Oribatida mites are usually slow moving, long lived fungivores and decomposers, and do not breed rapidly. Astigmatina mites are a subdivision of oribatida adapted to rapid breeding, dispersal and more temporary habitats. Poduroidea (= Poduromorpha) and Entomobryoidea (= Entomobryomorpha) are decomposer springtails that breed relatively rapidly, but can also attain larger sizes, and are associated with rapid nutrient turnover. The Mesostigmata are mostly predatory mites that feed on springtails and other mesofauna. Prostigmata includes a wide range of mites, including some that show similar dispersal characteristics to the Astigmatina, and can have very high rates of breeding. Neelipleona are tiny springtails, whose ecology is little known, and Symphypleona are globular springtails which include both tiny soil-dwelling forms and larger competitive species.

However, there is currently no consensus on the most appropriate metric to describe the mesofauna community overall or how to apply benchmarks to measures of the mesofauna community, or component species, to show whether a soil has ‘good’ biological health. Various metrics have been proposed including: abundance, biomass, density, species richness, trophic/guild structure, food web structure, keystone species and ecosystem engineers, and Stork & Eggleton (1992) suggested that measurements of keystone species, taxonomic diversity and species richness could provide a ‘simple guide’ to soil biological health. However, a potential interpretation framework for these metrics in the context of UK agricultural soils has not been developed.

2.2.2. Role of molecular biology in monitoring soil microbiology

New molecular tools are allowing a step-change in the study of microbial populations associated with soil and roots (Project 3 report; Elphinstone et al., 2017). DNA detection first requires extraction of DNA from the substrate. For soil or water, all extraction methods usually recover the microorganisms (bacteria, fungi, archaea, viruses etc), alive, dead or decaying, as well as residual DNA left from higher organisms, such as trails of mucus and sloughed off cells left behind by earthworms or nematodes; together this extracted DNA is termed ‘environmental DNA’ (eDNA). Metabarcoding then targets specific organisms or groups of organisms based on the DNA sequence of a common or ubiquitous region (also known as an amplicon). Since assessment of DNA and DNA sequencing are highly technical disciplines, there are a number of terms that warrant definition to aid the reader. The terms have been ordered by complexity, from the biological communities to specific techniques.

Term	Definition
Soil microbial community	The living microscopic biological entities within the soil, includes bacteria, archaea and fungi.
Microbiome	A microbiome is the collection of microorganisms that live together in a given environment. The term soil microbiome is now used more commonly but is synonymous with the earlier term soil microbial community.
eDNA	Environmental DNA: DNA collected from an environmental sample e.g. water, sediment, soil, rather than directly sampled from an individual organism. This may include DNA from microorganisms (bacteria, fungi, archaea, viruses etc), alive, dead or decaying, as well as any residual DNA and/or extracellular DNA e.g. from animal residues and microbial DNA.
Sequencing	Molecular technique used to decipher the genetic code of a given target. The sequencing machine used in this project has the tradename of MiSeq (Illumina™). This defines the type of sequencing where very short-regions of DNA are decoded and are assembled to longer regions that can then be mapped to a reference dataset for identification.
Metabarcoding	High-throughput sequencing approach that determines the composition of a community of organisms, based on the DNA sequence of a common or ubiquitous region (also known as an amplicon)
ASV	Amplicon Sequence Variant: DNA sequence identity of an organism to the genera or sub-genera level of taxonomy. Although indicative of a species type, amplicon-type high-throughput sequencing approaches do not provide sufficient resolution for absolute species identity
16S	A DNA region encoding a ribosomal gene ubiquitous in bacteria that is used for taxonomic identification, in conjunction with a database of known bacterial ribosomal DNA sequences (Silva)
ITS	A DNA region between ribosomal genes ubiquitous in eukaryotes, including fungi & oomycetes, used for taxonomic identification in conjunction with the Unite database
COI	A DNA region encoding cytochrome oxidase subunit 1 gene ubiquitous in animals that is used for taxonomic identification in conjunction with a CO1 database developed at Fera Science Ltd.
Alpha-diversity	Variation in microbes within a single sample of the microbiome Comprises two measurements, ‘richness’ for the number of different species in the sample; and ‘species diversity’, for how evenly the microbes are distributed in a sample, using statistical indices
Beta-diversity	Variation in microbial communities between separate samples of the microbiome, measured using statistical indices
qPCR	Amplification of known targets of DNA using PCR reactions for their absolute quantification. It differs from sequencing in that it is a diagnostic technique that does not decipher sequences where the target is unknown.

Extraction of eDNA followed by metabarcoding has enabled the identification of animal organisms from complex populations, including water samples from ponds, rivers, lakes (Valentini et al., 2016) and deep-sea sediments (Lecroq et al., 2011). This approach has the potential to circumvent identification of higher organisms with time consuming methods that examine morphological characteristics (Lawton et al., 1998), which requires a high degree of taxonomic expertise (Andre et al., 2001), and can only identify species from adult specimens (Griffiths et al., 2002). As such, morphological analyses are generally resolved relatively coarsely leaving any associated ecological analysis potentially ambiguous or superficial (Yeates and Bongers, 1999). The use of eDNA (Bohmann et al., 2014) also has the potential to dramatically reduce the cost of analysis. Methods that target extracellular DNA in the soil matrix can use a simple buffer extraction instead of a more protracted extraction of DNA from within living cells. Buffer-only extractions coupled with high-throughput DNA sequencing were used successfully by a project partner (Goddard, University of Lincoln) to study soil fungal communities in New Zealand to reveal landscape-scale patterns, and the forces that drive them (Morrison-Whittle and Goddard, 2015). The eDNA method, however, still needs to be fully evaluated for soil monitoring purposes, and considered in the context of soil health.

Practicalities that need to be considered include how to make the DNA sample representative. Research at Fera has overcome this by increasing the amount of soil that can be feasibly processed by an order of magnitude from 0.25 – 0.5 g soil to up to 10 g soil. This is important because small amounts of soil are less likely to be sufficient to give representative samples for nematodes and other indicator fauna, as recent research indicates that 200 g samples are required (Wiesel et al., 2015). The breakthrough by Fera has been the introduction of routine DNA extraction from 250 – 500 g soil for quantification of target organisms (qPCR), and from 10 g for overall taxonomic composition, ensuring a representative sample is obtained. It has been well documented for any nucleic acid-based assessment that the extraction method can impact the output and hence interpretation of the results. A key aim of this project is to identify the best approaches that enable DNA analysis to be incorporated as a part of routine soil analyses.

Molecular approaches for biological soil health assessment have been reviewed in detail as part of Project 3 ‘Molecular approaches for routine soil-borne disease and soil health assessment – establishing the scope’. This covers considerations for sampling, and for DNA extraction and purification. It discusses how targeted approaches such as quantitative PCR compare with untargeted community analyses and lists biological indicators suited for soil health assessment. Soil-extracted DNA is used to detect and quantify most of the populations and functions that are needed to measure soil’s biological functioning. These include microbial biomass (Plassart et al., 2012), bacterial and fungal biomass (Bru et al., 2011), nematode abundance (Wiesel et al., 2015) and specific nematode taxa (Vervoort et al., 2012). We can also determine relative community composition of bacteria, fungi and nematodes (Plassart et al., 2012). Quantification of functional

genes related to key processes such as nitrification and denitrification (Bru et al., 2011) is also commonplace. Testing for specific pathogens and beneficial organisms in soil, based on a DNA extract from that soil, has been made available commercially in some countries, e.g. in Australia (www.pir.sa.gov.au/research/services/molecular_diagnostics/predicta_b). However, this field is still in an early stage of development and therefore the SBSH Partnership provided an excellent opportunity to determine whether these molecular-based techniques could be adopted in a wider assessment of soil health and soil function.

Use of high throughput sequencing and metabarcoding to study impacts of soil management

Lauber et al. (2009) carried out one of the first studies using high throughput sequencing approaches from a wide array of ecosystem types to describe a correlation between pH and soil bacterial community structure. More generally, recent soil DNA sequencing studies have shown significant but small and inconsistent differences between fungal (Hannula et al., 2021, Morrison-Whittle and Goddard, 2015) and bacterial (Hendgen et al., 2018) diversity when comparing soils under conventional management approaches compared with those of conservation agriculture. However, some studies have reported that fertiliser application increases richness and diversity (Pan et al., 2020, Wang et al., 2017) while others found no significant influence of fertilisers (Yao et al., 2018). Studies of microbial communities have included simultaneous analyses of bacteria and fungi (Hartmann et al., 2015), and total soil biology across time and space (Giraldo-Perez et al., 2021b), and include attempts to analyse the functions of these communities (Harkes et al., 2019). Recent studies have also shown how fungal community structure and functionality (Hannula et al., 2021), as well as bacterial diversity (Hartmann et al., 2015), have been affected by different long-term agricultural practices such as tillage, cover cropping and organic amendment. Moreover, Hannula et al. (2021) and Giraldo-Perez et al. (2021b) concluded that different components of soil biodiversity responded differentially to agricultural practices depending on geographic location and time of year. In general, studies to date have focused on different agricultural system types, and only evaluated soils in one or a few locations and timepoints and have lacked methodological and analytical standardisation making it comparison across studies difficult.

Linking DNA-based measurements with the measurements made using current methods for assessing soil health on the experimental sites in Project 4 has therefore allowed a further evaluation of the effects of soil management practices (rotations, amendments and cover crops) on the soil biological community and its function. This allowed an updated evaluation of the Visual tool to support management decisions and also allowed a focussed evaluation of whether more detailed information on the soil biological community (whether obtained through microscopy or molecular testing) should become part of the soil health scorecard.

3. Materials and methods

3.1. Selection of targets and sites for soil ecosystem-based analysis

A short-list of targets for DNA-based analysis was determined from analysis of the published literature, this detailed review of targets for DNA analysis was reported in Project 3.

A selection of experimental sites from Projects 4 and 5 were selected to provide a breadth of soil health, soil conditions and farm management types to allow comparison of methods across a range of contrasting soil management approaches. A summarised version is shown Table 1, with more detail included in the report of Projects 4 (Craibstone, Gleadthorpe, Harper Adams, Terrington) and Project 5 (Cranfield Asparagus).

Standard Operating Procedures for practical work were agreed at the outset to ensure standardisation. Soil chemical and microbiological assessments were undertaken on a single representative topsoil sample per replicate plot at each site, taken from 0-15 cm. Samples were also taken at the same time for nematode community analysis (representative mixed soil sample) and other mesofauna (intact cores).

3.2. Extraction and identification of nematodes and mesofauna from soil

Nematodes and mesofauna were measured at the Project 4 long-term experimental sites (Table 1 – excluding the site at Cranfield) with the addition of a long-term no-till field at GWCT Loddington which had had ploughed treatments imposed as part of Project 4. The range of soil conditions (nutrient inputs, tillage, pH, crop rotation) at these sites was expected to give rise to different nematode and mesofauna community structures.

Nematodes

Nematodes were extracted from a representative c. 200 g subsample of topsoil collected alongside the Project 4 soil health assessments at each of the long-term experimental sites. The Seinhorst two flask technique was used for small and juvenile nematodes (Seinhorst, 1955), with the larger species extracted using a modified Cobb decanting and sieving method with Baermann funnel extraction (Flegg, 1967). Sieve apertures were 53 microns for the small species and 150 microns for the large species. The numbers of free-living plant feeding nematodes, bacterial feeders, fungal feeders, omnivores and predatory nematodes in a sample were determined using a microscope at x400 magnification. An analysis of the community composition was then undertaken using the 'Ninja' software (Sieriebriennikov et al. 2014) to derive enrichment and structure Indices for each of the soils. Statistical analysis of the results was undertaken by ANOVA and t-test, depending on the number of treatments at each site.

Table 1: Trial sites used for assessment of detailed soil biological parameters, with associated main trial treatments. At Harper Adams, soil biological parameters were determined on all treatments in 2017 (just before cultivation of the grass-clover ley phase of the rotation), but only treatments 1-3 in 2020 (following winter wheat, potatoes and spring barley).

Site	**Treatment variables**	**Soil texture (% clay)**
Craibstone-2018	pH: 7 pH subplots 4.5-7.5 in 0.5 increments Large plots following the rotation: **Potatoes**, **W. Wheat**, Hay, Pasture, **Pasture**, **Oats**, Swede, S. Barley (each crop present every year) Measures in 4 crops (in bold) & 4 pH levels (4.5, 6.0, 6.5 & 7.5)	Sandy loam (12% cl)
Craibstone-2019	Fertiliser: 1. No fertiliser 2. Complete fertiliser (N, P, K) Fertiliser subplots in large plots following the rotation: **Swede**, **Barley**, Hay, Pasture, **Pasture**, **Oats** (each crop present every year) Measures in 4 crops (in bold) & 2 fertiliser treatments	Sandy loam (12% cl)
Gleadthorpe	Organic amendments: 1. Manufactured fertiliser (RB209 recommendations) 2. 10 t/ha Broiler litter 3. Green compost (@250 kg N/ha; c.25 t/ha) 4. Cattle FYM (@250 kg N/ha; c.35 t/ha) 5. Cattle slurry (@250 kg N/ha; c.80 m^3/ha) 3 replicates (15 plots)	Loamy sand (6% cl)
Harper Adams	Organic amendments: 1. Manufactured fertiliser (RB209 recommendations) 2. Green compost (@250 kg N/ha; c.25 t/ha) 3. Cattle FYM (@250 kg N/ha; c.35 t/ha) 4. Cattle slurry (@ 250 kg N/ha; c. 60-80m^3/ha) 5. Green/food compost (@250 kg/ha N; c. 20 t/ha) 6. Food-based digestate (@180 kg/ha N; c. 40m^3/ha) 3 replicates (18 plots)	Sandy loam (12% cl)
Terrington	Organic amendments: 1. Manufactured fertiliser (RB209 recommendations) 2. Green compost (@250 kg N/ha; c.25 t/ha) 3. Pig FYM (@250 kg N/ha; c.35 t/ha) 3 replicates (9 plots)	Silty clay loam (28% cl)
Cranfield - Asparagus	Re-ridging compared with zero tillage Split plots with six treatments using inter-row companion crops or mulching: 1. Bare soil 2. Bare soil with shallow soil disturbance. 3. Mustard (*Sinapis alba* L. var. Severka) companion crop (broadcast in late August and removed the following March). 4. Rye (*Secale cereale* L var. Protector) companion crop (broadcast in late August and removed the following March). 5. Mulch of PAS100 green compost with shallow soil disturbance. 6. Mulch of straw with shallow soil disturbance.	Sandy silt loam (12% cl)

Mesofauna

Intact soil cores (5 cm diameter x 5.7 cm deep or 7.5 cm diameter x 5 cm deep, depending on the site) were collected from each replicate plot/sub-plot for the determination of mesofauna community composition. Cores were placed upside down in Tullgren funnels for extraction over five days, with the mesofauna collected in 70 % alcohol (Crotty et al., 2012), prior to being counted and identified using a microscope. For most sites and sampling times, in keeping with the soil mesofauna sampling conducted as part of the Countryside Survey (1998 & 2007; Emmett *et al.,* 2007), taxonomic classification was at a "major group" level, with Collembola identified to the level of Order and Mites grouped into orders Mesotstigmata and Endeostigmata, suborders Prostigmata and Oribatida, excepting the cohort Astigmatina, which were counted separately, and all other micro-arthropods identified to Order. For mesofauna sampled at the Craibstone pH experiment in 2017 and Gleadthorpe in 2020, mesofauna were identified to the highest level of taxonomic detail available (Collembola, Oribatida and some Mesostigmata to species, Prostigmata to species, genus or family, Astigmatina to genus). Statistical analysis of the results was undertaken by non-metric multidimensional scaling (NMDS) to understand the relative similarity between mesofauna communities at each of the sites and treatments, and by ANOVA, comparing mesofauna community species richness, population and diversity (Shannon H diversity index), with all analyses performed in R.

For the mesofauna data, a cross site analysis was also performed whereby data from Harper Adams organic amendments experiment, Gleadthorpe organic amendment experiment, Loddington tillage (direct drill or ploughing) and Craibstone pH /rotation plots were combined using the following process. Acari data were split into Mesostigmata (order), Non-astigmatine Oribatida (Suborder), Astigmatina (cohort) and Prostigmata (suborder) plus Endeostigmata (sub-order). Collembola were split into the four orders (Entomobryomorpha, Poduromorpha, Symphypleona and Neelipleona). All other animals were grouped by order (Araneae, Coleoptera, Diptera, Psocoptera, Hemiptera, Thysanoptera, Hymenoptera, Lepidoptera, Haplotaxida, Enchytraeida, Pauropoda, Isopoda, Diplopoda, Symphyla). All data were converted and expressed as individuals per m^2, because the samples were taken using different sized rings. The overwhelming majority of individuals in soil mesofauna populations tend to be springtails and mites, but analyses of communities may be unduly influenced by presence or absence of individual specimens of other taxonomic groups unique to particular sites. Furthermore, there was some disparity between the presence or absence of Endeostigmata, which may have been recorded as Prostigmata during some analyses. Therefore a further analysis was carried out using only acari and collembola and grouping together Endeostigmata and Prostigmata.

These data were analysed using NMDS and the scores for each sample were grouped as means for each combination of site treatment and sampling period. These data were compared with other soil

parameters collected and reported as part of Project 4 using either Pearson r on raw or log/square root transformed data, or using Spearman's Rank correlation. Due to the high numbers of comparisons, Bonferroni's correction was applied to reduce type 1 errors and maintain the family-level error rate at 0.05 for comparison of each test statistic against the range of soil parameters, so that only correlations with a P value of <0.00143 were considered significant.

Other mesofauna community parameters, that could be calculated from mesofauna data from all sites without using NMDS, were explored to determine whether they could be used to broadly represent the axis scores derived from NMDS, and again, correlations were explored between these, and the other soil parameters reported in Project 4 using correlations, again applying the same Bonferroni correction to reduce type 1 errors.

Differences in NMDS scores and associated mesofauna community parameters were compared between treatments using ANOVA, with all analyses being conducted in R or using https://www.statskingdom.com/two-way-anova-calculator.html.

Because of the relatively limited range of sites available, data from the experimental plots were summarised to report proportions of each of the 4 orders of springtail, and proportion of total acari, and these data were collated into a single dataset alongside:

- 22 samples of soil mesofauna communities analysed using metabarcoding of COI genes, from arable land sampled for Countryside Stewardship monitoring by Natural England at six locations around the UK. See Figure 2 for sample locations.
- 14 samples of soil mesofauna communities from three lowland low intensity grassland sites forming part of Natural England's Long Term Monitoring Network (LTMN) analysed by metabarcoding of COI genes. See Figure 2 for sample locations.
- 308 samples of soil mesofauna communities sampled as part of the Countryside Survey 1998, from Arable and horticulture, Calcareous grassland, Neutral Grassland and Improved Grassland habitat types, and identified by microscopy.
- 273 samples of soil mesofauna communities sampled as part of the Countryside Survey 2007, from the same habitat types as above, and identified by microscopy.

The collated dataset was analysed by NMDS to provide a broader context for the variability within the experimental sites, and scores on the 2 NMDS axes explored to detect correlations with other measured soil properties and using ANOVA to detect any impact of treatments.

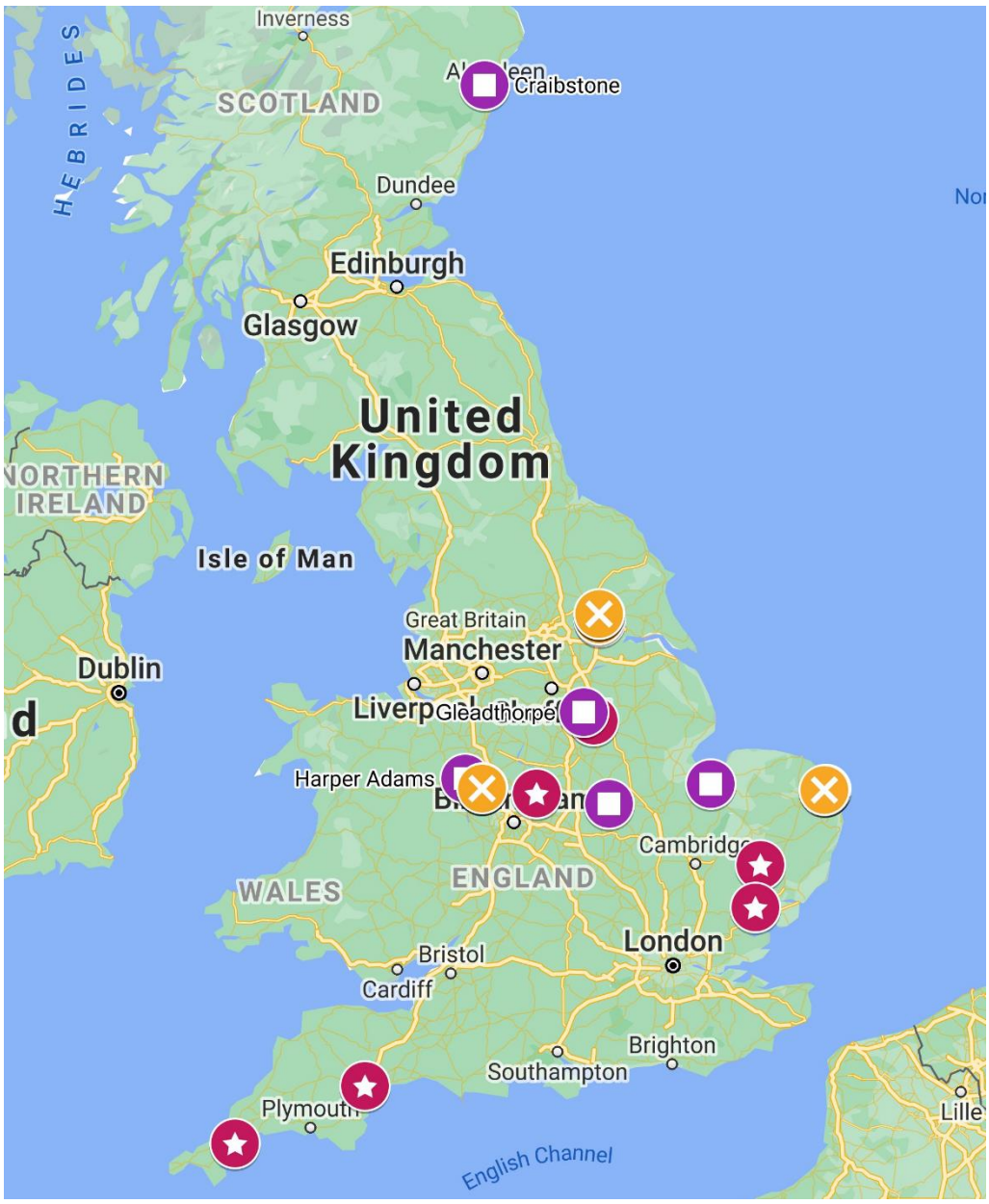


Figure 2: Location of AHDB experimental sites (shown as purple circles with white squares), LTMN plots (orange circles with white crosses) and Countryside Stewardship monitoring plots (red circles with white stars). Countryside Survey 1998 and 2007 data was also included in this analysis of collated results, and these sites are likely to be distributed widely across Great Britain, but their exact locations are not available.

3.3. DNA extraction methods

Initially, soil samples collected in October 2017 from the Harper Adams experimental site, were used to compare results from soil extracted DNA and eDNA, in a fully factorial manner (Figure 3). The extraction was validated by a quantitative PCR approach (qPCR) for microbial (bacterial and fungal) plant pathogens (methods described in Project 5). The extracted DNA was also used for detection of microbial and mesofauna community structure using high-throughput sequencing.

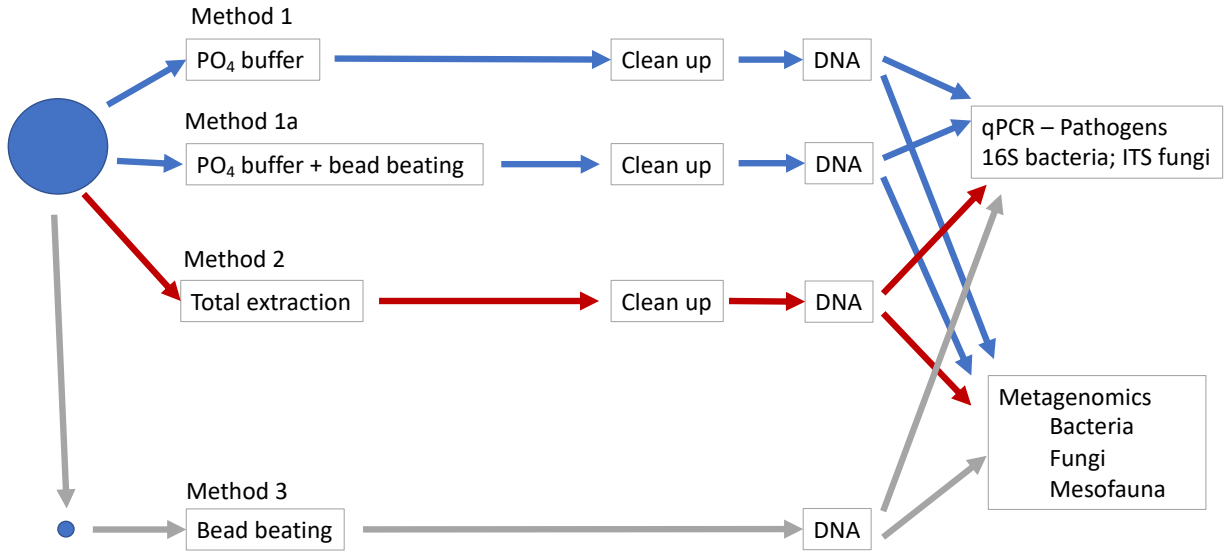


Figure 3: Diagram explaining how an extraction targeted at extracellular DNA (PO_4 buffer – Method 1) was compared with two conventional methods of DNA extraction (Methods 2&3). Method 1 was adapted to include a bead-beating stage (Method 1a). The size of the blue circles shows the amounts of soil used for extract and reflect the fact that typically tiny (< 1g) amounts of soil are used in many cases.

Three different extraction methods were compared to assess the impact on the DNA sequence determination (Figure 2). Methods 2 & 3 incorporated steps that released cellular DNA and so were more similar to contemporary approaches that quantify and describe microbial communities. In contrast, Method 1 was designed so that it was more likely to yield only the extracellular DNA. However poor DNA yields from this method led to the introduction of a bead-beating step (Method 1a). Following DNA extraction, bacterial (16S) and fungal (18S) rRNA gene targets were analysed using quantitative detection (qPCR) and community sequencing (metabarcoding) techniques at the Fera laboratories. All extracted eDNA was archived at -80 °C.

3.3.1. Method 1 (buffer only PO_4 - SRUC)

The soil was used fresh and sieved to remove all large stones. A phosphate buffer was prepared by dissolving 1.96 g NaH_2PO_4 and 14.68 g Na_2HPO_4 in 1 L of distilled water and adjusting pH to 8 with NaOH. Soil was homogenised by mixing 250 mL of phosphate buffer to 250 g of soil, using a planetary ball mill PM400 (Retsch, Haan, Germany) for 2 min at 300 rpm. The mixture of soil and buffer was then incubated at 4°C for 20 min under constant agitation. DNA extraction was carried out using the NucleoSpin®Soil kit (Macherey-Nagel, Düren, Germany) following manufacturer's instruction but omitting the initial sample preparation step. DNA purification was carried out using the Genomic DNA Clean & Concentrator kit (Zymo Research, Irvine, CA, USA) following manufacturer's instruction. DNA was quantified using a QuBit3 fluorometer (Invitrogen, Carlsbad, CA, USA).

The method was adapted Method 1a to include bead beating with the addition of glass beads for 5 minutes at the start of the agitated incubation stage.

3.3.2. Method 2 (total extraction - Fera)

Soil was air-dried before being weighed out into appropriate samples with gravel and coarse organic debris removed by hand. DNA was extracted from 250 g sub-samples of dried soil using a method adopted from Woodhall, *et al*., 2012 (Fera Science, unpublished). Sub-samples were homogenised in 500 mL grinding buffer (120 mM sodium phosphate buffer pH8, 2 % cetrimonium bromide, 1.5 M sodium chloride) and 15 mL Antifoam B by shaking in a paint shaker (Merris Engineering Ltd, Ireland) for 4 mins with 20 stainless steel ball bearings (25.4 mm diameter). A 50 mL sub-sample was then centrifuged at 5000 g for 5 min. A 20 mL sub-sample of the supernatant was added to 2 mL of 5M potassium acetate, incubated on ice for 10 mins and centrifuged for 5 min at 12,000 g. The supernatant was shaken for 15 min. at 11 rpm with 15 mL isopropanol and 800µl 100 % silicon dioxide suspension (Sigma), followed by 5 min centrifugation at 12,000 g. The supernatant was discarded, and 2 mL Buffer A (Promega Wizard Food Kit) was added to the remaining silicon dioxide particles and the tubes placed in a shaking incubator on their side for 10 min at 65 °C at 100 rpm. The silica particles were then separated by centrifugation for 5 min at 12,000 g. A sub-sample (1000 µL) of the resulting supernatant was processed according to the manufacturer's instructions for the Wizard Food kit (Promega) in a Kingfisher ML magnetic particle processor (Thermo Electron Corporation) and purified DNA was finally eluted into 200 µL of TE buffer.

3.3.3. Method 3 (bead-beating – University of Lincoln)

Soil was used fresh with stones removed by hand. One mL of pre-heated sterile extraction buffer (0.1 M Fe(III)Cl, 0.1 M Tris base, 0.1 M EDTA, 1.5 M NaCl, 4% SDS AND 0.5 M Sucrose) was added to a tube containing glass beads (<0.1mm and 0.5 mm) and 200 mg of soil. After bead-beating (30 secs, max speed, 5 times), samples were incubated for 10 mins at 70 °C and centrifuged at 5000 g

5 mins. 550 µl of supernatant were transferred in a new tube and placed in ice, then 35 µl of 8M Potassium Acetate were added and samples were centrifuged at 4C for 20 mins at 13.2 rpm. 500 µl of supernatant was transferred into a new tube and 250 µl 30% PEG solution was added. Samples were mixed by inverting 10 times and incubated at room temperature for 1-2 hours. Samples were centrifuged at 10000 g 15 mins and supernatants were discarded. The pellet was resuspended in 100 µl of TE 1:100. DNA was purified by using Beckman Coulter™ Agencourt AMPure XP magnetic beads. The integrity and the yield of the extracted DNA were checked by agarose gel electrophoresis, UV quantification (NanoDrop ND-1000, Germany) and fluorimetric quantitation (Quant-iT dsDNA Assay Kit, high sensitivity - Thermo Fisher Scientific).

3.3.4. Bioinformatics and statistical analysis

Sequence quality was assessed with FastQC (Andrews, 2010). QIIME version 2 was used to investigate the microbial communities (Bolyen et al., 2019). Paired end sequences were denoised by using dada2 (Callahan et al., 2017). ASVs were clustered using vsearch with an identity of 0.97 (Rognes et al., 2016). Variance-stabilising normalisation (Muletz Wolz et al., 2018) on the raw sequence counts was performed in R using CSS normalisation using metagenomSeq and phyloseq package (McMurdie & Holmes, 2013; Paulson et al., 2013; Weiss et al., 2017). The ASV table was normalised in R using phyloseq package (v 1.20.0; McMurdie & Holmes, 2013). ASVs were annotated using q2-feature-classifier plugin and gg_13_8_otus and unite_20.11.2016 databases (Bokulich et al., 2018). Animal ASVs were annotated using BLAST (Altschul et al., 1990).

All analyses were conducted in R (R Core Team, 2013), two-way full factorial permutational multivariate ANOVA (permanova) (Anderson & Walsh, 2013) on binary Jaccard dissimilarities s and PCA plots (Hellinger standardization) were calculated using the ‘vegan’ package (Oksanen et al., 2018). Indicator species of cloacal samples were identified by using the indicspecies package (ver. 1.7.1, Dufrêne & Legendre, 1997).

3.4. Extraction of total soil DNA and evaluation of the impact of management on the soil microbiome

Soil samples were collected from trials at Craibstone, Aberdeenshire, that that have managed soil pH (sampled autumn 2018) and inorganic fertiliser input for over 50 years (sampled autumn 2019). In addition, new bioinformatic approaches were evaluated to determine whether observed diversity within the DNA of microbial communities can be used to predict shifts in their biological functions. Once validated, this standardised approach was then further used to compare microbial communities in soil samples from 3 long-term trial sites comparing the effects of different organic amendments (sampled autumn 2020) and in a management trial with asparagus (sampled July 2020, for full trial details see Project 5).

3.4.1. DNA extraction

The DNA extraction method used for the taxonomic community assessment (metabarcoding) was chosen taking the findings described in Section 4.2 into account. The method uses a soil sample of 10 g (intermediate between the very large and very small samples used in that study) and DNA is extracted using the DNeasy PowerMax Soil Kit (Qiagen, Carlsbad, CA, United States) following the manufacturer's instructions. This method was optimised in Project 5 for qPCR and was selected for use throughout the project as it was rapid and reproducible, and resulted in suitable DNA quantity and quality for high throughput DNA sequencing using MiSeq technology. Soil samples were thoroughly mixed and maintained at 4 °C for up to two weeks before extraction.

3.4.2. Metabarcoding

Bacterial and fungal V4-16S and ITS1 rRNA barcodes were amplified from the total extracted DNA via PCR using 515F (Parada et al., 2016) and 806R 16S primers (Apprill *et al.*, 2015), and ITS1-F_KY02 ITS forward primer (Toju *et al.*, 2012) with a modified ITS2 (Table 3). For high throughput sequencing of the PCR products, indexed libraries were prepared using Nextera (Illumina) adapters following the manufacturer's protocol. A PhiX internal control was added to the library pool before sequencing with 600 cycles reagent kit v3 (2x300 PE) using a MiSeq sequencer (Illumina, San Diego, CA, USA). DNA from each trial was sequenced and analysed separately.

Table 2: Primer sequences used for amplifying the bacterial and fungal gene barcode regions that were subsequently sequenced using the metabarcoding approach.

16Sv4	F	515F	5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA-3'
	R	806R	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACNVGGGTWTCTAAT-3'
ITS1	F	ITS1-F_KY02	5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTAGAGGAAGTAAAAGTCGTAA-3'
	R	ITS2_Wobble	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCWGYGTTCTTCATCGATG-3'

3.4.3. Data analysis

Each of the two Craibstone trials was first analysed independently. Sequences of all PCR products from amplification of the total fungal and bacterial DNA in each soil sample were compared using the microbiome bioinformatics platform QIIME2 v. 2020.6 (Bolyen *et al.*, 2019) complemented with various R (R Core Team, 2021) statistics packages. Only 16S and ITS sequence reads >50 bp were compared. 16S reads were truncated to 253 bp but ITS reads were not trimmed. Tables of the numbers of various amplicon sequence variants (ASVs), with exact sequence matches, were compiled using DADA2 software (Callahan *et al.*, 2016).

Taxonomic assignment of each ASV was attempted using a naïve bayes classifier trained with either a 16S sequence database (release 138) from SILVA (Quast *et al.*, 2012) or a modified ITS database (v. 8.2) from Unite (Nilsson *et al.*, 2018). Only bacterial ASVs that could be recognised at phylum level with 70% confidence were further considered. Only fungal ASVs recognised with 95% confidence as belonging to the kingdoms Fungi or Stramenopila (including oomycetes) were further considered. A full account of the standardised bioinformatics procedures and subsequent statistical analyses used has been submitted for publication (Fernadez-Huarte et al., 2022).

3.5. Standardised pipeline for analysis of combined data across different trials

A standardised approach for comparison of ASV communities between soil samples was based on three main classes of soil community metrics, as defined by Morrison-Whittle *et al.* (2017), Morrison-Whittle and Goddard (2018) and Giraldo-Perez *et al.* (2021), and this comprises the core of the standardised analysis we propose (Figure 4). These were:

1. Numbers: Differences in the total number (absolute richness) of different ASVs (representing different taxa) and their distribution across sample treatments.
2. Types: Differences in the presence/absence of different types of taxa (clusters of related sequences) between samples.
3. Abundances: Differences in the relative proportions of taxa between samples inferred by sequence read counts. Where significant differences in taxa abundances are revealed, then this is followed by Indicator taxa and functional prediction analyses to highlight which taxa are significantly overrepresented in particular treatments and what functions are ascribed to these.

Differences in absolute taxa numbers were evaluated with Kruskal-Wallis tests and effect sizes calculated with E2 = H/((n2-1)/(n+1)), where n = number of observations and H is the Kruskal-Wallis H-test statistic (Tomczak & Tomczak, 2014), and distributions of counts were analysed using Shannon (Shannon, 1948) diversities. In addition, phylogeny relatedness within different taxa was assessed using Faith's phylogenetic diversity analysis (Faith, 1992).

Differences in types and abundances of taxa were evaluated using weighted and unweighted Jaccard distance matrices respectively with PERMANOVA (Anderson, 2014) with 999 permutations to determine pseudo-F ratios to calculate P and R2 estimates of effect sizes. In addition, unweighted UniFrac was used to analyse phylogenetic differences between communities, and weighted UniFrac that also takes abundances into account giving more weight to most abundant taxa (Lozupone & Knight, 2005).

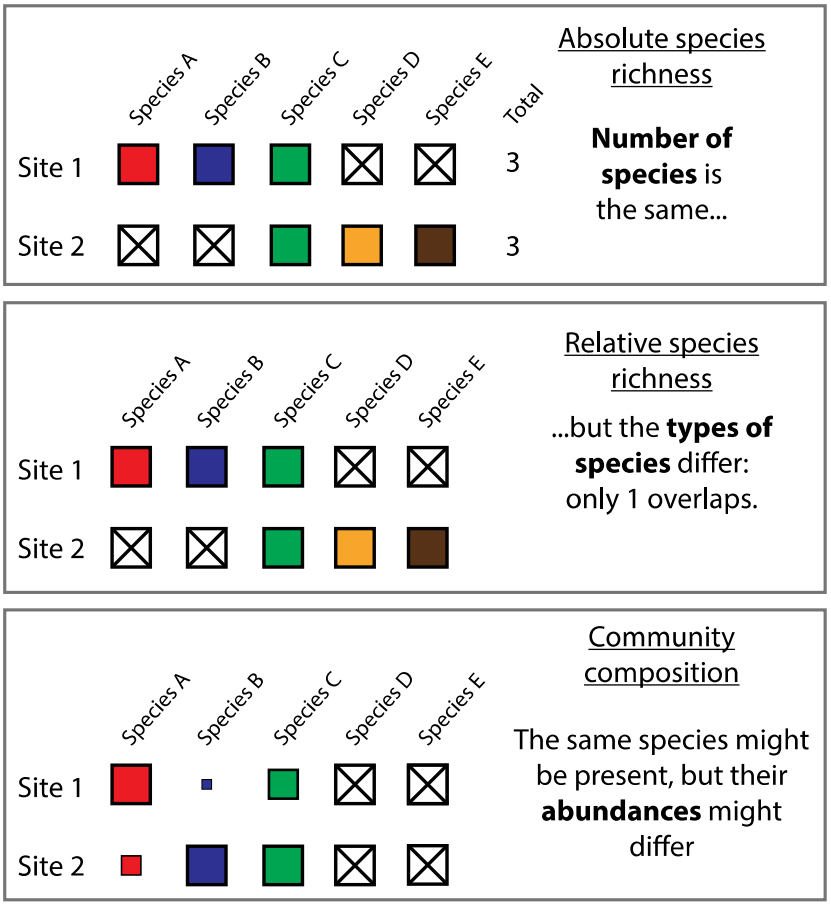


Figure 4: The three standard measures of biodiversity proposed for the soil DNA biodiversity standard analysis method following Morrison-Whittle et al. (2017), Morrison-Whittle and Goddard (2018).

To identify those individual ASVs (potential biomarkers) that were most likely to explain observed differences in diversity between soil samples, an algorithm known as LEfSe (Linear discriminant analysis Effect Size) was tested (Guo & Gao, 2021; Segata *et al.*, 2011). Further software packages were also evaluated for prediction of possible biological functions from identified ASVs. Bacterial functional predictions were estimated using the FAPROTAX tool (Louca *et al.*, 2016) on ASVs annotated using the SILVA database. For fungal functional predictions, the FunGuild tool (Nguyen *et al.*, 2016) was used on ASVs annotated using the UNITE database.

Analyses of the combined data from both trials were also performed. For comparing datasets from different trials, the sequence reads were first independently processed using DADA2, to account for potential differences between sequencing runs, prior to merging the datasets for analysis. Subsequently, the two generated tables were merged using the 'table merge' tool in Qiime2, which identifies identical ASVs between tables. To account for minor differences between ASVs from each trial, potentially resulting from differences in original orientation or trimmed length, ASVs in the

merged table were subsequently clustered into groups of greater than 98% genetic identities and termed merged ASVs (mASVs) to delineate them from ASVs. mASV representative sequences for taxonomic identification were derived by VSEARCH (Rognes *et al.*, 2016) de novo clustering approach based on 98% identity threshold. A scaling with ranked subsampling (SRS) normalisation method was then applied to the merged table (Beule & Karlovsky, 2020) to obtain an identical sample size for each data set as required in ecological analysis of taxa count. This SRS normalisation method provides a more conservative approach than the traditional rarefying approach (random subsampling without replacement).

3.6. Evaluating new biological indicators for soil health

We evaluated the outcomes from measurements made across the experimental sites, as described above, in the light of the requirements for soil health indicators (described in Project 2). The criteria used consider relevance to both agricultural production and environmental impact as well as practical aspects including sample throughput; sample storage; necessity of single or multiple visits for sampling; ease of use; ease of interpretation; sensitivity; cost; standardisation and UK availability of analysis. Measures of nematodes and microarthropod communities had scored highly during the initial logical sieve analysis and the potential of DNA-based approaches was recognised (Project 2). None of these approaches was ready for widespread deployment at the outset of the project, hence we included a further evaluation at the end of this project to consider whether these indicators could now be recommended for routine use on farm.

3.7. Correlations between microscopic observation and metabarcoding of mesofauna diversity

A second DNA metabarcoding run was performed with samples from the Craibstone pH trial to explore changes in mesofauna diversity in response to changes in soil pH levels. This work also allowed a comparison of results determined by sequencing and the direct microscopic assessment performed by Natural England (Section 3.2).

3.7.1. Metabarcoding

An animal cytochrome oxidase (CO1) barcode was amplified from the total extracted DNA using PCR primers mlCOIintF (Leray *et al.*, 2013) and HCO2198 (Vrijenhoek, 1994) (Table 4). For high throughput sequencing of the PCR products, indexed libraries were prepared using Nextera (Illumina) adapters following the manufacturer's protocol. A PhiX internal control was added to the library pool before sequencing with 600 cycles reagent kit v3 (2x300 PE) using a MiSeq sequencer (Illumina, San Diego, CA, USA). DNA from each trial was sequenced and analysed separately.

Table 4: Primer sequences used for amplifying the arthropod gene barcode regions that were subsequently sequenced using the metabarcoding approach.

CO1	F	mlCOIintF	5'- TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**TAAACTTCAGGGTGACCAAAAAATCA**-3'
	R	HCO2198R	5'- GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**GGWACWGGWTGAACWGTWTAYCCYCC**-3'

3.7.2. Data analysis

Analyses were conducted following the standardised pipeline and ASVs were assigned using a naïve bayes classifier trained with a COI database developed at Fera Science Ltd. Only arthropod ASVs recognised at phylum level with 70% confidence were considered. Arthropod ASVs were subsequently clustered by VSEARCH (Rognes *et al.*, 2016) *de novo* clustering based on 97% identity threshold to emulate taxonomic assignment at species level.

Differences in taxa numbers were evaluated with Kruskal-Wallis tests and effect sizes calculated with $E^2 = H/((n2-1)/(n+1))$, where n = number of observations and H is the Kruskal-Wallis H-test statistic (Tomczak & Tomczak, 2014). Data from metabarcoding and microscopy approaches were compared according to (a) whether pH level or cropping stage of the rotation significantly affected arthropod richness, and (b) whether the results from pairwise tests showed similar patterns of differences in richness at species level. Correlations in types of taxa were evaluated with a two-sided Mantel test (Mantel, 1967) to identify any correlation between their binary Jaccard distance matrices.

3.8. Evaluating the outputs of the visual tool to show impacts of management on the soil biological community

The development of the Visual tool was reported in Project 1. The initial management decisions the visual tool covered were: Reduced tillage, High Carbon (C) Organic Matter, High Nitrogen (N) Organic Matter and use of a Cover Crop. Additionally, options for the soil texture (Light/Sandy, Medium or Heavy/Clay), regional climate (Cold Wet, Cold Dry, Warm Wet or Warm Dry) and cropping (Arable-combinable, Arable-roots or Grass) were included to allow changes that could be more relevant to certain conditions to be highlighted (Figure 5). During the SBSH Partnership, two further management options (Increased pH (Addition of lime) or Improved Drainage) were added.

Field Conditions		
Please enter the conditions for your field		
Soil Type i.e. Light/Sandy, Medium or Heavy/Clay	Light/Sandy	
Climate i.e. Cold Wet, Cold Dry, Warm Wet or Warm Dry	Warm Dry	
Cropping i.e. Arable-combinable, Arable-roots or Grass	Grass	
Management Change		
Please enter **Management Change**	High N Organic Matter	
This can be:		
No Tillage		
Reduced Tillage		
High C Organic Matter i.e. Digestate, Biochar		
High N Organic Matter i.e. Farm Yard Manure, slurry		
Cover Crop		
Increased pH (Addition of lime)		
Improved Drainage		

Figure 5: Example of inputs sheet for Soil Management Tool.

A combination of information from published literature and expert knowledge was used to assess the outcomes for a range of variables that help the farmer or consultant understand the advantages and potential challenges for each of the management decisions. The specific management decisions chosen were those that had sufficient information from the literature to inform the management outcomes. The research papers that were used to define the outcomes were used in a database to allow those interested to read the papers if they wanted more data and the reasoning behind the consequences of a management change.

Additional information was included as 'pop-up' screens for the Soil Type and Climate to help in clarifying the extent of these (Figure 6).

Soil Types

Soil Type	Notes	Risk of Compaction
Light (sandy soil) Low clay content less than 18%	Freely drained and less organic material Low aggregate stability	Less prone to compaction due t better drainage. If compacted sandy soils are often slow to recover, mechanical interventior may be necessary
Medium (silty/loams) Clay content between 19 and 35% in the topsoil	Can be slow draining Greater aggregate stability Prone to capping	Susceptible to compaction wher wet
Heavy (clay) Clay content greater than 35%	Poor natural drainage Remain wet longer Variable porosity	High chance of soil compaction when wet, although naturally crack in dry and frost conditions helping to alleviate compaction

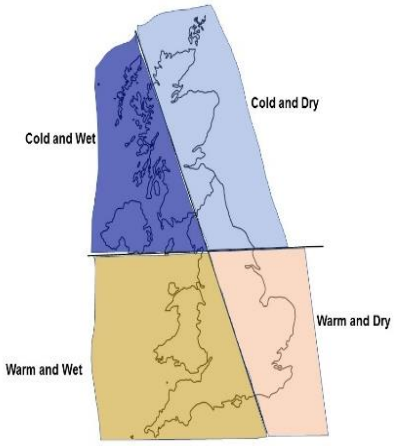


Figure 6: a) Soil Type/Texture options b) UK Climate areas.

The effects of the agricultural management options (management variables) were modified by agronomic conditions across the UK. Expert knowledge of the underlying principles of soil/plant/organism/environment interactions was used to interpolate and extend the literature review findings, to give a full coverage for UK agricultural systems and to consider how any general findings would be adjusted as a result of soil/ region specific nuances. Thus, the outcomes included yield, soil biota (positive biology and slugs), disease, weeds (plus the reduction of herbicide use), chemical parameters of pH, N, P, K, cation exchange capacity, SOM and physical parameters of water infiltration, soil structure and trafficability. The effects of the management options (i.e., positive or of concern) was further modified by regional climate, soil type and farming system prevalent in the UK.

The expert opinion of the project partners was used to apply modifications of how the effect of each driver was moderated or exaggerated by:

- soil type (simplified to light (sandy or light silty), medium (loams) or heavy (clay-rich),
- regional climate (simplified to cold & wet; cold & dry; warm & wet; warm & dry)
- main agricultural systems (arable - combinable, arable including root crops, grassland

Data tables were developed for the various possible scenarios and delivered outcomes from a management change that were coloured, for visual ease, to highlight a positive or a challenge from the new management (Figure 7).

Figure 7: Example of an outcome display for an option to incorporate High N Organic Matter on a Light/Sandy soil texture, in a Warm Dry regional climate and a Grassland system.

A summary of the advantages and challenges was provided along with a selection of the references from the literature. An estimation of effect of profit margins was also introduced for the short-term (1 year) and the longer-term (5 years).

4. Results

4.1. Short-list targets for analysis for DNA-based approaches

A systematic review for DNA extraction methods was carried out in conjunction with Project 5 as the outputs were inherently relevant to both projects. In summary, the review highlighted changes in procedures, optimised over time by the scientific community, that are most likely to accurately quantify the DNA of target organisms in soil. Key elements found essential for method standardisation were identified in Project 5 and can be summarised as: sample size; sample storage; sample homogenisation; cell lysis; and DNA purification.

Soil contains a wide range of biological organisms, with a wide range of potential targets for DNA analyses. The consideration of each group in terms of their biological function and community composition in soil is explained in more detail in Project 3. Since bacteria and fungi are key components of the soil that make up a significant proportion of the microorganism biomass (Murphy et al. 2003), and nematodes and annelids comprise the majority of the mesofauna biomass (Wallwork, 1976), they were selected for detection. To enable molecular detection, three different amplicon primer sets were defined to provide the broadest taxonomic coverage, targeting ubiquitous DNA sequencing in each group: ribosomal 16S for bacteria covering variable regions V4; the intergenic ribosomal region for fungi and some oomycetes, ITS1; the animal cytochrome oxidase subunit 1 gene, COI. It is acknowledged that these targets do not provide 100 % coverage for each group, but each are well accepted by the microbiome research communities to describe taxonomic compositions.

4.2. Comparison of soil DNA extraction methods

Extraction of DNA from soil was shown to be improved by the inclusion of a brief bead-beating step (Method 1a), compared with extraction in phosphate buffer alone (Method 1). This allowed recovery of DNA reproducibly from this sandy loam soil (12 % clay). Inclusion of the additional bead-beating step would potentially release some cellular DNA and so the final extract is not expected to be entirely composed of only extracellular eDNA. However, Method 1 was not reproducible, and did not yield suitable DNA in quantity and quality for high throughput DNA sequencing. There was no significant difference between the amount of DNA extracted by Method 1a, Method 2 and Method 3 (ANOVA, $F= 0.628$, $P = 0.432$). Specifically, DNA concentration was, 40.21 ± 6.64 ng/µL for Method 1a, 23.67 ± 11.61 ng/µL for method 2 and 36.65 ± 14.95 ng/µL for Method 3. However, Method 2 was time consuming because the first step (*i.e.* matrix disruption using the ball mill and incubation in phosphate buffer) was only possible for four samples at a time, thus extending the processing time. This method is not rapid enough for routine use.

The total number of bacterial, fungal and animal species identified were 4,831, 1,516 and 3,442 respectively. Comparison of extraction methods showed no effect on the number of species identified; however, extraction methods significantly affected the relative species richness and community composition.

For bacterial, Method 1a registered a significantly higher abundance of phyla FCPU426, Planctomycetes, OD1 and WS2 and significantly less abundant Cyanobacteria and Proteobacteria. Method 2 registered a higher abundance of phyla Bacteroidetes, Chlorobi, Cyanobacteria, Elusimicrobia, Fibrobacteres, Gemmatimonadetes, Proteobacteria, Spirochaetes, WS2 and WS6. but Actinobacteria, as well as Chloroflexi were significantly less abundant when compared with both Methods 1a and 3. Method 3 registered a higher abundance of Acidobacteria, Actinobacteria, Chlamydiae, Crenarchaeota, Euryarchaeota, Firmicutes and TM6. Armatimonadetes, OD1, Spirochaetes and WS3 were significantly less represented in Method 3 when compared with Methods 1a and 2. These differences are summarised in Figure 8a.

When considering fungal communities, Method 1a revealed a significantly higher abundance of Basidiomycota, Glomeromycota and Zygomycota, and a significantly lower abundance of Ascomycota. Method 2 registered significantly more abundant Basidiomycota and Glomeromycota and significantly lower abundance of Chytridiomycota, Rozellomycota and Zygomycota. Method 3 identified a significantly higher number of Rozellomycota and Zygomycota and significantly less abundant Basidiomycota and Glomeromycota. These differences are summarised in Figure 8b.

Extraction method had a larger impact on the nature of the estimated biological community than the range of long-term organic amendment treatments tested at the Harper Adams site (Figure 9). Therefore, standardisation of DNA extraction methods is important when comparing samples collected from different locations and at different times. It is likely that the different extraction approaches target different parts of the community.

Where the outputs from each extraction method were combined, either by pooling the extracted DNA by mixing the samples in equimolar concentrations (followed by a single sequencing step) or by pooling the sequencing results using bioinformatics, the additive results from combined methods are more consistent, this was also found when using quantitative PCR approaches as reported in Project 5. The total number of species recorded in the pooled samples was significantly higher than the one recorded in the samples extracted with the three different methods (data not shown). However, the pooled data also showed no of organic amendment treatment on the total number of species. The treatments had a small effect on relative species richness and there was no significant change in community composition identified. A summary of this work was published as a Case Study in 2019.

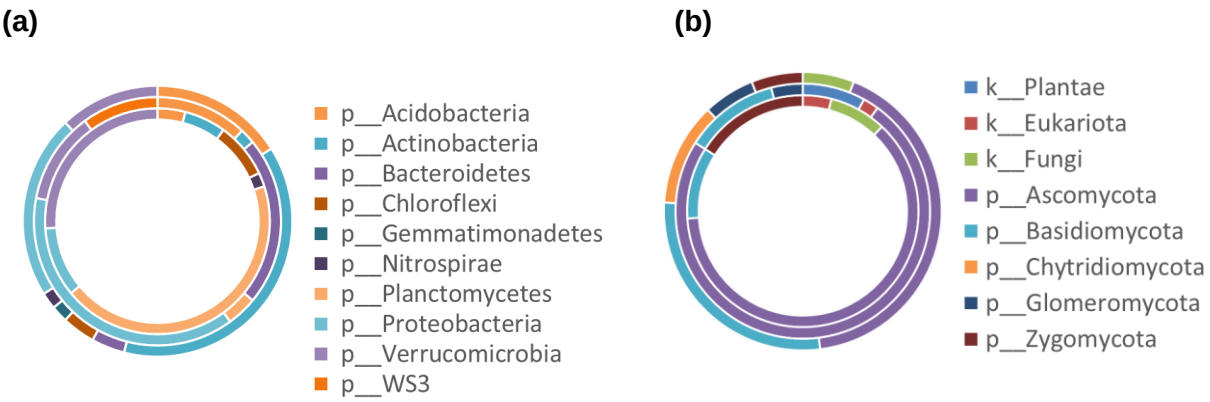


Figure 8: Relative proportions of the most abundant types of bacteria (a) and fungi (b) estimated using three different DNA extraction methods. Estimates obtained using methods 1, 2 and 3 are shown from the inner to outer rings respectively. The coloured bars in each ring represent the proportion of the total community for each type of bacterium or fungus. K – kingdom, P – phylum.

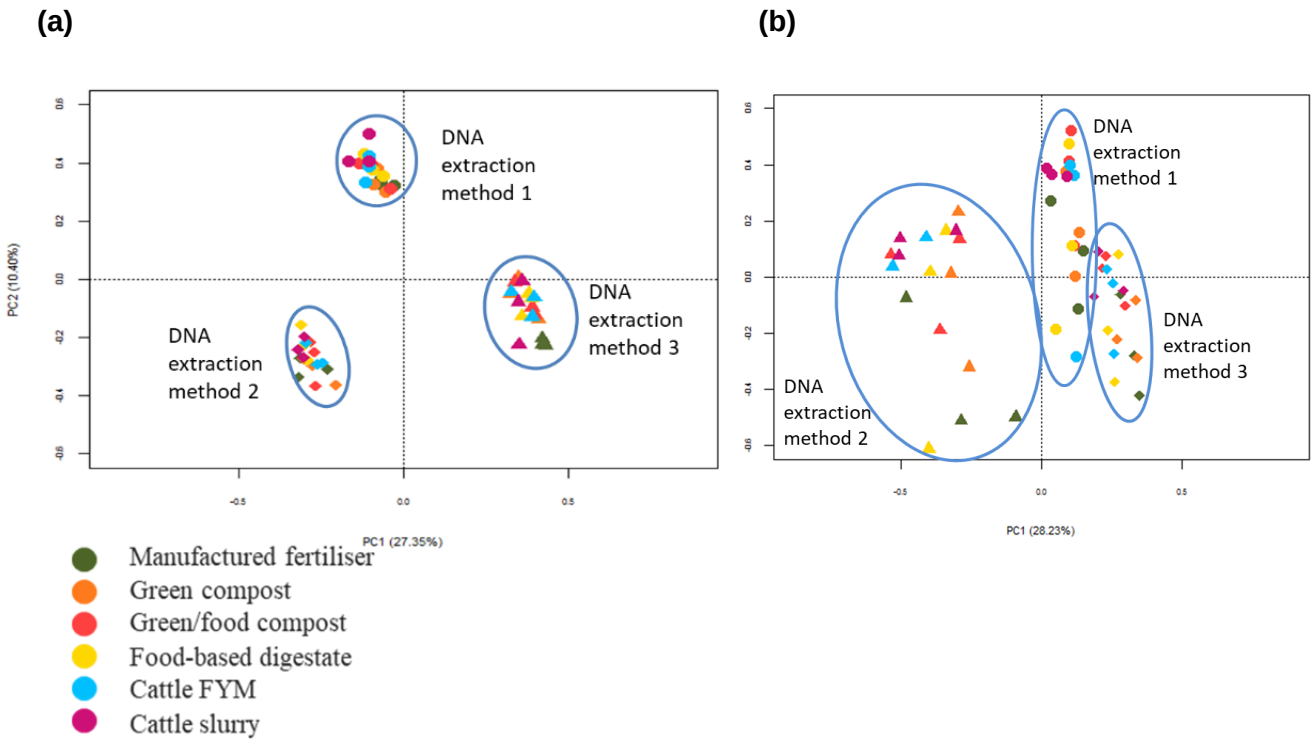


Figure 9: Principal component analysis comparing relative abundances of different (a) bacteria and (b) fungi in soils treated with different organic amendments, estimated by metabarcoding with three different DNA extraction methods.

4.3. Evaluation of the effects of different soil management practices on the soil biological community and its function

4.3.1. Effects on nematode communities

Effect of long-term pH manipulations

At Craibstone, the average structure index by crop type was highest in the grass ley of the pH experiment (P<0.05), with no difference in enrichment scores due to crop type (Table 4). pH had no effect on the structure index, but enrichment indices were lower at pH 7.5 (P=0.07), although this varied with crop type; the effect was apparent in all but the wheat crop. The majority of soils were considered to have a nematode community structure close to the idealised optimum (scores of 1 or 2), except in the grass ley at pH 6.5 (due to a high enrichment index) and in the oat and potato crops at pH 4.5 (high enrichment index, as well as a high structure index in the oat crop).

Table 4: Structure (SI) and enrichment (EI) indices at the Craibstone experimental sites and their associated scores colour coded according to the proposed scorecard traffic lights: green – continue rotational monitoring, amber – review, red – investigate.

Treatment	**Grass**			**Oats**			**Potatoes**			**Wheat**		
	SI	**EI**	**Score**	**SI**	**EI**	**Score**	**SI**	**EI**	**Score**	**SI**	**EI**	**Score**
pH experiment												
pH 4.5	82	31	2	92	85	4	63	74	3	65	43	1
pH 6.0	84	60	2	47	58	2	52	38	1	67	50	1
pH 6.5	76	68	3	62	54	2	45	43	2	61	41	1
pH 7.5	80	17	2	65	43	1	60	24	2	42	54	2
P[1] Crop	*0.01*	*NS*										
P[1] pH	*NS*	*0.07*										
P[1] Crop x pH	*NS*	*0.07*										
Old Rotation												
Treatment	**Grass**			**Oats**			**Potatoes**			**Barley**		
	SI	**EI**	**Score**	**SI**	**EI**	**Score**	**SI**	**EI**	**Score**	**SI**	**EI**	**Score**
No Fertiliser	39	0	3	79	0	3	60	6	3	58	0	3
NPK	63	17	2	64	0	3	61	0	3	58	0	3
P[2]	NS	0.05										

[1]Statistical analysis undertaken using a two factor ANOVA; there were three replicates of each treatment; *NS:* No significant difference (P>0.05).
[2]Statistical analysis undertaken using a single factor ANOVA; there were three replicates of each treatment; *NS:* No significant difference (P>0.05).

Effect of long-term fertiliser application

The adjacent old rotation experiment at Craibstone (with the same soil type but sampled a year after the pH experiment) had a very different nematode community structure, particularly as the species

that contribute to the enrichment index were absent from a lot of the treatments (*P*=0.05; Table 4). As a result, the majority of the soils were given an intermediate score for soil biological health.

Effects of organic amendments

As observed with most of the other soil health assessments from the long-term experiments (as reported in Project 4), 'site' differences dominated, with relatively few differences between treatments at an individual site. This can be clearly seen at the long-term organic amendment sites (Table 5 & Figure 10). The highest structure and enrichment indices were observed in the grass ley at Harper Adams in 2017, compared with the same site after 3 years of arable cultivation in 2020, and the lowest indices were observed on the light textured continuous arable soils at Gleadthorpe in 2019. There was no consistent difference between the long-term organic amendment treatments at these sites, although the control (no organic amendment) and slurry treatments at Harper Adams in 2017 (grass ley) were allocated a 'poorer' soil biological health score than the other treatments due to a very high structure index and in the case of the slurry treatment a high enrichment index also (Table 5). However, this was not the case in 2020. At Gleadthorpe, the control, green compost and broiler litter treatments had higher scores due to low enrichment indices on the control and green compost, and a low structure index on the broiler litter treatment, although these differences were not statistically significant. The FYM and compost treatments at Terrington were also allocated a poorer score for biological health based on the nematode data than the control, due to lower enrichment indices on these treatments (Table 5; *P*<0.05 for the compost treatment).

Table 5: Structure (SI) and enrichment (EI) indices at the long-term organic amendment experimental sites and their associated scores colour coded according to the proposed scorecard traffic lights: green – continue rotational monitoring, amber – review, red – investigate.

Site/Treatment	Harper 2017 (grass)			Harper 2020 (cereal stubble)			Gleadthorpe 2019 (cereal stubble)			Terrington 2020 (cereal stubble)		
	SI	**EI**	**Score**	**SI**	**EI**	**Score**	**SI**	**EI**	**Score**	**SI**	**EI**	**Score**
Control	92.1	43.3	3	41.2	44.0	2	53.1	0	3	37.6	28.2	2
FYM	80.5	62.8	2	52.9	39.3	1	55.8	13.6	2	34.7	20.1	3
Slurry	91.9	73.1	3	nd	nd	nd	52.7	19.5	2	nd	nd	nd
Green compost	78.3	57.2	2	73.1*	47.4	1	56.7	2.3	3	53.9	8.8*	3
Green/food compost	86.5	47.6	2	nd	nd	nd	n/a	n/a	n/a	nd	nd	nd
Food-based digestate	82.8	59.2	2	nd	nd	nd	n/a	n/a	n/a	nd	nd	nd
Broiler litter	n/a	n/a	n/a	n/a	n/a	n/a	31.9	19.6	3	n/a	n/a	n/a
P[1]	*NS*	*NS*		0.005	*NS*		*NS*	*NS*		*NS*	0.002	

[1]Statistical analysis undertaken using ANOVA; there were three replicates of each treatment; * shows significantly different from control; *NS:* No significant difference (*P*>0.05).

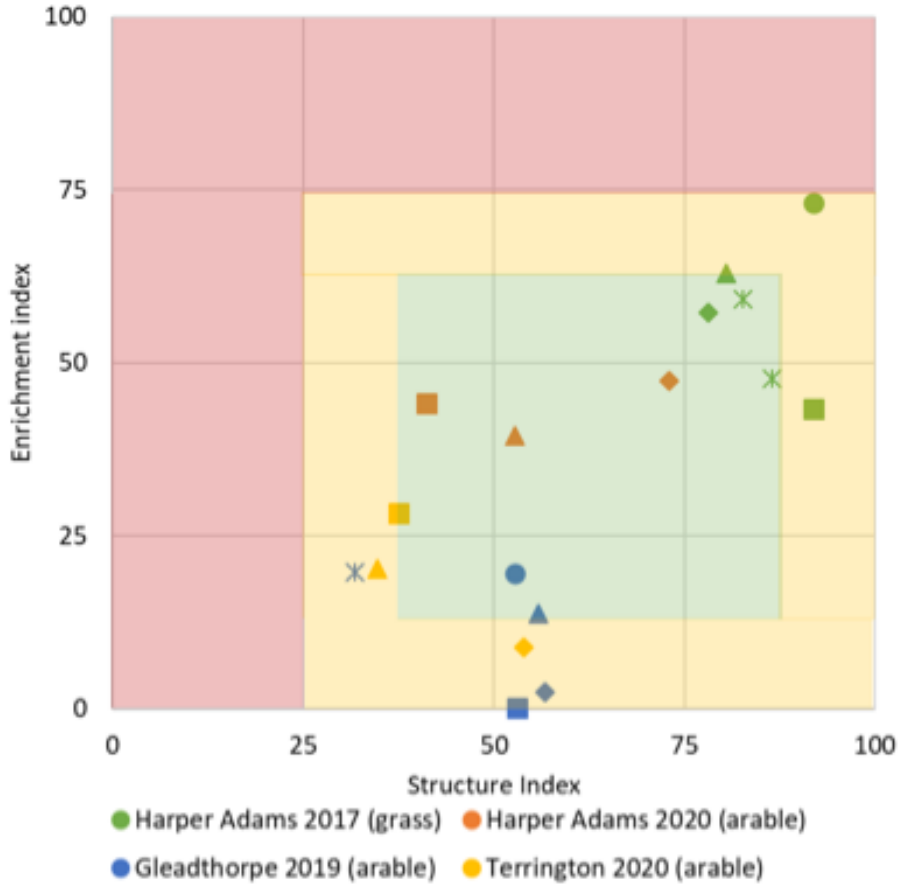


Figure 10: Enrichment and structure indices at the long-term organic amendment experimental sites. The position of the index values on the grid was scored and coded using traffic lights: green area of the grid – healthy biological community, close to the idealised optimum composition; amber – intermediate biological health, red – least healthy community, either highly enriched, disturbed or stressed community. Colours of the symbols indicate the site; different symbols represent the organic amendment treatments (squares: control; triangles: FYM; diamonds: green compost; circles: slurry; stars: other treatments e.g. digestate or green/food compost at Harper Adams).

4.3.2. Effects on mesofauna communities

Effect of long-term pH manipulations with taxonomic detail

The soil mesofauna samples from the pH trials at Craibstone were identified to the highest mesofauna taxonomic level possible and the community was characterised in terms of numbers of individuals of each taxa. Some samples, particularly those from the under sown oat crop, contained very large numbers of above-ground dwelling herbivorous mites. Because the project was focused on exploration of the links between soil mesofauna communities and soil conditions, these herbivorous mites (in the prostigmata families, Phytoptidae, Siteroptidae, and Tarsonemidae) were

excluded from the data. Removal of the herbivorous mites resulted in better separation between treatments in the NDMS plots (Figure 11). NMDS distributions of community samples from under potatoes showed no overlap with those taken from under winter wheat, when herbivorous mites were removed, but considerable overlap when these were included in the data. The samples taken from the under sown oats areas were more similar to the grass treatment when the herbivores were removed from the data, probably reflecting the association between *Siteroptidae* mites and the oats on which they are known to feed.

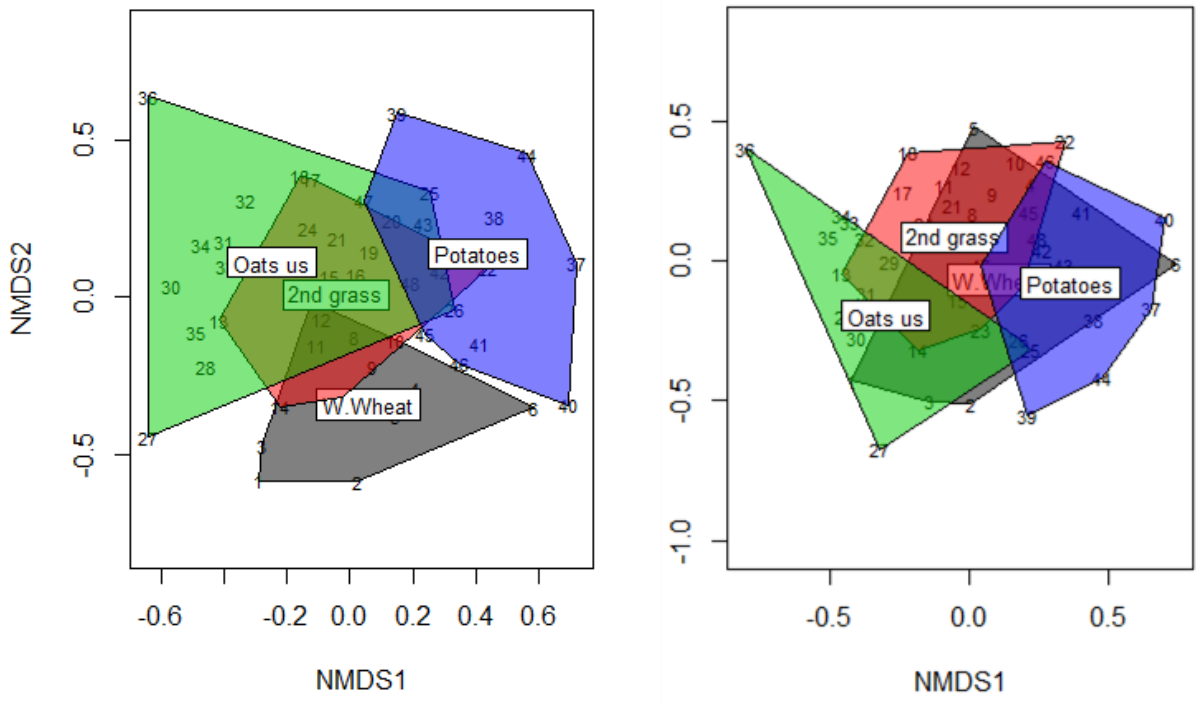


Figure 11: NMDS analysis axis loading of soil mesofauna communities grouped at family level, from Craibstone 2018 sampling showing the influence of crop type, without herbivore mite families included (left) or including all families (right).

Removal of the herbivores from the dataset when the NDMS factors 1 and 2 for mesofauna communities were grouped by pH treatment (Figure 12) had the effect of making the mesofauna communities more similar in terms of their response to pH. While there was relatively good separation between the pH 4.5 treatments and the more neutral treatments (pH 6 - 7.5) when herbivore mites were included, this effect was absent when herbivorous mites were removed. This suggests that one impact of low pH may be to make the crops weaker and thereby more susceptible to herbivory.

There was a significant difference between mesofauna community species richness, population and diversity due to crop type, but not pH (Figure 13). Because the crop treatments at Craibstone were rotational, they effectively provide true replicates for the pH treatment, while samples taken within crop treatments (which are single blocks) are pseudoreplicates. Bulked data for mesofauna at family level, without above-ground plant mites, gave broadly overlapping ranges on the 2 main axes, with no clear separation of communities evident due to pH (Figure 14). The ANOVA was repeated using this bulked data, but no significant effects of pH on diversity, population or richness were observed.

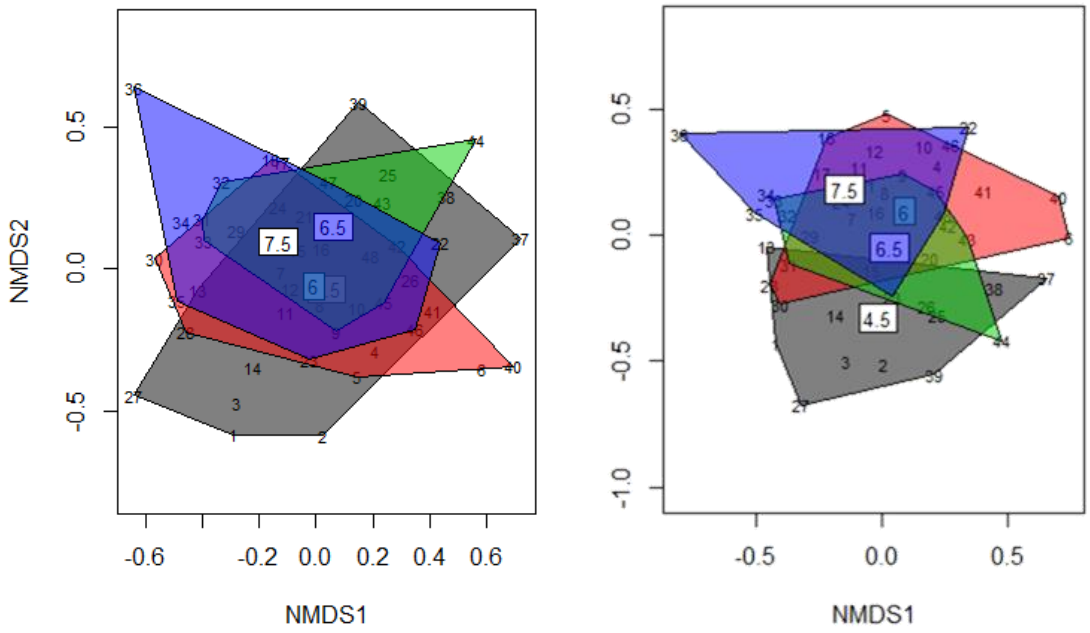


Figure 12: NMDS analysis axis loading of soil mesofauna communities, grouped at family level, from Craibstone 2018 sampling showing impact of removing herbivore mite families (left) or including all families (right) on the influence of pH treatment.

The diversity, population, richness and loadings of each sample were compared to other soil parameters, calculated as averages for each crop/pH combination. Notably, total soil mesofauna catch (without plant mites) correlated positively with clay content, bulk density, penetration resistance and with total earthworms, adults, epigeics and anecics. Mesofauna richness (at family level) correlated positively with bulk density, total earthworms, juveniles and endogeics, whilst mesofauna diversity (at family level) correlated positively with VESS score. These changes in soil physical properties are also likely to be correlated with a more favourable moisture regime for soil mesofauna in this sandy loam site.

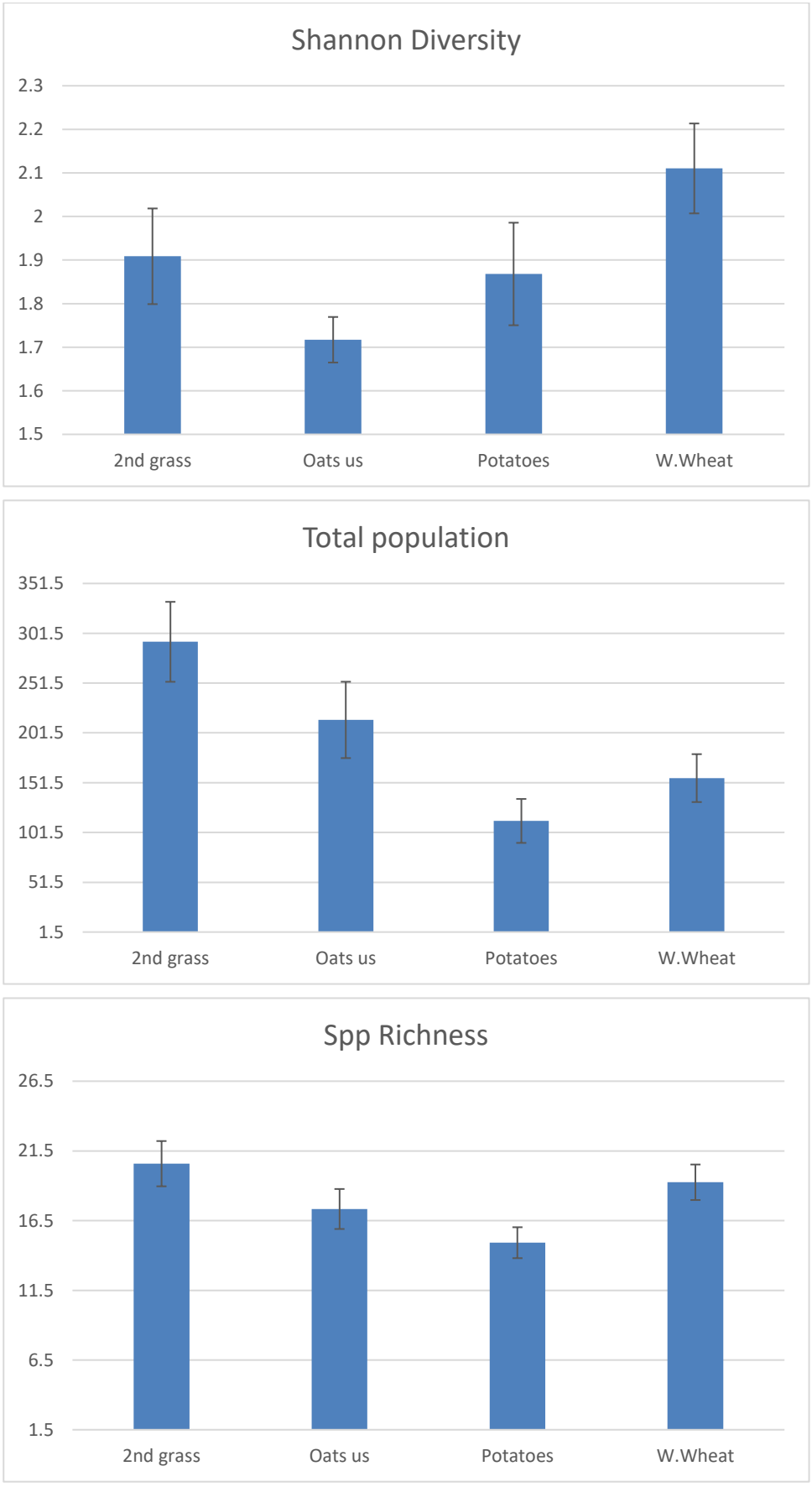


Figure 13: Effect of crop type on mesofauna community metric, shown as mean ±SD. Treatment was found to have a significant effect on all three metrics ($P<0.05$),

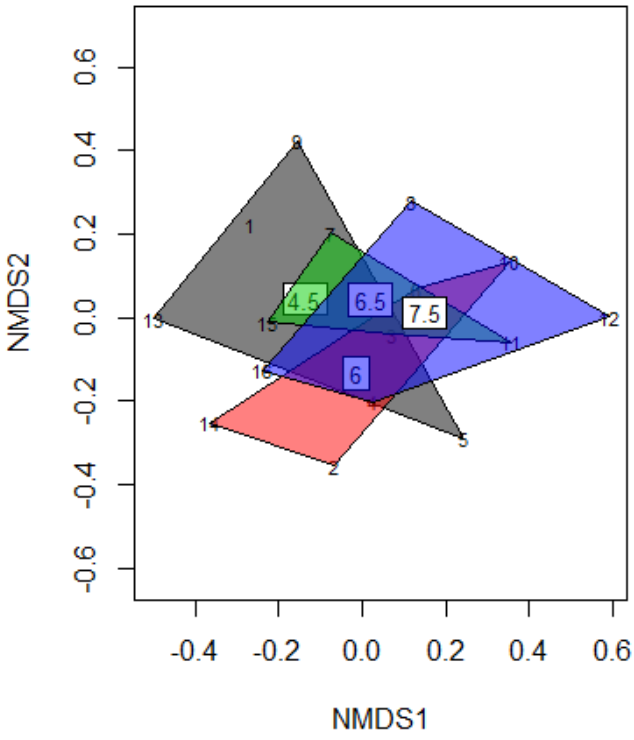


Figure 14: Distribution of mesofauna communities from Craibstone where all data from single plots were bulked, and plots with different crops (at different stages in the same rotation) used as true replicates for the different pH treatments.

Cross site analysis, including the effects of organic amendments & tillage

NDMS analysis of all mesofauna data from all sites and periods sampled (including any herbivores, Figure 15) showed a separation of the communities at Harper Adams in 2017, when the site was sampled under a grass-clover ley. All other sites/sampling periods showed some overlap between sites, with the community characteristics of Terrington 2020 overlapping with Harper Adams 2020, which overlapped with Gleadthorpe 2020 which overlapped with Craibstone. There were only 2 samples from Loddington, but these were reasonably close to the Craibstone samples.

Samples from Harper Adams in 2017 had communities ~5 times larger than the sites sampled in 2020, and other samples from 2017 were also larger (Figure 16). This is reflected by the strong Spearman's rank correlation between NMDS1 and log-transformed total soil mesofauna population density ($P<0.00001$), but this axis correlated even more strongly with total numbers of Entomobryomorpha and Poduromorpha springtails ($P<0.0000001$) (hereafter referred to as NEPS).

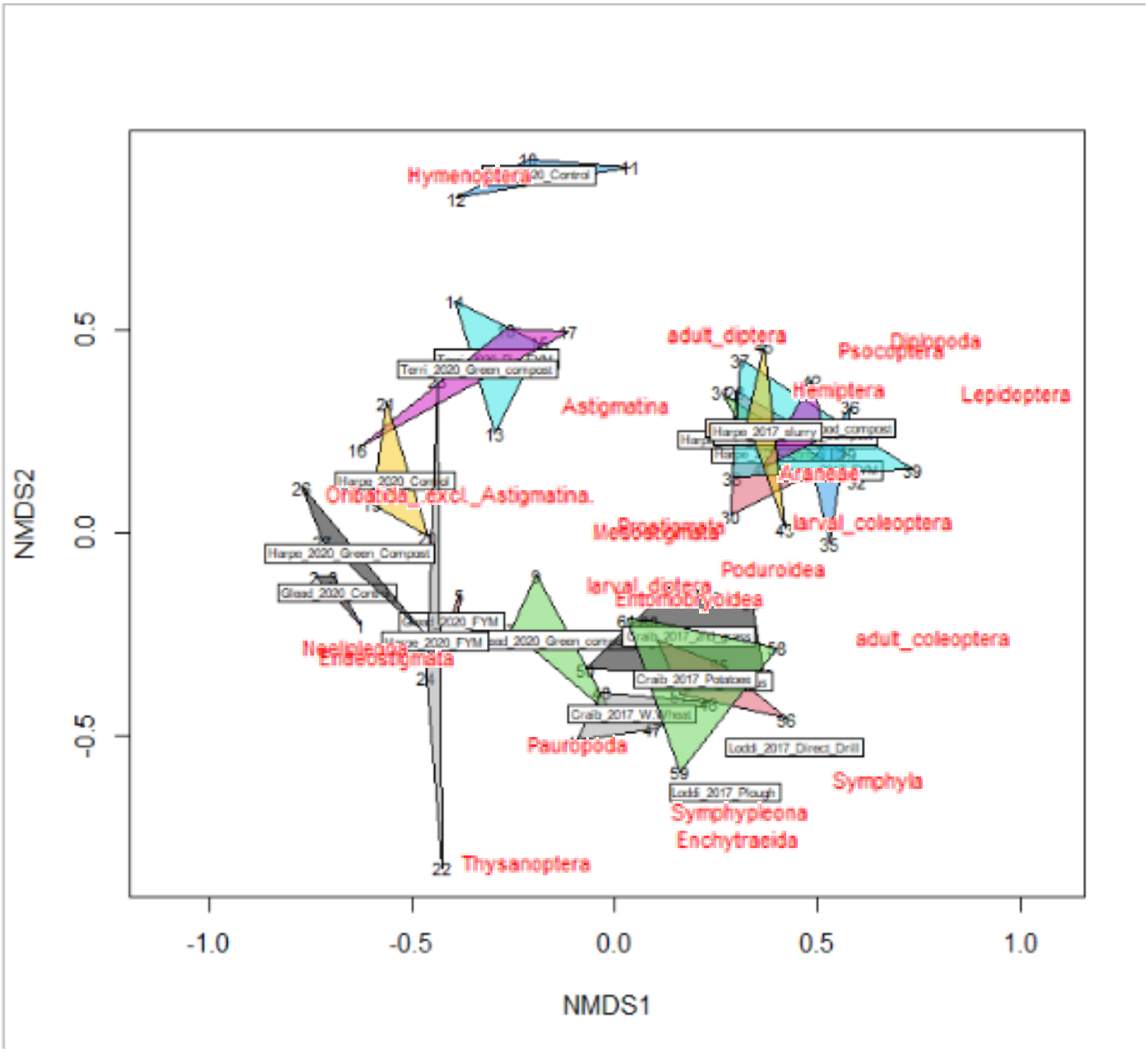


Figure 15: NMDS plot of the full mesofauna dataset from all sites and sampling times. Bounded areas show samples taken from the same sites, years and treatments, with a label identifying these at their centroid. Text in red indicates how the taxa are associated with the axes (note that Chilopoda are associated with high scores of NMDS2 (~1) and mid (zero) scores of NMDS1, and are not shown.

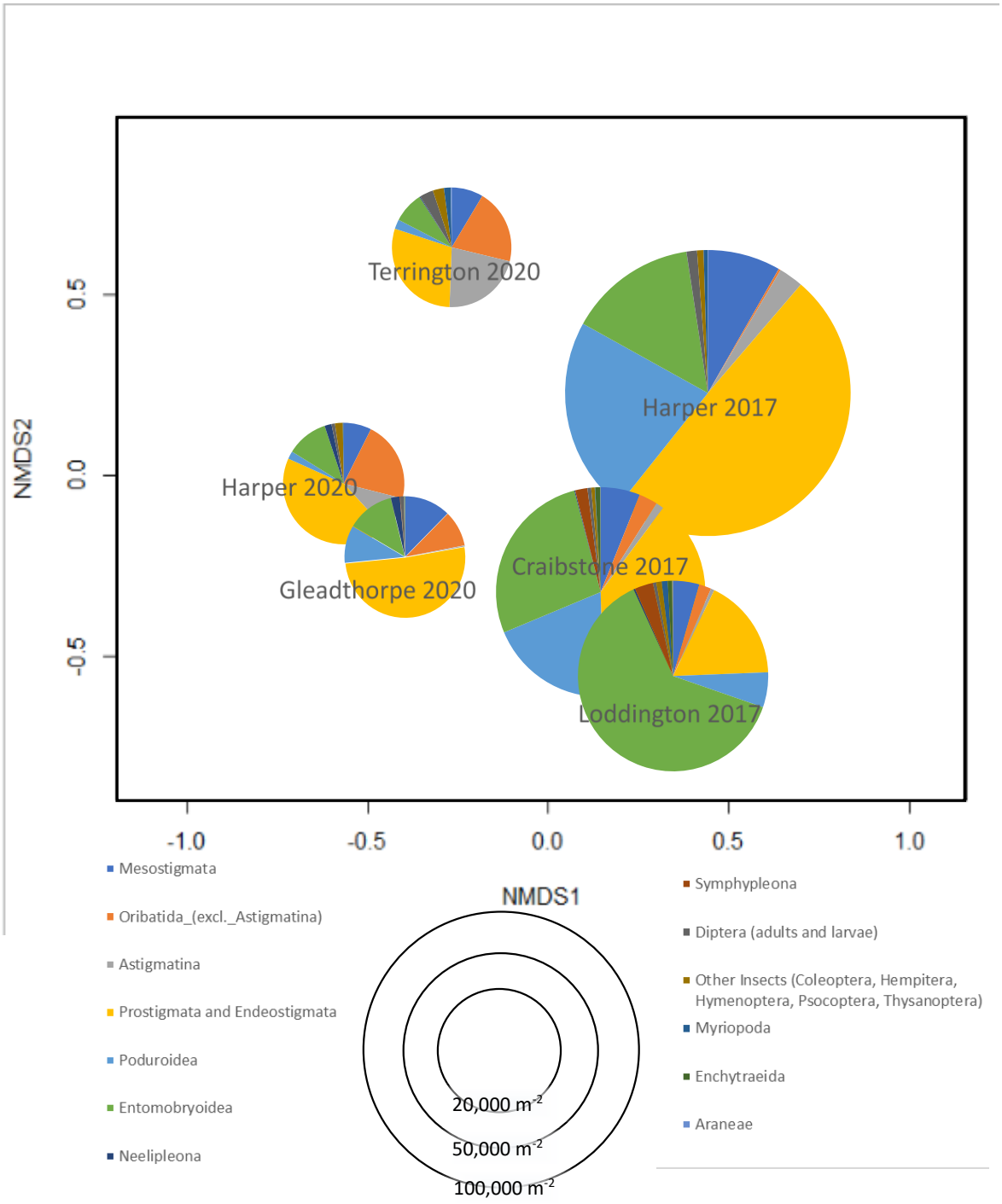


Figure 16: Pie charts showing proportions of different groups of mesofauna found at each site and sampling year. The relative areas of the charts show the total population of mesofauna identified per square metre and their location on the chart shows similarities in community composition, using the approximate centres of each site as characterised using NMDS axes.

Control, Green compost and FYM treatments at Harper Adams in 2020 gave distinct mesofauna communities, and all three treatments at Gleadthorpe 2020 showed strong separation from each other; data from an NMDS using the Gleadthorpe mesofauna data at the most detailed level of taxonomy are shown in Figure 17. Mesofauna communities where soils were amended with Green Compost or FYM treatments at Terrington were also strongly different from those under the control treatment. A stronger response to treatments is seen at Gleadthorpe in comparsion with Harper Adams and Terrington. In contrast communities under the wide range of organic matter treatments sampled in at Harper Adams in 2017 and under different crops or pH groupings at Craibstone showed similar, overlapping, community characteristics.

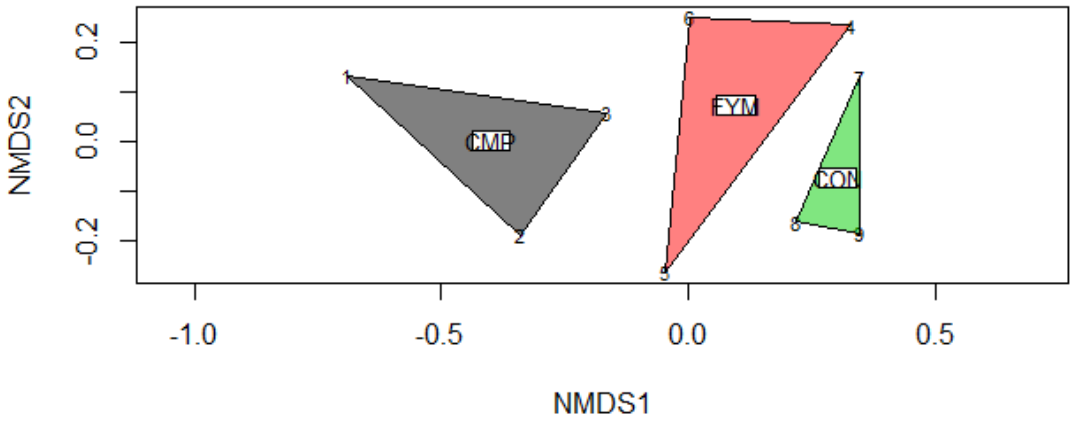


Figure 17: Scores for compost (CMP), farmyard manure (FYM) and control (no addition, CON) treatments at Gleadthorpe on the 2 axes produced by NMDS analysis of soil mesofauna communities, identified to the most detailed level of taxonomy available.

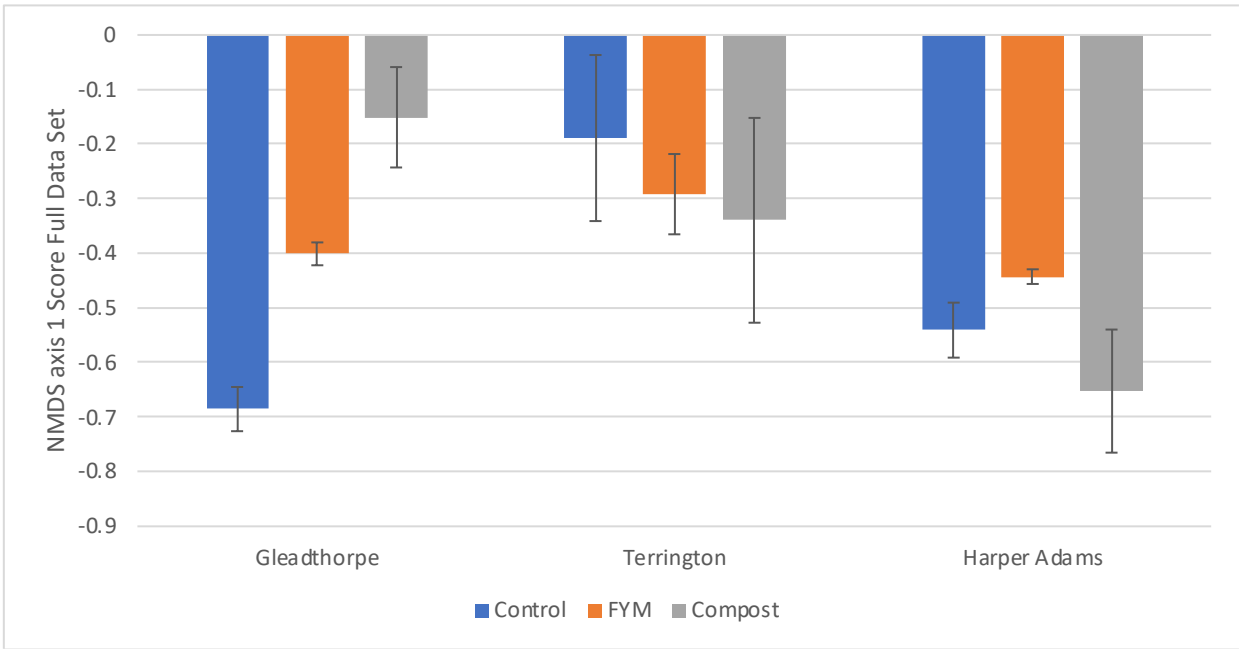
Sample scores on the overall NMDS axes were averaged for each combination of site treatment and year (Figure 18). For the full dataset, including all mesofauna, scores on axis 1 were found to have very highly significant (P<0.00143) correlations, using Spearman's Rank, with:

- +ve, CO_2-C burst (p<0.0000001)
- +ve, Maximum penetration resistance (p<0.000001)
- -ve, Root lesion nematodes (p=0.0000001)
- -ve, VESS (p<0.0001)
- -ve, Total nematodes (p= 0.0005)

NMDS axis 2 showed significant (P<0.00143) positive Spearman's rank correlations with:

- +ve, Bulk density (p= 0.00027)
- +ve, Nematode structure index (p= 0.00076)
- -ve, Soil water content (p= 0.0011)

a)



b)

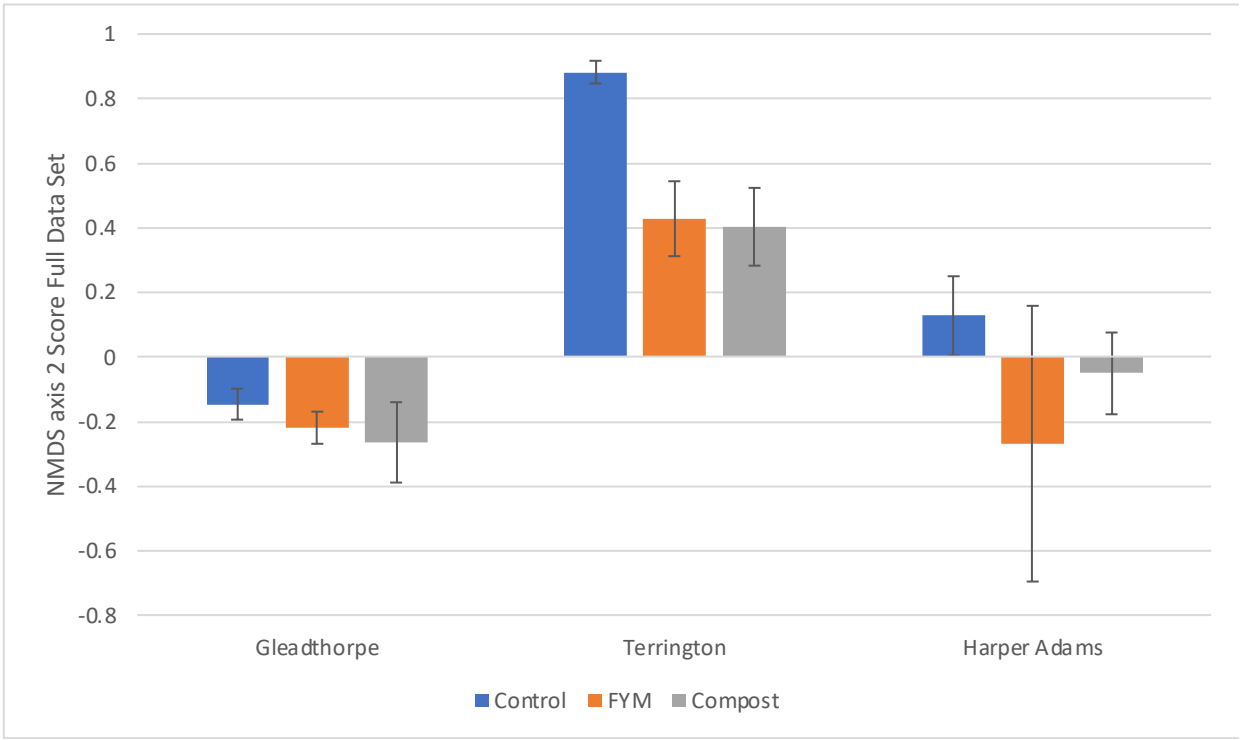


Figure 18: Scores on NMDS axis 1 (a) and NMDS axis 2 (b) based on analysis of the full soil mesofauna community across all experimental sites. Graph shows score for 3 sites sampled in 2020, receiving treatments of either farmyard manure (pig manure at Terrington), green compost or no treatment (control). Bars show standard error (n=3).

Analysis of the reduced dataset containing only springtails and mites, and identified to broad taxa, showed more commonality between sites. A little over 97% of the specimens identified in all sites and sampling times were acari or collembola, with 57.1 percent of these being acari. Because of the overwhelming numerical dominance of the acari and collembola in these samples, the pattern of sample sizes (Figure 19) was similar to that seen in the full data set (Figure 16).

Again, there were strong separation between sites. For Harper Adams, the 2017 mesofauna communities were distinct from those sampled in 2020, one exception was a plot from Harper Adams sampled in 2020 receiving farmyard manure which resulted in this treatment strongly overlapping the 2017 range for this site. These data showed similar treatment effects in response to organic amendments at Harper Adams, Gleadthorpe and Terrington. At Craibstone, this analysis separated more strongly, with communities under potatoes being fully distinct from wheat and oats, while the 2nd year grass (following oat under sowing) overlapped communities associated with all other treatments.

In the reduced dataset containing only acari and collembola, significant ($P<0.00143$) correlations, using Spearman's Rank, were found between NMDS axis 1 and:

- +ve, CO_2-C burst Spearman $p<0.0000001$
- +ve, Max penetration resistance $p<0.00001$
- -ve, VESS $p=0.0006$
- -ve, Root lesion nematodes $p=0.00001$

NMDS axis 2 showed significant ($P<0.00143$) positive correlations with:

- Stunt/spiral nematodes

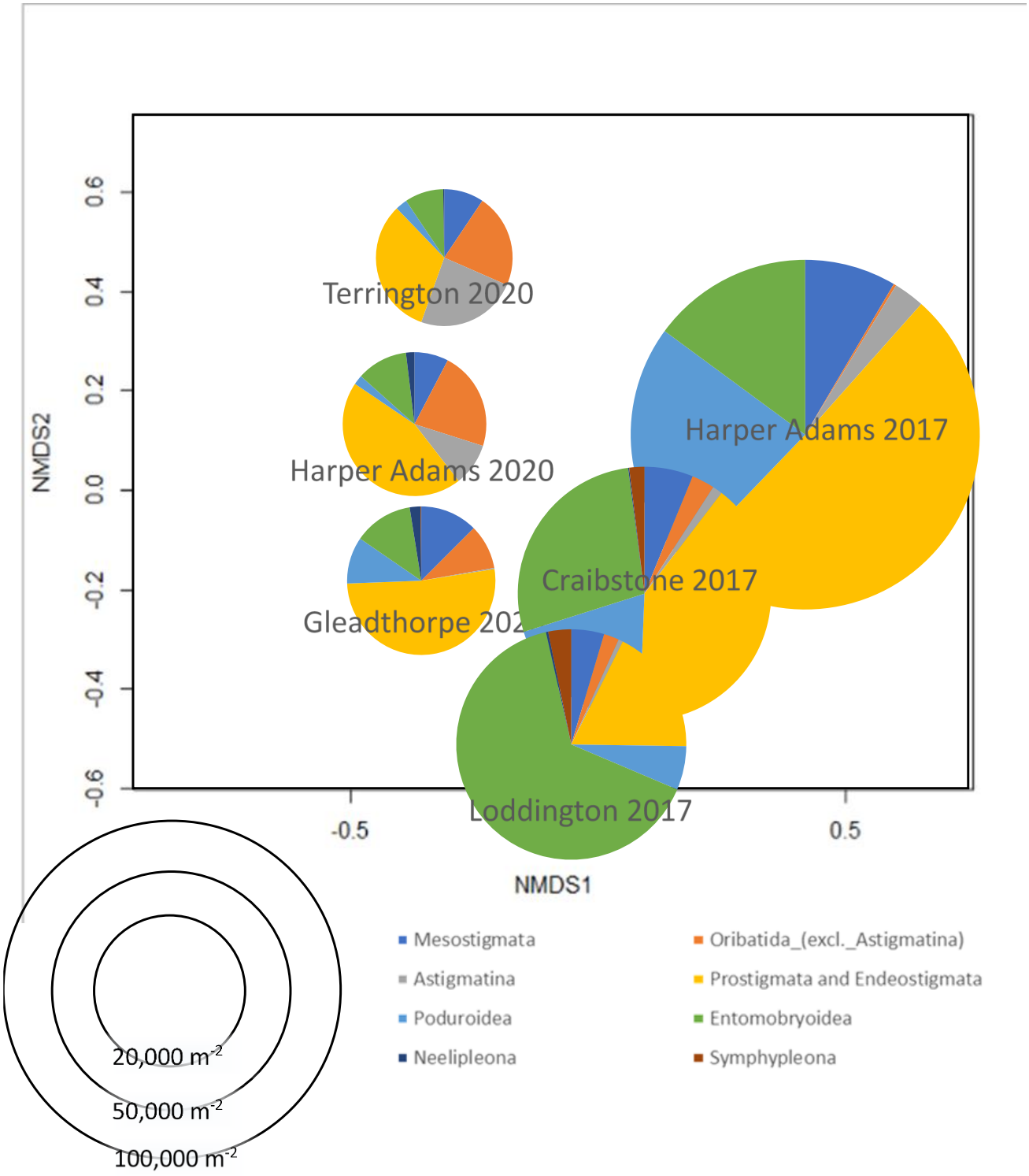


Figure 19: Pie charts showing proportions of different groups of collembola and acari major groups found at each site and sampling year. The relative areas of the charts show the total population of individuals identified per square metre and their location on the chart shows similarities in community composition, using the approximate centres of each site as characterised using NMDS axes.

Exploring simpler mesofauna community indicators

One problem with analysis of data using NMDS is that the axes generated are unique to the datasets used to generate them, and therefore to interpret new observations against this framework, a new analysis would be needed. The datasets were explored to identify if simple directly measurable indicators could be used as a proxy for the NMDS axes.

It was found that the total population density (n m^{-2}) of Entomobryomorpha and Poduromorpha springtails (hereafter referred to as NEPS) was strongly correlated with NMDS axis 1 and its use as a proxy indicator for the effects observed on this axis was explored by looking for correlations with measured soil properties, and by analysing the impacts of treatments. There were significant ($P<0.00143$) positive Spearman's rank correlations between NEPS and a range of soil properties. Addition of predatory mesostigmata mites to this indicator resulted in similar patterns but failed to improve the strength of correlations except for a slight improvement in that with total earthworms. But the observed patterns were not the same as found for NMDS axis 1. A balanced two-way ANOVA was conducted looking at the response of NEPS to control, manure or compost treatment across the three sites sampled in 2020 found no significant differences in NEPS due to treatment, site or the interaction between them (Figure 20).

The second axis of the NMDS correlated negatively with the % of collembola in the mesofauna community (hereafter referred to as %Collembola). %Collembola showed a significant negative Spearman's rank correlation with:

- Bulk density (P=0.00012)

%Collembola also showed significant positive Spearman's rank correlations with:

- %SOM (P=0.00012)
- % Total N (P=0.00037)
- Potentially mineralisable N (PMN; P=0.00088)

Analysis of variance between the treatments sampled in 2020 showed significant differences in mean value of %Collembola between sites (P=0.0029), treatment (P=0.026) and the interaction between these (0.01734).

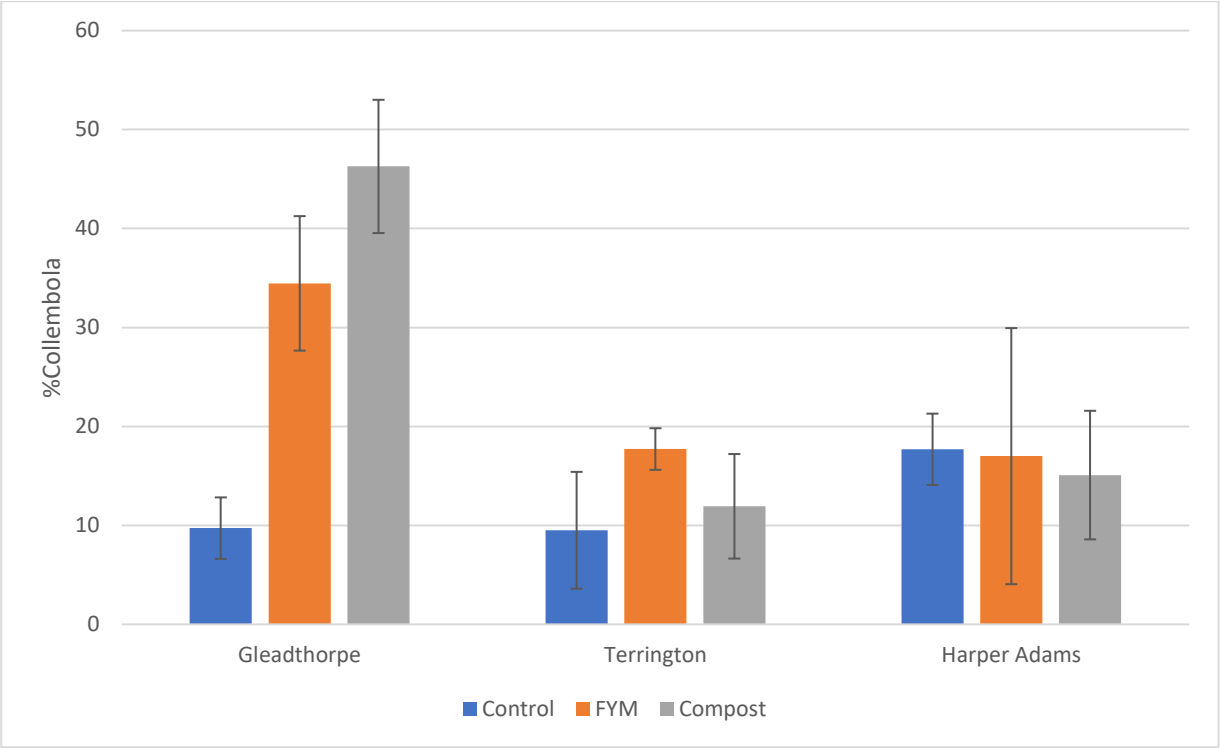


Figure 20: %Collembola (Collembola numbers as a % of total numbers of Collembola and mites) at 3 sites sampled in 2020, receiving treatments of either farmyard manure (pig manure at Terrington), green waste compost or no treatment (control). Bars show standard error (n=3).

Contextualising experimental plot mesofauna data against national datasets

It is possible that strong differences between site characteristics in both mesofauna communities and soil properties may result in correlations and conclusions that hold true only for the experimental sites. Considering the mesofauna data within a much larger dataset of 617 other soil mesofauna communities, from a wider range of arable, agricultural and semi-natural grasslands should help to show how these sites vary in line with nationally observed patterns of mesofauna community characteristics and help to increase the possibility that mesofauna community indicators of soil health are applicable nationally. These data were expressed as proportions, rather than total numbers, to enable valid comparison between datasets.

An NMDS was calculated for this wider national dataset, axis 1 from this NMDS was found to be positively correlated with the NDMS axis 1 for the full data set from the experimental data only, and negatively correlated with NDMS axis 2 from the same analysis, and therefore seems to capture some variability from both these axes. NMDS axis 1 was strongly influenced by the percentage of the collembola and acari community represented by Collembola, as shown in Figure 21.

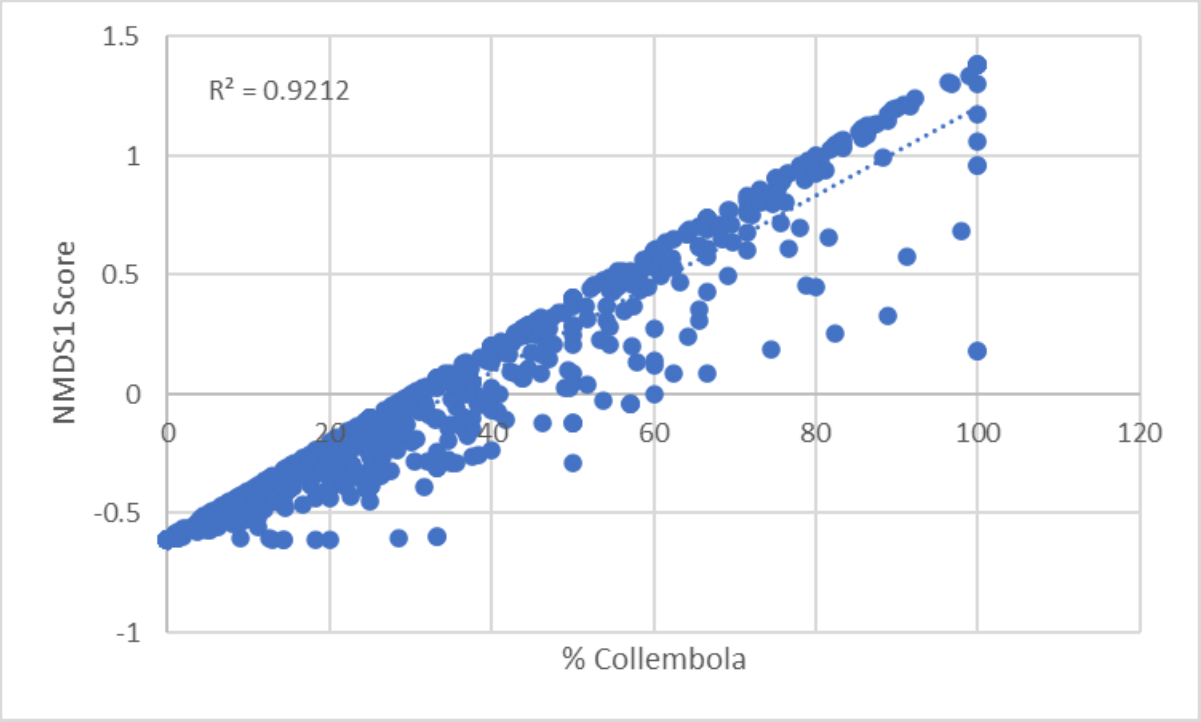


Figure 21: Relationship between NMDS axis 1, based on data from the plot experiments, combined with data from Natural England metabarcoding studies and the 1998 and 2007 countryside survey, and the percentage of collembola of the community of collembola and acari.

The mean values for axis 1 of for each combination of experimental site, treatment and sampling date, showed significant (P<0.00143) positive Spearman rank correlations with:

- SOM (p=0.0003)
- Total N (p=0.0005)

A negative Spearman's rank correlation was observed with Bulk density (P=0.0005). NMDS axis 2 was not found to be correlated with any other soil parameters at the Bonferroni adjusted threshold.

NMDS axis 1 scores for mesofauna data from the 2020 sampling were analysed using balanced two-way ANOVA, which detected significant differences between mean values between sites (P=0.0159), treatments (0.04879) and due to the interaction between site and treatment (P=0.01226). Scores for 2020 mesofauna communities on NMDS axis 2 were also analysed by ANOVA which detected significant effects between sites (P=0.001627) and treatments (P=0.03125). Mean scores on these axes for the sites and treatments included in this analysis are shown in Figure 22.

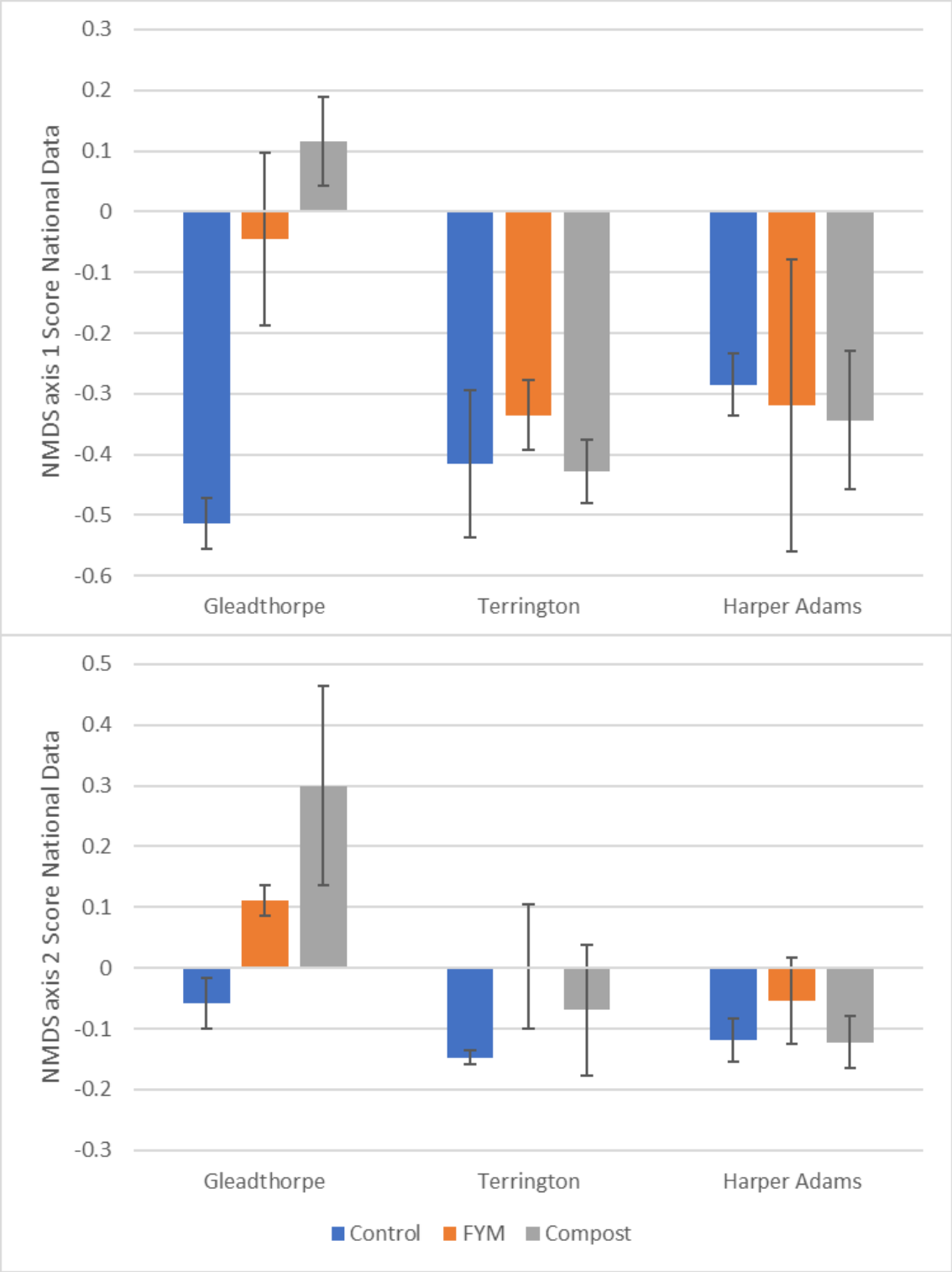


Figure 22: Scores on NMDS axis 1 (a) and NMDS axis 2 (b) based on analysis of the proportions of springtails from 4 order, and total mites, from all experimental sites in this study, combined with data from NE metabarcoding pilots on arable land and seminatural grassland, and combined with Countryside Survey lowland grassland and arable land data from 1998 and 2007. Graph shows only the scores for the 3 sites sampled in 2020, receiving treatments of either farmyard manure (pig manure at Terrington), green waste compost or no treatment (control). Bars show standard error (n=3).

Comparison of different sampling years at Harper Adams

A two-way ANOVA of the measurements of soil biology between 2017 and 2020 under control, FYM and compost treatments at Harper Adams, revealed that, compared with 2017, Harper Adams in 2020 had significantly higher total nematodes, stubby root nematodes, root lesion, root knot nematodes, and significantly lower earthworms, stunt/spiral nematodes and total mesofauna. More detailed resolution showed that in all treatments (control, FYM and compost) there was a significant decrease in richness of major taxon groups, mesostigmata mites, poduromorph and entomobryomorph springtails and an increase in oribatida mites and neelipleona springtails. These differences are consistent with those expected as a result of the difference in land-use at the time of sampling, with the 2017 sampling undertaken at the end of a 2 year grass/clover ley, and the 2020 sampling following 3 years of arable cropping (plough-based), including potatoes.

Interactions between year and treatment were found for Mesostigmata, Prostigmata and Entomobryomorpha springtails. To explore these interactions and indicate to what extent the treatments made the samples more or less resilient to the impacts of management between 2017 and 2020, individual plot differences in soil parameters from 2017 to 2020 were calculated as a single metric and analysed by one-way ANOVA against treatment effects. None of the soil physico chemical parameters showed directional responses that were different between treatments, indicating that the management impacts of the cropping rotation affected all treatments equally. Only the soil mesofauna parameters %Collembola and a similar parameter, which included mesostigmata mites with entomobryomorpha and poduromorpha springtails, showed significant differences, as did the 2nd axis on the NMDS analysis including national data. In these parameters. %Collembola and %NEPS+Mites changed least in the compost treatment and reduced moderately in the control treatment and most in the FYM treatment. A similar pattern was observed in the NMDS 2nd axis using the national soil data. These data suggest that different organic amendments may confer resilience to soil mesofauna communities in different ways.

Correlations between microscopic observation and metabarcoding of mesofauna diversity

Absolute richness

Both pH gradient and cropping stage significantly affected mesofauna sample diversity, as measured from CO1 ASV richness and the associated Shannon entropy ($P < 0.041$; Table 6) at Craibstone. Whereas no effect of pH level was observed on species richness recorded in the microscopy dataset (Section 4.3.2), a significant effect of cropping stage on species richness and Shannon entropy was apparent in both microscopy and the metabarcoding dataset. Although there was no significant effect, pairwise comparison showed similar trends between the different pH levels in both datasets and opposite patterns resulting from on crop when results from sequencing and microscopy data were compared (Figure 23).

Table 6: Kruskal-Wallis results for richness and Shannon entropy comparing the effect of pH and crop stage at Craibstone on the diversity of CO1 ASVs and microscopically observed species (values at P<0.05 shown in bold).

Method		**Sequencing**		**Microscopy**	
Treatment		**pH**	**Crop**	**pH**	**Crop**
Richness	**P**	**0.023**	**0.003**	0.248	**0.009**
	***E*²**	**0.148**	**0.258**	-	**0.194**
Shannon	**P**	**0.002**	**0.041**	0.152	**0.004**
	***E*²**	**0.267**	**0.120**	-	**0.229**

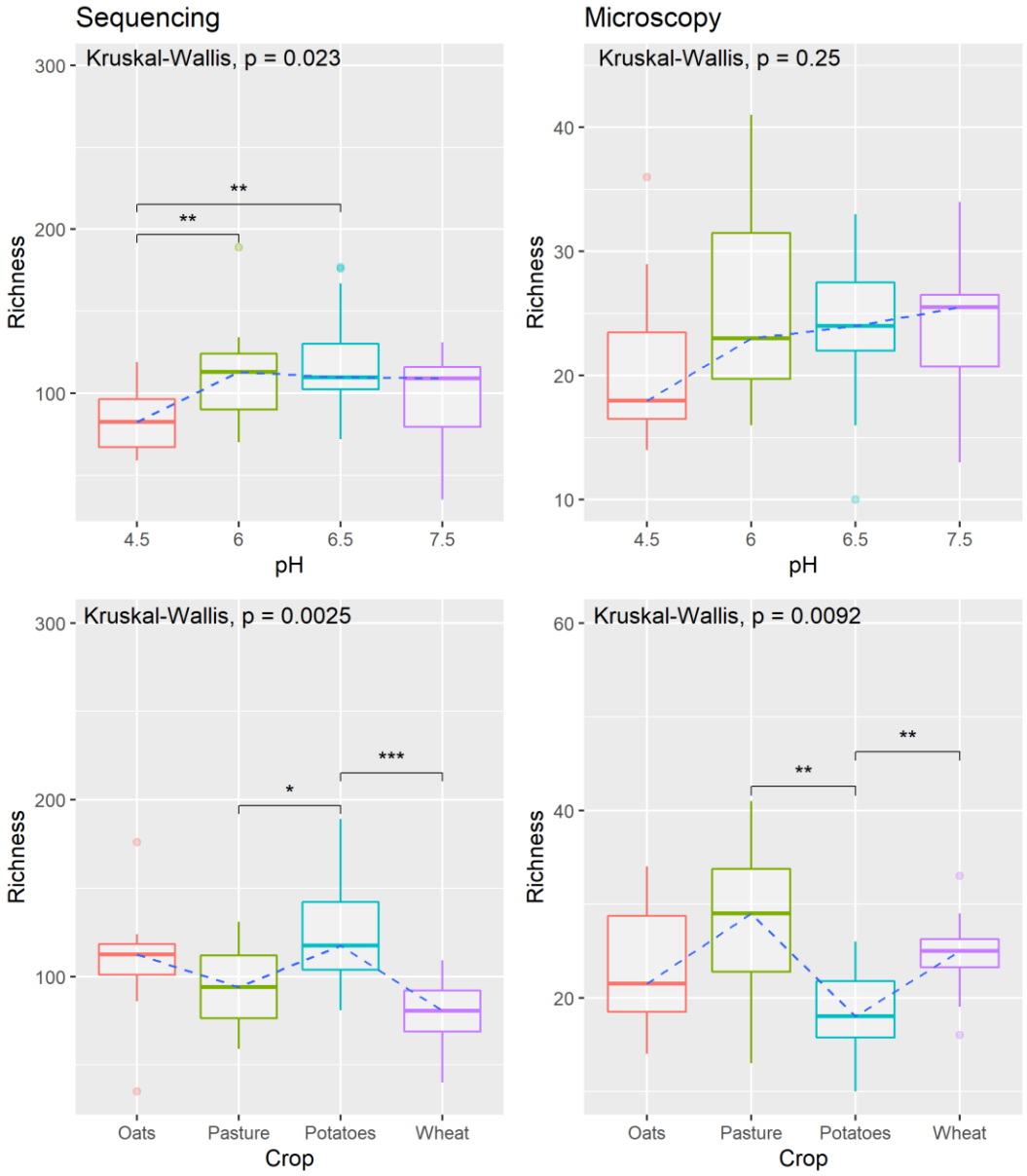


Figure 23: Absolute richness of arthropod AVS from sequencing (left-hand pair of graphs) and microscopically observed species (Microscopy, righthand pair of graphs) as affected by pH (upper pair of graphs) and cropping stage of the rotation (lower pair of graphs).

Correlation analysis: No correlation was apparent between sequencing and microscopy datasets for both observed features and Shannon entropies (Figure 24; Table 7). However, when both unweighted Jaccard distances were tested with Mantel, a significant but weak correlation was observed (Figure 25).

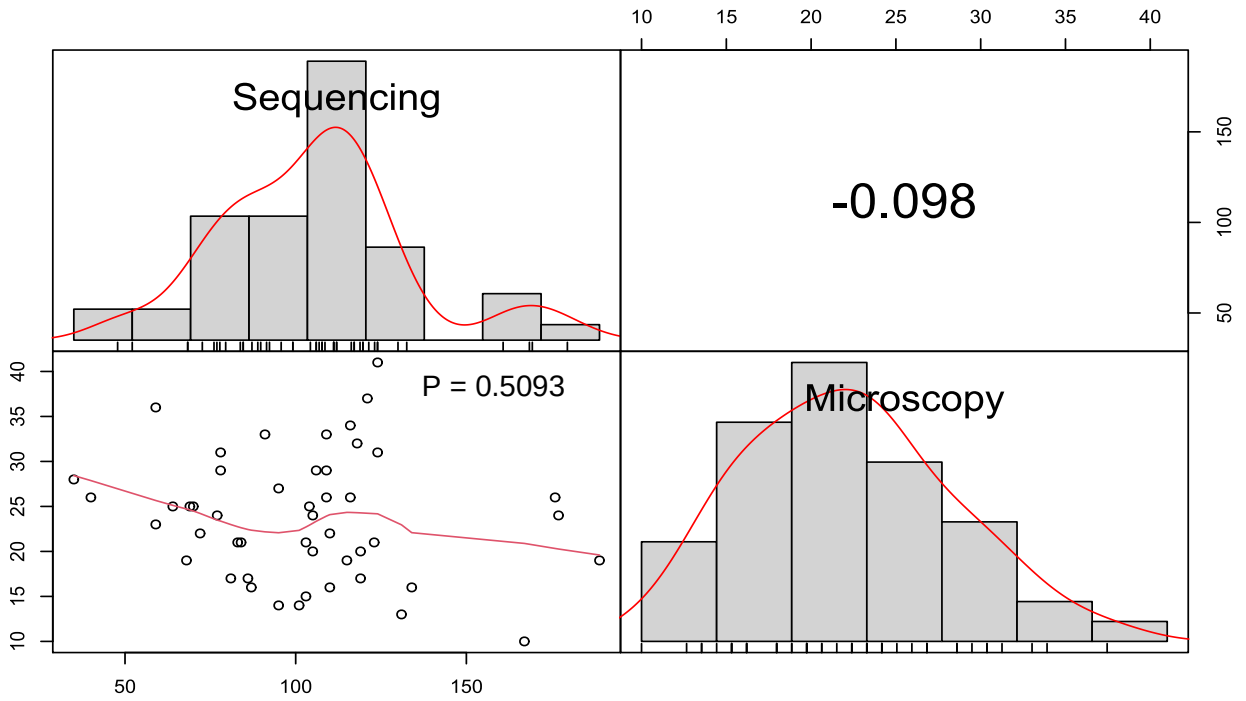


Figure 24: Spearman correlation chart comparing sequencing and microscopy richness data. Bar-plots show a normal sample distribution for both datasets. The scatter plot indicates the lack of correlation between richness values from both datasets.

Table 7: Spearman and Pearson correlation results between sequencing and microscopy richness and Shannon entropy data.

Method		**Spearman**	**Pearson**
Richness	**P**	0.509	0.315
	cor	-0.1	-0.15
Shannon	**P**	0.130	0.173
	cor	0.267	-0.2

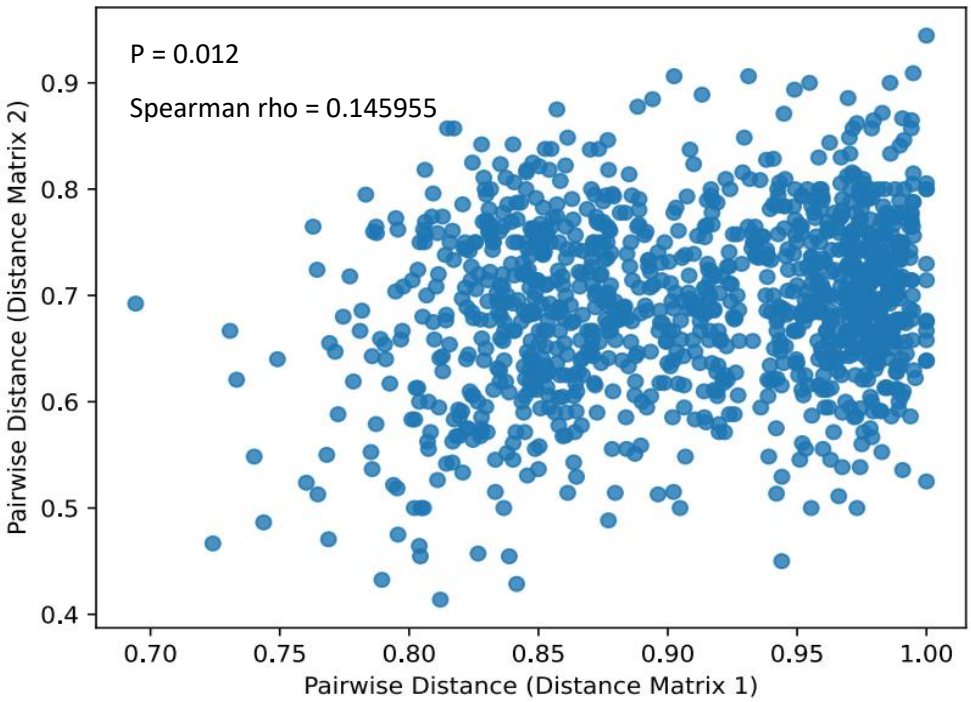


Figure 25: Correlation plot representing unweighted Jaccard pairwise distances from sequencing and microscopy datasets.

4.3.3. Effects on the soil microbiome

Effect of long-term pH manipulations on soil bacterial and fungal diversity

Overall, there were significant effects of pH on both soil bacterial and fungal diversity measured in the Craibstone: pH gradient experiment). A total of 5,852,860 16S rRNA and 5,291,924 ITS sequence reads were generated from the 48 pH trial soil samples which clustered into 8,353 and 4,975 different ASVs assigned to bacteria and fungi-like organisms respectively. The 16S ASVs were allocated phylogenetically into 38 phyla, 120 classes, 274 orders, 414 families, and 684 genera. ITS ASVs represented 13 phyla (including the phylum Oomycota of the kingdom Stramenopila), 38 classes, 100 orders, 217 families and 394 genera.

Numbers: Soil pH has long been known to impact community composition, and as expected a significant effect of pH extremes on both bacterial / 16S (P = 0.006) and fungal / ITS (P = 0.0005) ASV numbers (absolute richness) was observed (Table 8, Figure 26). Kruskal-Wallis pairwise comparisons showed the effect of pH mainly manifested as lower numbers of ASVs at pH 4.5, with a minor significant difference also observed between pH 6 and 7.5 level for ITS ASVs (Figure 26). On average, there were 450 fewer 16S ASVs and 135 fewer ITS ASVs from pH 4.5 samples than from samples at the other three pH levels. On average 27% of the variance in ASV numbers was explained by the effect of pH (Table 9).

Table 8: Effect of long-term pH and fertiliser manipulations on soil microbial diversity.

Treatment		pH		Fertiliser		Differences
Barcode		16S	ITS	16S	ITS	
Numbers	P	**0.005**	**0.001**	0.267	0.644	Only pH sig.
	E^2	**0.230**	**0.338**	-	-	-
Type	P	**0.001**	**0.001**	**0.017**	0.083	pH sig. effect on both; fertiliser only on bacteria
	R^2	**0.387**	**0.276**	**0.056**	-	pH ~7x greater size effect on bacteria only
ASV relative abundance	P	**0.001**	**0.001**	**0.037**	**0.001**	pH greater sig.
	R^2	**0.446**	**0.348**	**0.06**	**0.149**	pH ~4x greater mean size effect

The effect of long-term pH and fertiliser manipulations on soil microbial diversity and a comparison of the effects. P values from Kruskal-Wallis tests for ASVs richness and PERMANOVA test on binary (types) and abundance based Jaccard distances assessing the effect of pH and fertiliser manipulations on bacterial (16S) and fungal (ITS) communities. Effect sizes (E2 and R2) for significant differences at P<0.05 (in bold) are shown.

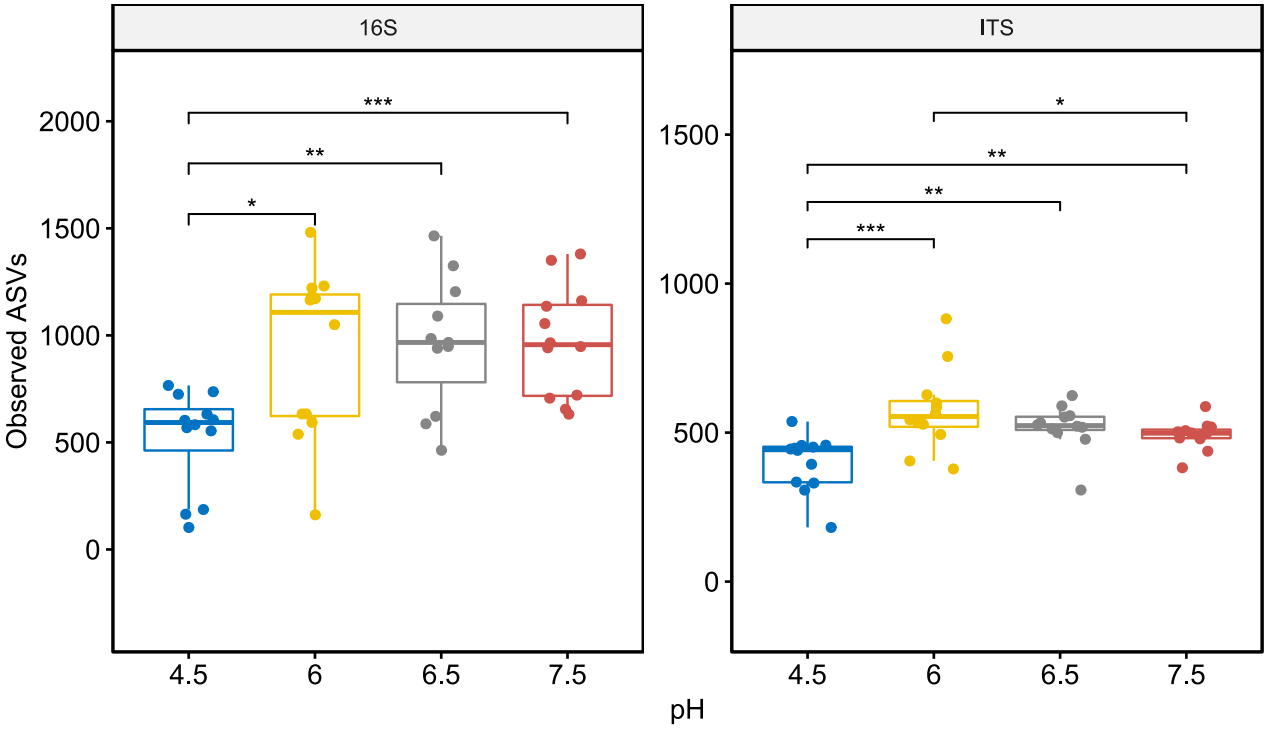


Figure 26: Boxplots comparing absolute ASV richness derived from bacterial (16S) and fungal (ITS) barcodes. Samples were grouped by soil pH level. Kruskal-Wallis pairwise results are indicated where adjusted p-value is significant (0 '***', 0.001 '**', 0.01 '*').

Types: PERMANOVA analyses revealed significant differences (P = 0.001) in the types (relative richness) of 16S and ITS ASVs present at each soil pH level (Table 8) and individual pairwise comparisons showed significant differences in the presences of 16S and ITS ASV types between all pH levels (P < 0.001; Table 8). Unweighted UniFrac analyses were also consistent with these findings Table 9 and on average, from the mean of R^2 values across PERMANOVA and UniFrac analyses, pH explained 33% of the variance in the presence of ASV types.

Table 9: Kruskal-Wallis results for Shannon and Faith PD estimated diversities; along with PERMANOVA results for UniFrac distances assessing the effects of soil pH and fertiliser application on bacterial (16S) and fungal (ITS) communities. Bold font indicates significant values where p-value ≤ 0.05.

Treatment		**pH**		**Fertiliser**	
Barcode		**16S**	**ITS**	**16S**	**ITS**
Shannon	**P**	**0.001**	0.415	0.268	**0.050**
	E^2	**0.308**	-	-	**0.130**
Faith PD	**P**	**2.75×10^{-5}**	**0.045**	0.538	0.488
	E^2	**0.484**	**0.115**	-	-
Unweighted UniFrac	**P**	**0.001**	**0.001**	0.377	0.26
	R^2	**0.46**	**0.315**	-	-
Weighted UniFrac	**P**	**0.001**	**0.001**	0.979	**0.003**
	R^2	**0.754**	**0.559**	-	**0.153**

Abundances: Significant differences were also observed in the relative abundances of bacterial / 16S and fungal / ITS ASVs at all pH levels (P = 0.001, Table 8) and pairwise analyses showed significant differences between all pH levels (P < 0.001). The separation of bacterial and fungal communities, especially at pH 4.5, was seen in PCoA plots (Figure 27). Again, UniFrac analyses were consistent with the signal from analyses of Jaccard dissimilarities, and collectively pH explained an average of 40% of the variance in ASV relative abundances.

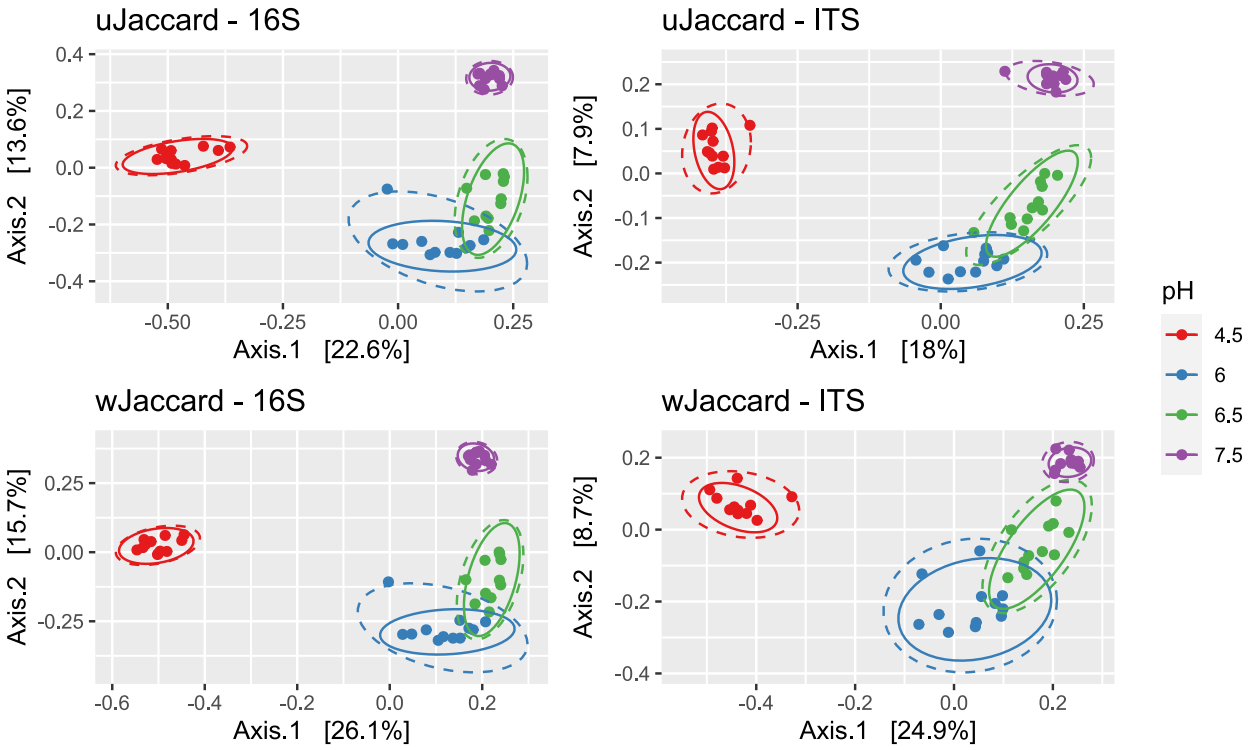
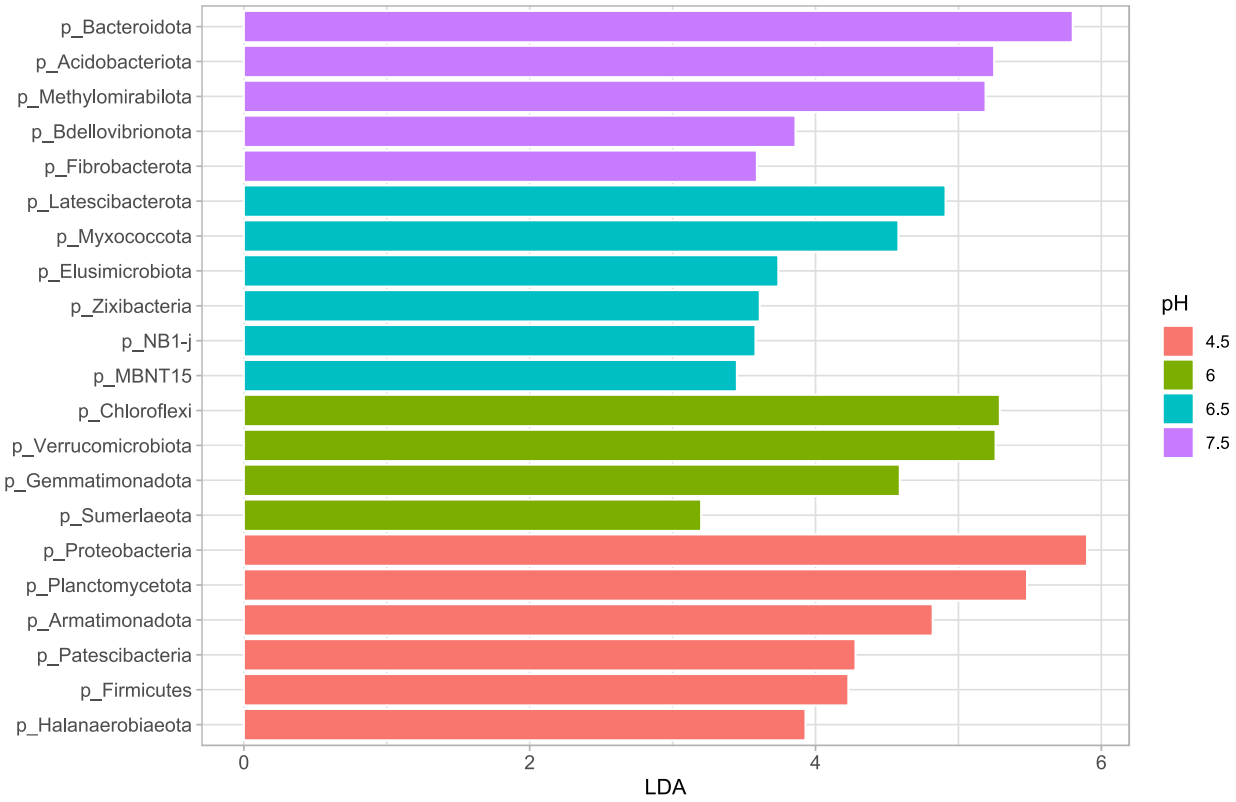


Figure 27: Principal Coordinate Analysis (PCoA) based on weighted and unweighted Jaccard distances matrices showing differences in relative abundances of 16S and ITS ASVs at each pH level.

Indicator taxa: LFsE predictions suggested that a total of 1,583 16S and 583 ITS ASVs significantly (p.adj < 0.05) differed in abundance between pH levels spanning 21 assigned bacterial phyla (Figure 28a) and 8 fungal phyla (Figure 28b). The number of 16S ASVs overrepresented at pH 7.5 (671) was greater than those at other pH levels (mean = 395.75) and spanned ASVs assigned to ten different phyla (Figure 29a). The pattern for fungi was different with larger numbers of ITS ASVs overrepresented at both pH 4.5 (204) and 7.5 (181) compared to pH 6 and 6.5 (mean = 145.75, 95 and 103 at pH 6 and 6.5 respectively; Figure 29b), and the largest number of differentially abundant ASVs were assigned to Ascomycota.

(a)



(b)

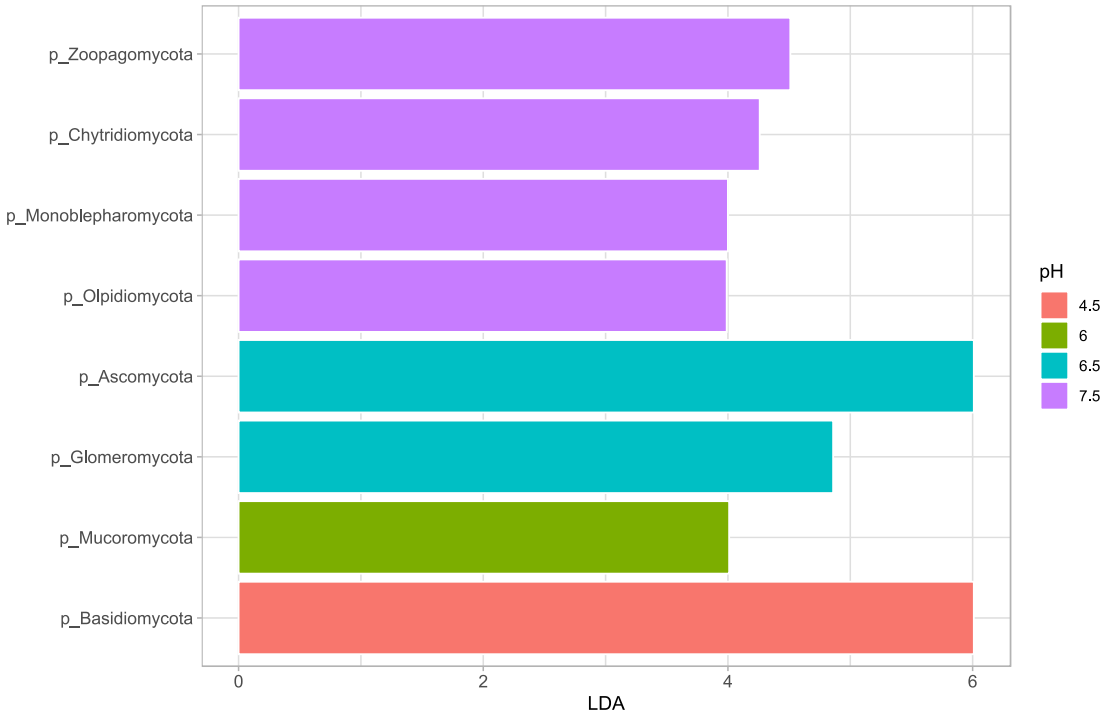
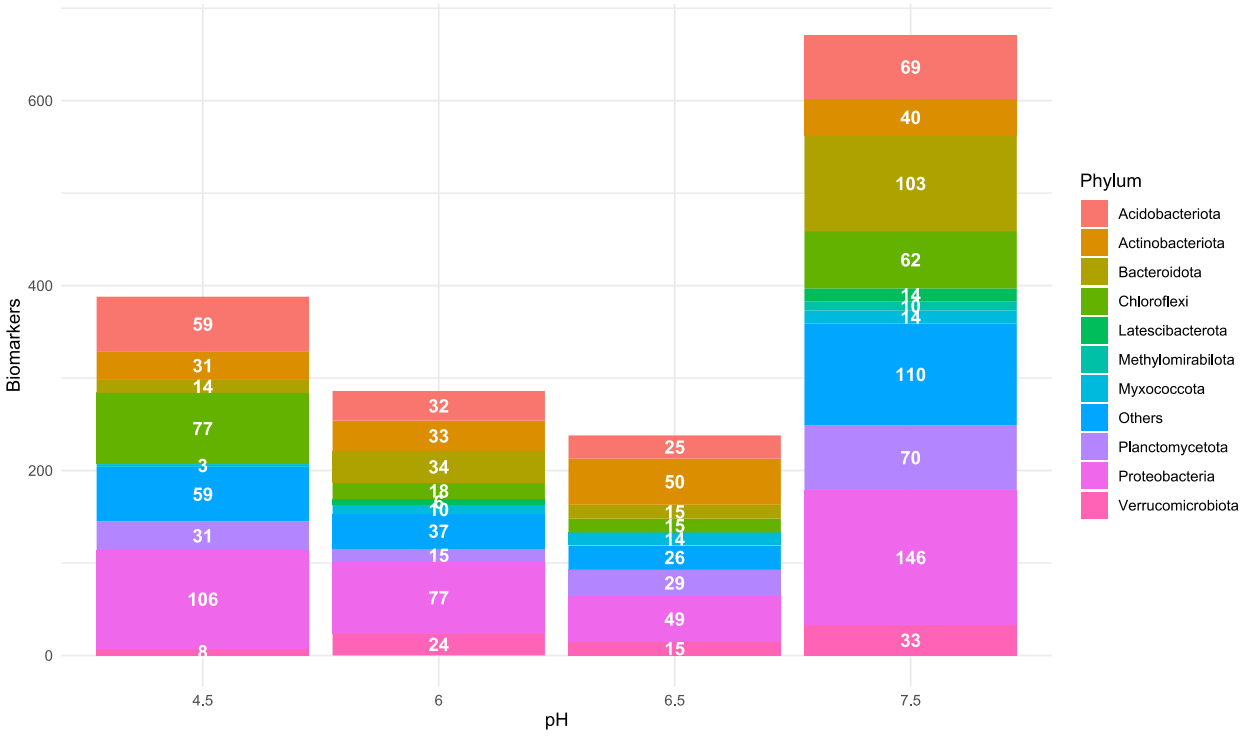


Figure 28: LEfSe prediction at phylum level of (a) 16S and (b) ITS ASVs as potential biomarkers of pH. Each predicted biomarker was ranked according to Linear Discriminant Analysis (LDA) effect size.

(a)



(b)

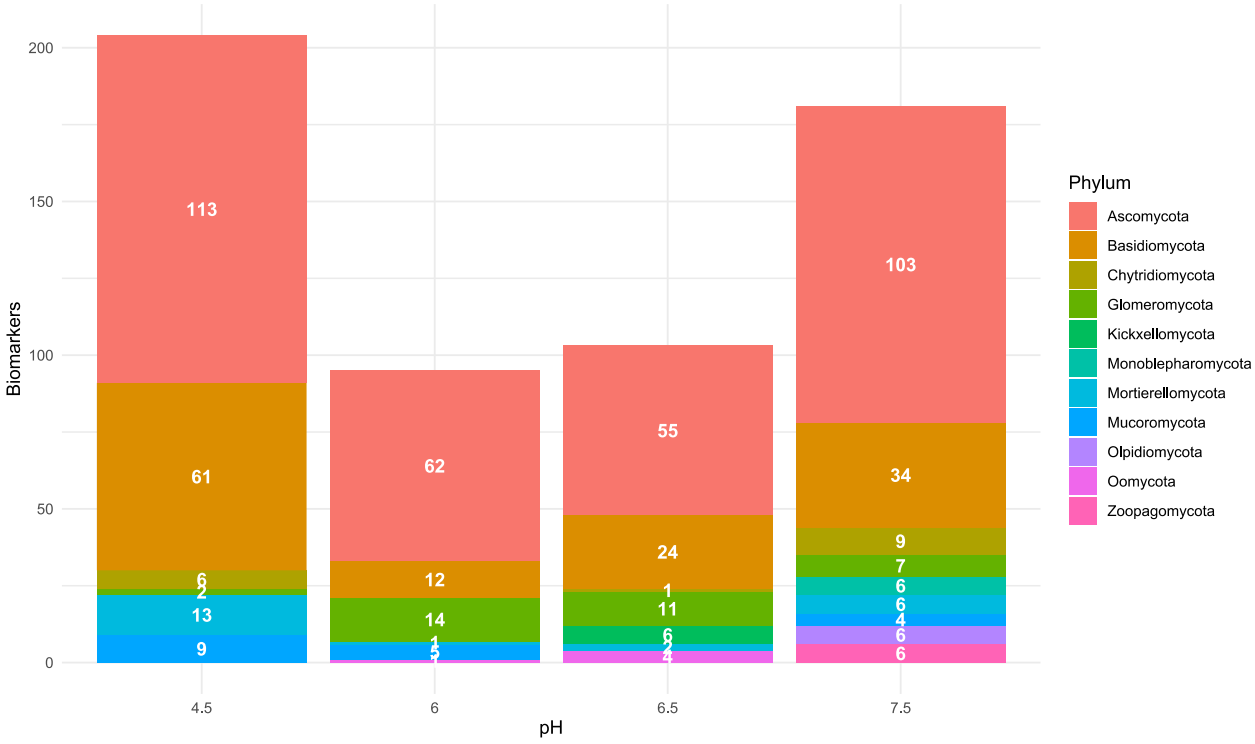


Figure 29: Histograms representing the number of total (a) 16S and (b) ITS ASVs per phylum predicted by LEfSe analysis as biomarkers of the different pH levels. “Others” includes a cluster of bacterial phyla, each with less than 40 total biomarkers.

Functional predictions: Due to the low resolution at species and genus levels using the SILVA reference database, only 12.7% of 16S ASVs could be assigned to a functional group by FAPROTAX. The largest assigned groups recognised were aerobic-chemoheterotrophs, animal parasites or symbionts and chemoheterotrophic bacteria (Figure 30a). A greater taxonomic resolution at family, genus and species levels allowed 65% of ITS ASVs to be assigned to a trophic mode and guild (Figure 30b).

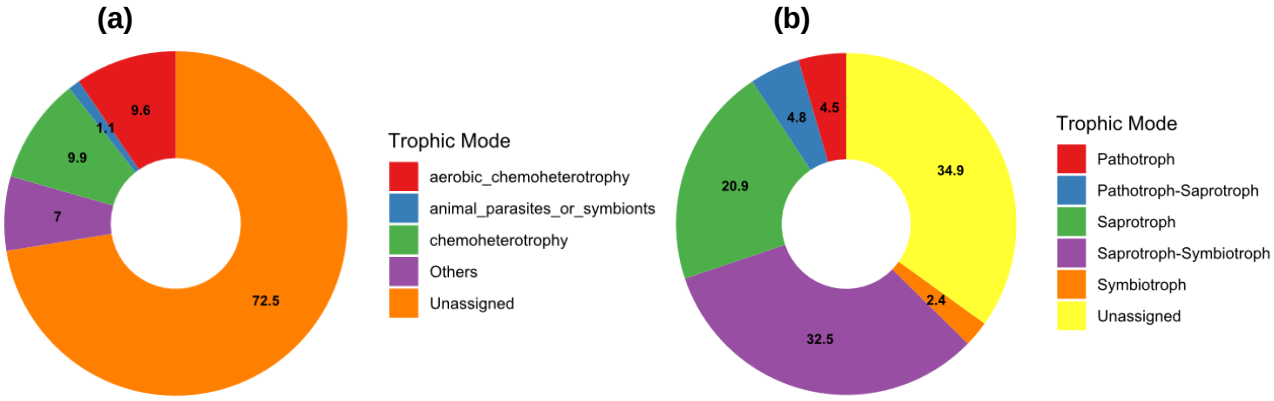
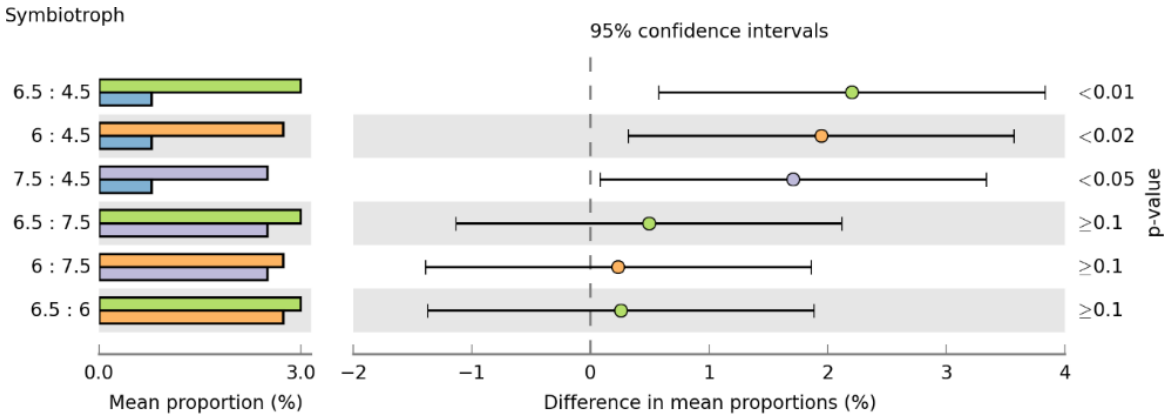


Figure 30: Proportion of (a) 16S and (b) ITS ASVs from the pH trial, assigned to functional groups and trophic modes. Bacterial functional groups with less than 0.01 abundance were clustered into the group “Others”. Fungal group “Unassigned” also contains ambiguous trophic modes assigned to “Pathotroph-Symbiotroph” and “Pathotroph-Saprotroph-Symbiotroph”.

Differential abundance analyses showed that fungal symbiotrophs where significantly ($p.adj < 0.05$) less abundant at pH 4.5 (Fig. 31a), and pH 7.5 contained significantly ($p.adj < 0.05$) more fungal pathotrophs than pH 4.5 and 6 (Fig. 31b).

(a)



(b)

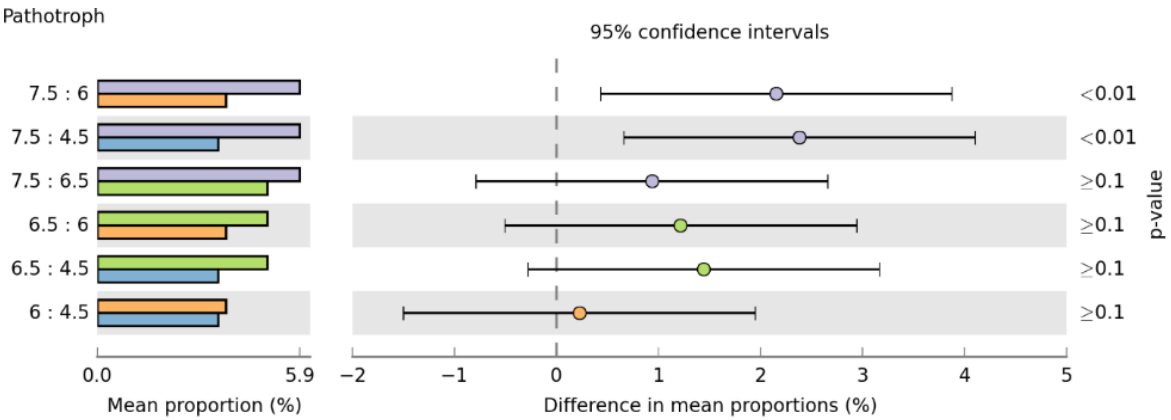


Figure 31: STAMP plot representing statistical pairwise comparing relative abundances of the trophic modes symbiotrophs (a) and pathotrophs (b) at different pH levels. Top bar of each pairwise comparison represent the treatment on the left while bottom bars represent the treatment on the right.

Effect of long-term fertiliser application on soil bacterial and fungal diversity

The 24 fertiliser trial samples yielded 2,596,928 16S and 2,347,133 ITS sequence reads which clustered into 5,188 16S and 3,458 ITS ASVs. 16S ASVs were phylogenetically allocated into 38 phyla (of which 9 were candidate phyla), 119 classes, 244 orders, 365 families and 568 genera. ITS ASVs represented 12 phyla (including the phylum Oomycota of the kingdom Stramenopila), 34 classes, 87 orders, 191 families and 336 genera.

Numbers: Analyses revealed no significant effect of fertiliser application on 16S or ITS ASV richness (P >0.25; Table 8) and this was consistent with Faith-PD analysis.

Types: PERMANOVA analyses revealed weakly significant differences between fertiliser treatments in the types of 16S ASVs (P = 0.017), which only explained approximately 6% of the variance in assigned bacterial types (Table 8). There were no differences in the types of ITS ASV between fertiliser treatments (P = 0.083), and unweighted UniFrac analyses revealed no effect of fertiliser treatment on either bacteria or fungi (Table 9).

Abundances: There were significant differences between fertiliser treatments in the relative abundance of types of both 16S (P = 0.037) and ITS (P = 0.001) ASVs which explained 6% and 15% of variance in ASV abundances respectively (Tables 8 and 9, Figure 32). These analyses were in line with weighted UniFrac (P = 0.003) for fungi but not bacteria (P = 0.98).

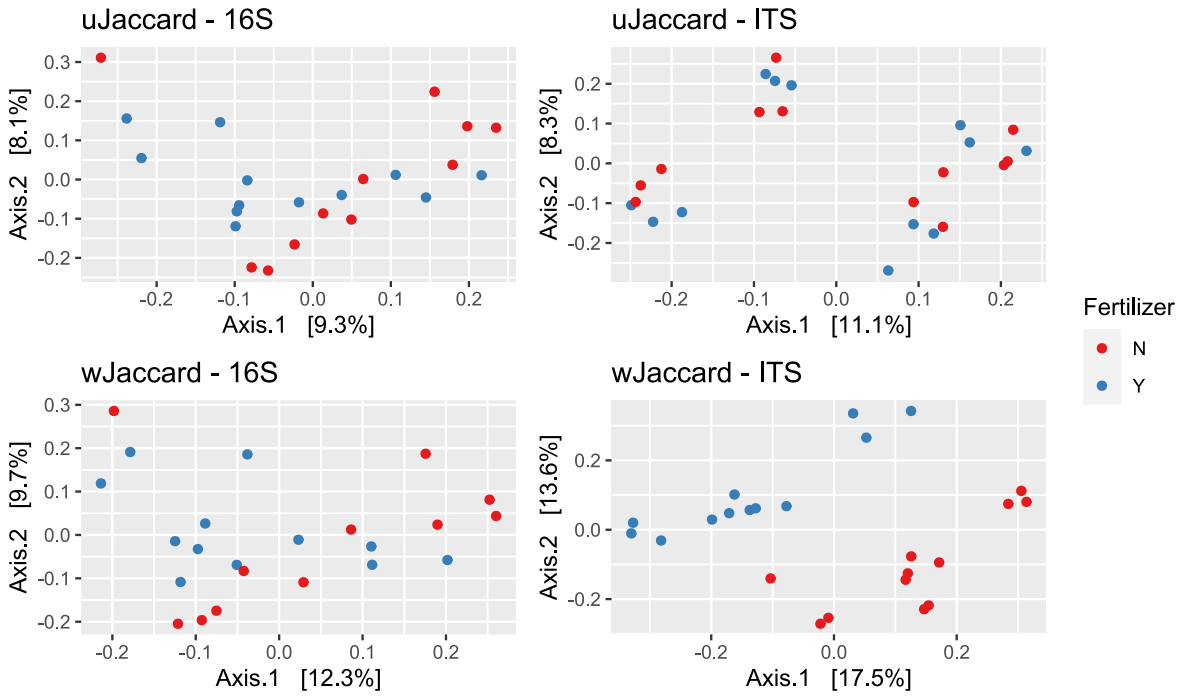


Figure 32: Principal Coordinate Analysis (PCoA) based on weighted and unweighted Jaccard distances matrices showing differences in relative abundances of 16S and ITS ASVs pH level. Colours represent soil from plots treated with (Y, Blue) or without (N, Red) fertiliser.

Indicator taxa: No 16S ASVs were identified as significantly overrepresented in one or other fertiliser treatment. Only 4 fungal Ascomycota ASVs, and one assigned to *Mortierella minutissima* (Mortierellomycota), were significantly associated with the no fertiliser control treatment.

Functional predictions: No bacterial or fungal functional groups were inferred to be over-represented between fertiliser treatments.

Standardised approach for combined analyses of the relative effects of pH and fertiliser across both trials

The new bioinformatics pipeline (as described in Section 3.5.4) was used to compare the diversity of bacteria and fungi across both Craibstone field trials. To understand which highly similar ASVs were common across the two independent trials, ASVs from both trials were merged and normalised together and those within 98% similarity were then clustered to form a new data set of merged ASVs, hereafter referred to as mASVs. Those mASVs common to both trials could then be selected and the relative effects of pH and fertilisation on their numbers, types and relative abundances could be compared using this standardised approach. When the distribution of individual mASVs (roughly equivalent to species) was considered, all diversity metrics were significantly affected by pH, fertilisation, and trial location (Table 10).

Table 10: The effect of long-term pH and fertiliser manipulations on soil bacterial (16S) and fungal (ITS) communities and the differences between pH and fertiliser trials overall and between all treatments from both trials using >98%mASVs derived from a merged dataset.

Treatment		All treatments		pH		Fertiliser		Trial site	
Barcode		16S	ITS	16S	ITS	16S	ITS	16S	ITS
Taxa richness	P	**1.44E-05**	**0.0003**	**0.001**	**0.001**	0.424	0.644	**0.003**	**0.003**
	E^2	**0.391**	**0.286**	**0.315**	**0.331**	-	-	**0.114**	**0.116**
Types of taxa	P	**0.001**	**0.001**	**0.001**	**0.001**	**0.032**	0.087	**0.001**	**0.001**
	R^2	**0.397**	**0.302**	**0.411**	**0.305**	**0.056**	**-**	**0.100**	**0.088**
Abundances of taxa	P	**0.001**	**0.001**	**0.001**	**0.001**	**0.033**	**0.002**	**0.001**	**0.001**
	R^2	**0.514**	**0.362**	**0.520**	**0.357**	**0.063**	**0.114**	**0.131**	**0.097**

P values from Kruskal-Wallis (taxa richness) and PERMANOVA tests for taxa types and abundance based Jaccard distances. Significant test statistics at P<0.05 are shown in bold and include effect sizes (E^2 and R^2).

Numbers: Analyses of the merged dataset indicated significant differences in total numbers of both bacterial / 16S and fungal / ITS mASVs between all treatments and across both trial sites ($P < 0.00026$, Figure 33). Similar differences were also observed following pairwise comparisons of all treatments when the two trials had been analysed separately. As in the previous analyses, no effect of fertiliser treatments on 16S or ITS mASV absolute richness was found. Additional significant differences between 16S mASV absolute richness in the fertiliser treated plots with all pH level groups except 4.5 were observed. However, in the case of ITS mASV absolute richness, fertiliser treated soil differed only with soil from the most extreme pH levels (4.5 and 7.5: Figure 33).

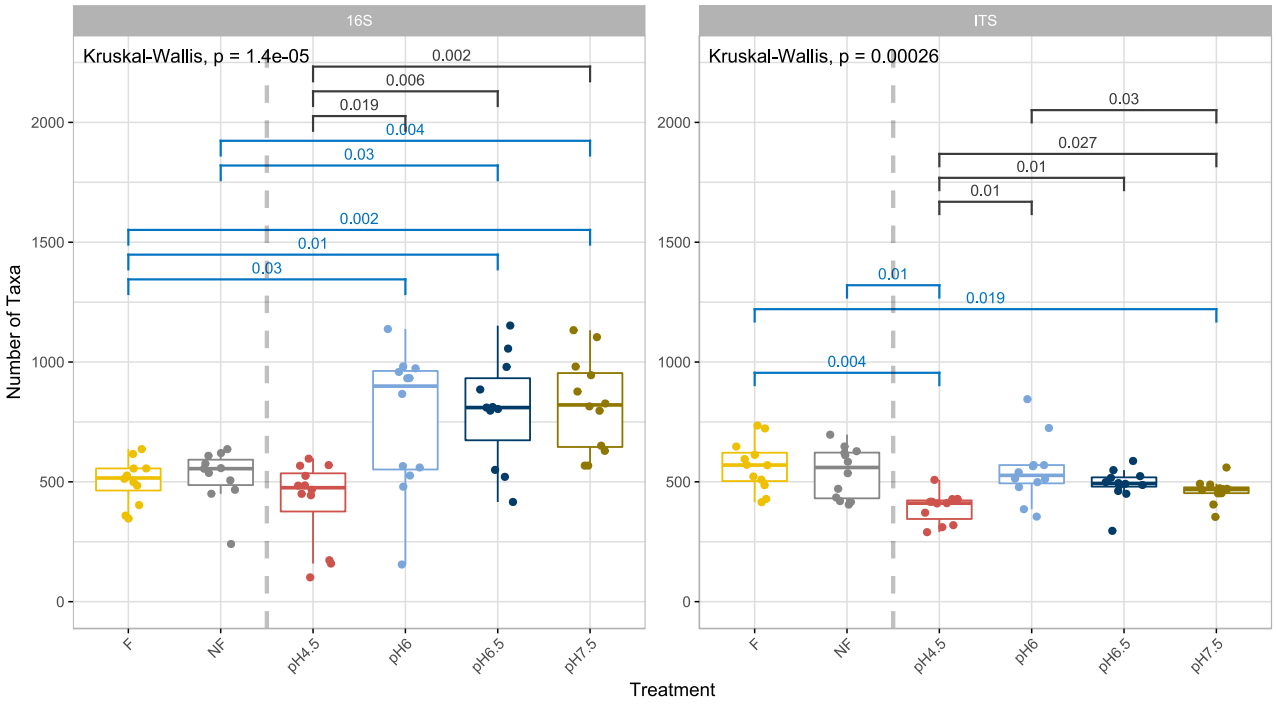


Figure 33: Comparing mASV richness for 16S and ITS barcodes from the merged dataset across all treatments. Significant differences in numbers of mASVs between treatments as revealed by Wilcoxon pairwise tests are shown in blue for between trial and in black for within trial comparisons (p-values have been adjusted using the Benjamini-Hochberg method).

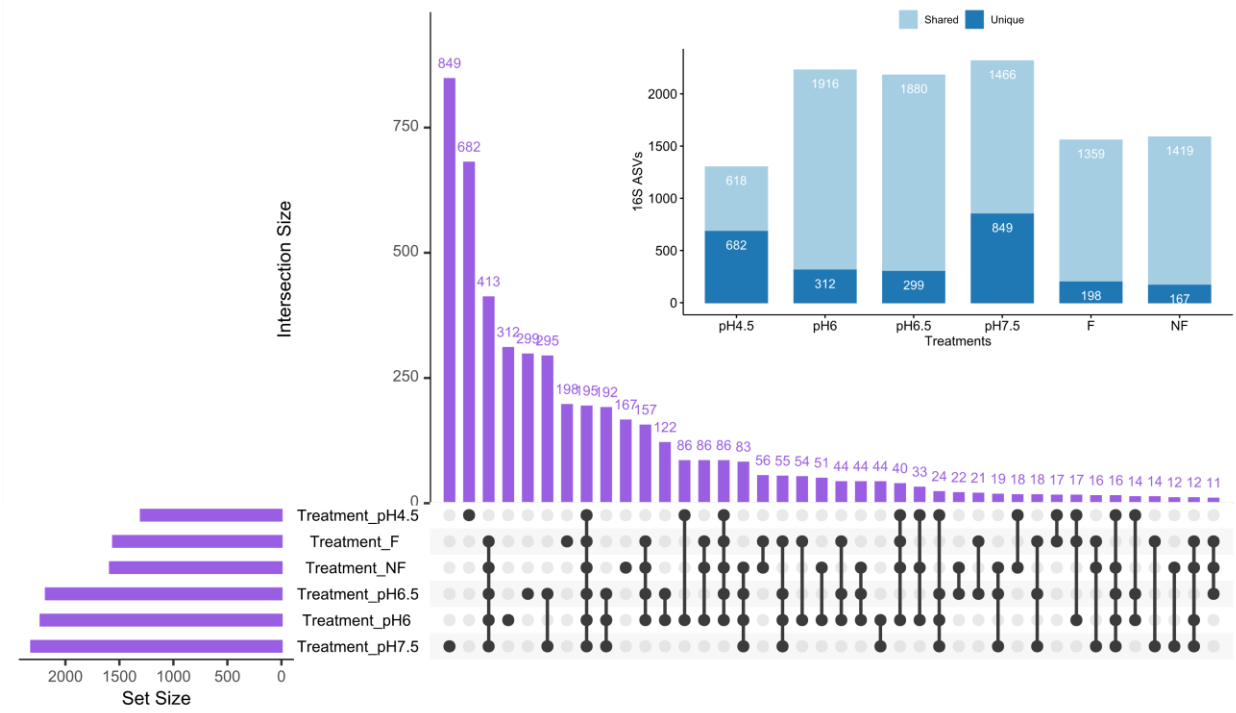
Types: PERMANOVA analyses revealed significant differences ($P < 0.001$) in the types of 16S and ITS mASVs present in soil following all treatments, except for the fertiliser treatments which (as in the separate analysis of the fertiliser trial) did not significantly affect the relative richness of ITS mASVs (Table 10). Approximately 40% of the variance in bacterial types and 30% in fungal types was explained overall. These results compare similarly to those obtained from the unmerged analyses of the separate pH and fertiliser trials, except that the size of the pH effect was even greater

than previously estimated. Figure 34 displays the proportion of mASVs overlapping between each of the 6 treatments, where most unique taxa were found in soil samples with extreme pH levels (4.5 and 7.5).

Abundances: As with the analyses of the individual trials, PERMANOVA analyses of the merged data set revealed significant differences between treatments ($P < 0.001$) in the relative abundances of 16S and ITS mASVs in soil; treatment explained approximately 30% of the variance in the relative abundance of bacteria and 36% of fungi-type organisms (Table 10). PCoA ordination plots confirm these differences (Figure 35) and all pairwise analyses revealed significant differences between all treatments ($P < 0.033$), although pH 4.5 clearly had a greater effect than any of the other treatments. The relative abundance of both 16S and ITS mASVs from the fertiliser treatments (maintained at pH 6) aligned most closely with those from the pH 6.0 treatments; again highlighting the importance of pH influencing the microbial diversity. Nevertheless, a significant difference in both 16S and ITS mASV abundances was observed between the two trials at pH 6.0. The effect of fertiliser/no fertiliser on relative mASV abundances, was more difficult to distinguish.

The community composition of the bacterial and fungal microbiomes of the soils in both trials was roughly similar when mASVs were clustered at phylum level (Figure 36), as may be expected from trials close to each other with similar soil types and conditions. Similarly, at this taxonomic level, the relative contribution of each phylum to the soil community was similar irrespective of pH or fertiliser regime (Figure 37).

a)



b)

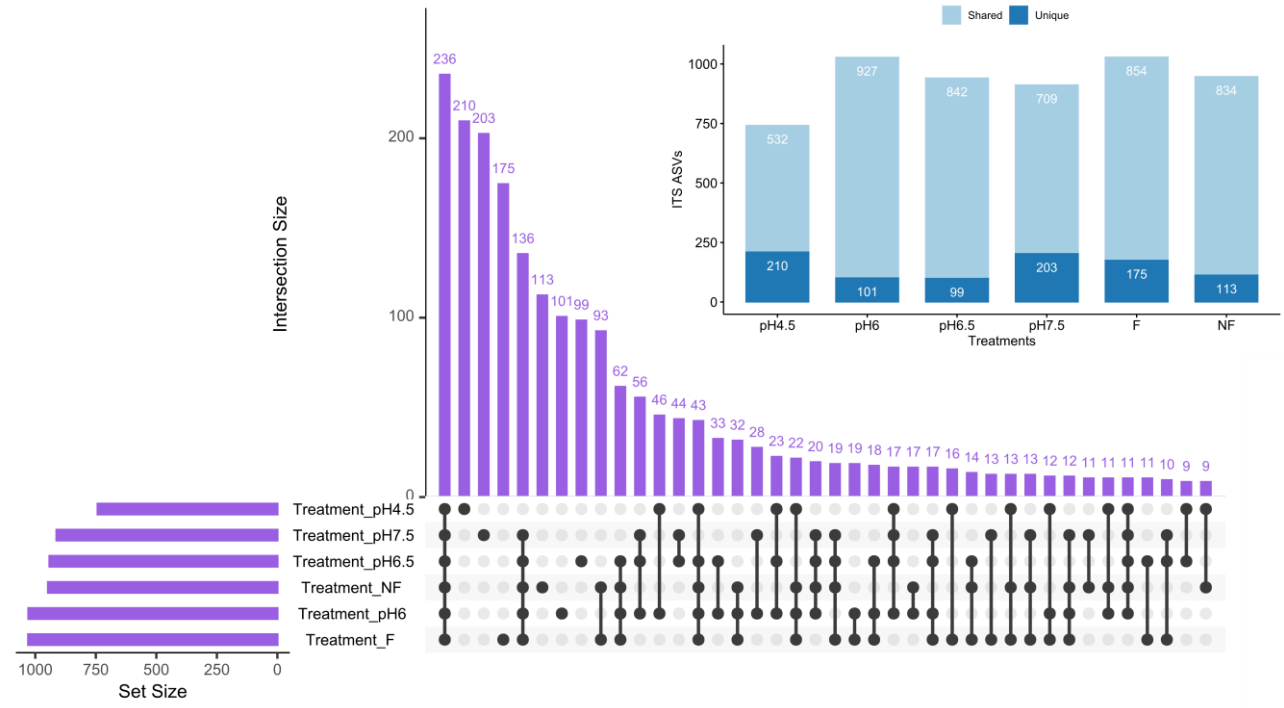


Figure 34: UpSet plots representing the proportion of different 16S (a) and ITS (b) mASVs across two field trials as influenced by six different soil management treatments: No fertiliser (NF), complete fertiliser (F), pH 4.5, pH 6.0, pH 6.5 and pH 7.5. Intersection Size bars show the number of mASVs in common between treatments indicated by the dots below. Set Size bars represent the number of observed different mASV associated with each soil treatment.

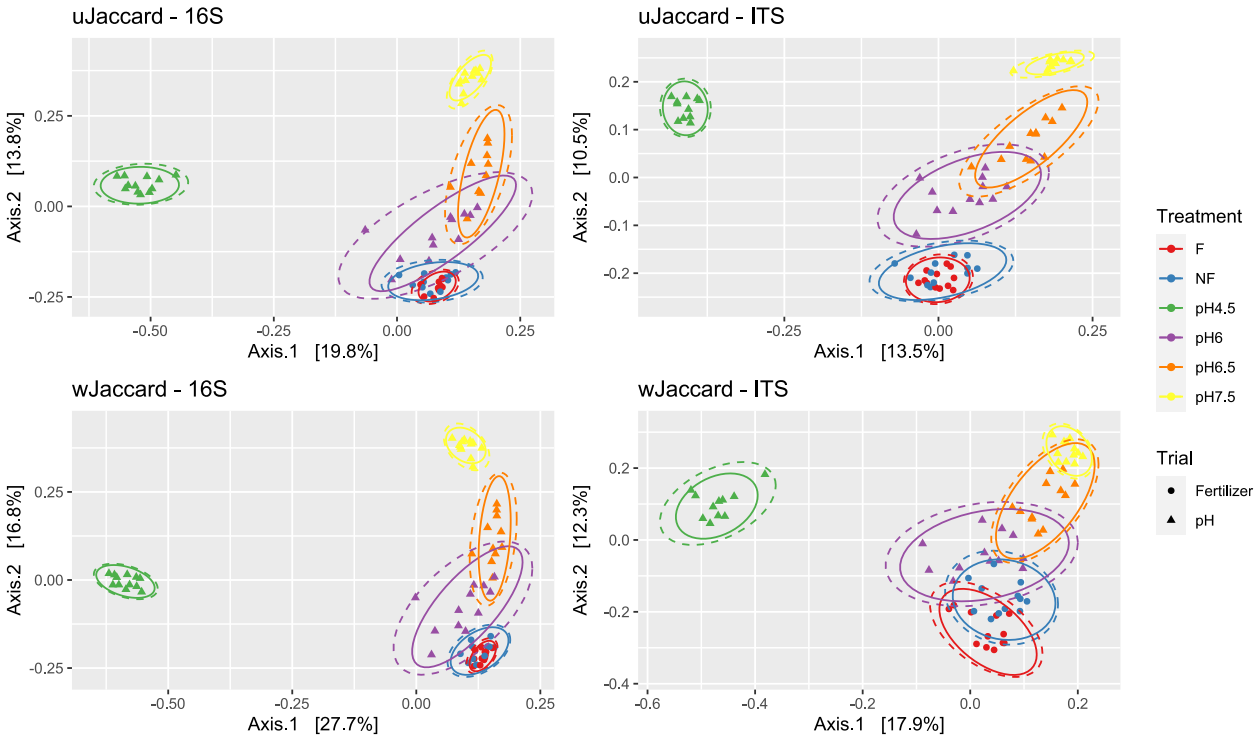


Figure 35: Principal Coordinate Analysis (PCoA) of the relative abundance of 16S and ITS mASVs based on weighted and unweighted Jaccard distance matrices the merged datasets from the two trials. No supplemental fertiliser (NF), supplemental fertiliser (F), pH 4.5, pH 6.0, pH 6.5 and pH 7.5.

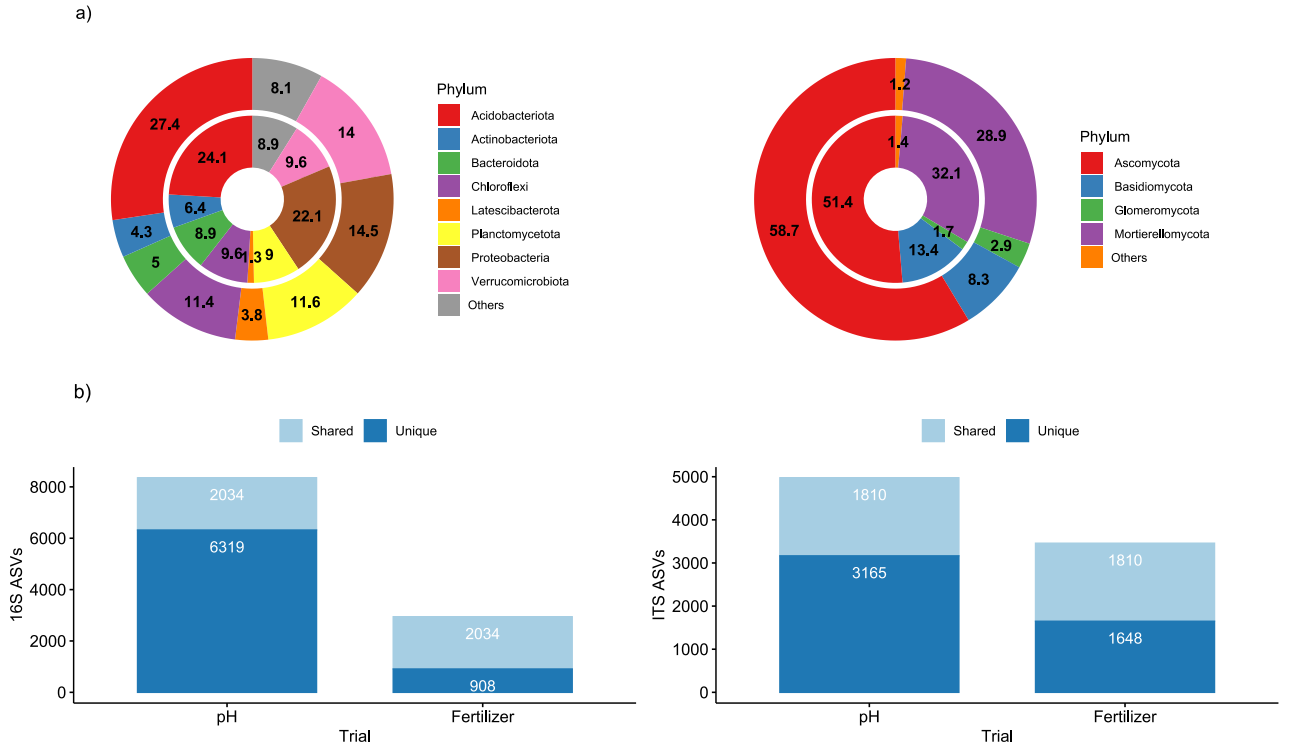


Figure 36: Comparison of bacterial and fungal diversity across two Craibstone field trials. (a) bacterial (left) and fungal (right) community composition: Relative abundance of indicated phyla in all samples from the pH trial (inner plot) and fertiliser trial (outer plot) and (b) the absolute richness of observed ASVs: Bar-plots represent the number of ASVs common to both trials (light blue) and the number of ASVs exclusive to each (dark blue).

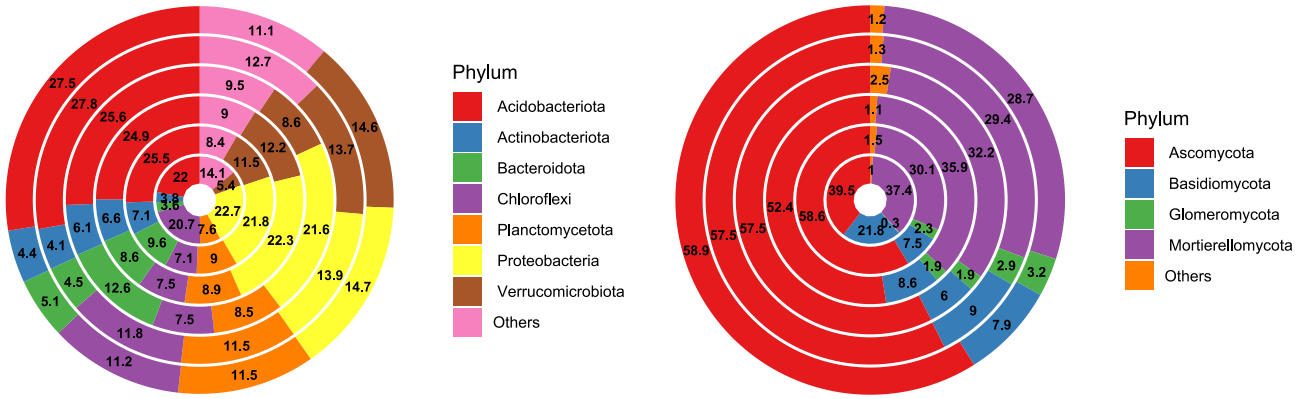


Figure 37: Comparison of bacterial (left) and fungal (right) diversity at phylum level across two field trials as influenced by different soil management treatments. Treatments represented by inner to outer plots are: (a) no fertiliser, (b) complete fertiliser, (c) pH 4.5, (d) pH 6.0, (e) pH 6.5 and (f) pH 7.5.

Effects of organic amendments on soil bacterial and fungal diversity

Amendment with organic materials (Harper Adams: organic amendment trial, 2017) had no significant impact on the fungal or bacterial communities, although there were some differences in organism types (Section 4.2). Estimated diversity of both bacteria and fungi were then independently compared across three long-term organic amendment trials at separate sites sampled in 2020 across England. Figure 38 shows that the community composition of the bacterial and fungal microbiomes of these soils was similar when ASVs were assigned at phylum level, despite the geographical distance from each other. However, there were significant numbers of both 16S and ITS ASVs and mASVs that were unique to each site (Figure 39), with only around 5% of ASVs being common to all three sites in each case (Figure 39) but higher proportions of mASVs common to more than one site (Figure 40). The highest numbers of unique taxa were found at Gleadthorpe, treatments sampled at this site also covered a more diverse range of input types.

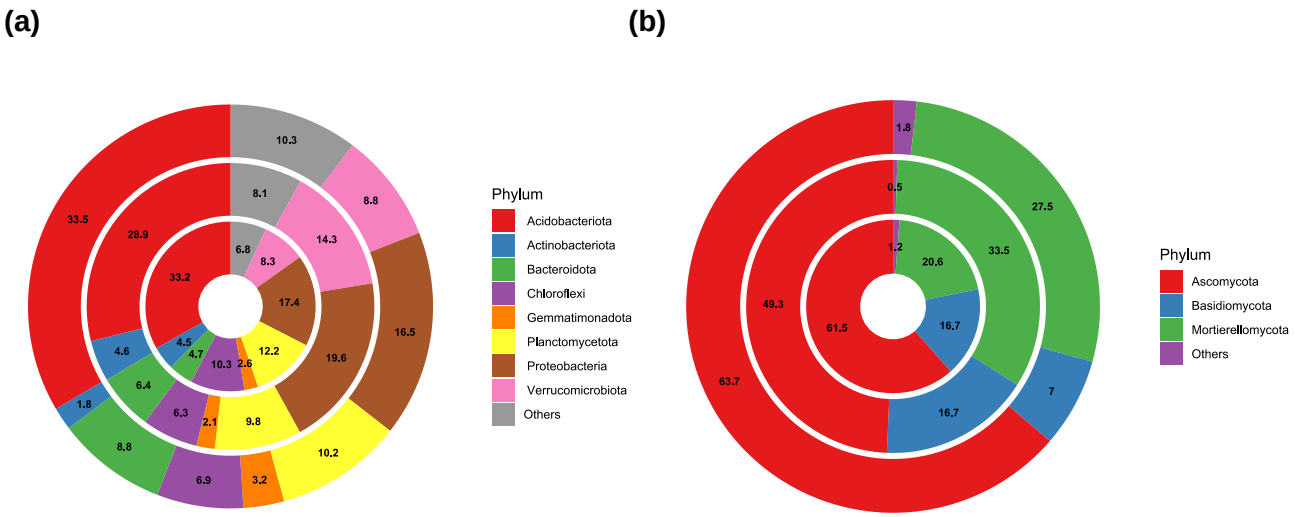


Figure 38: Comparison of bacterial and fungal soil community compositions across three long-term organic-amendment trials. Relative abundance of assigned phyla to (a)16S and (b) ITS ASVs from all samples collected from trials at ADAS Terrington (outer), Harper Adams University (middle) and ADAS Gleadthorpe (inner).

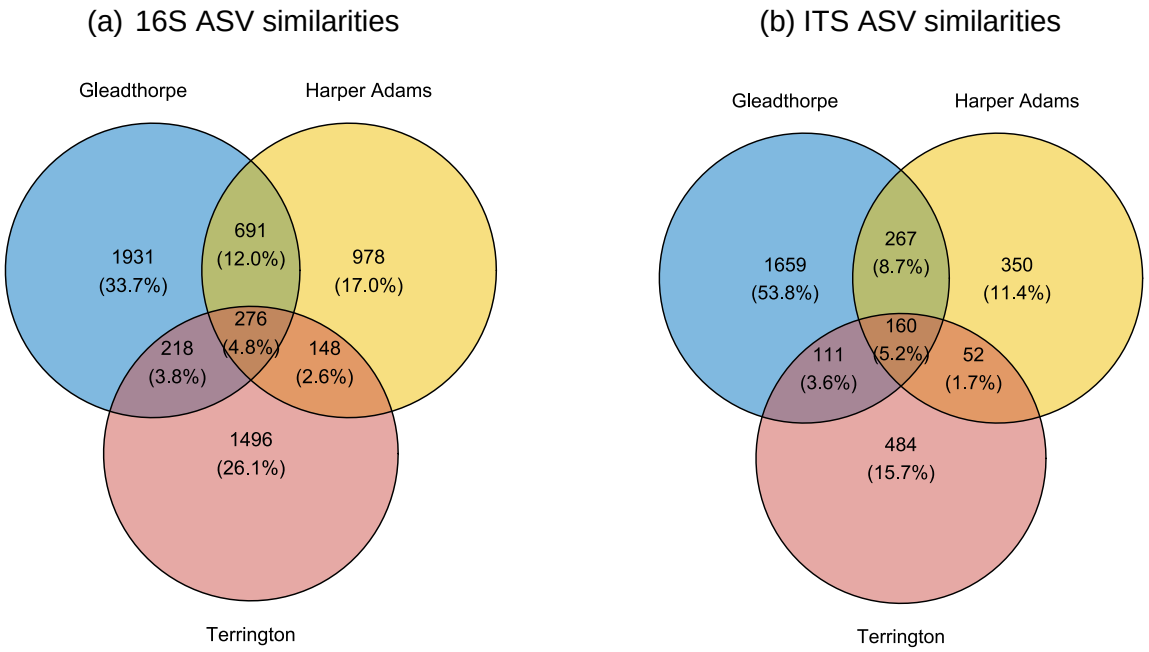


Figure 39: Commonality of (a) 16S and (b) ITS ASVs across three long-term organic amendment trials.

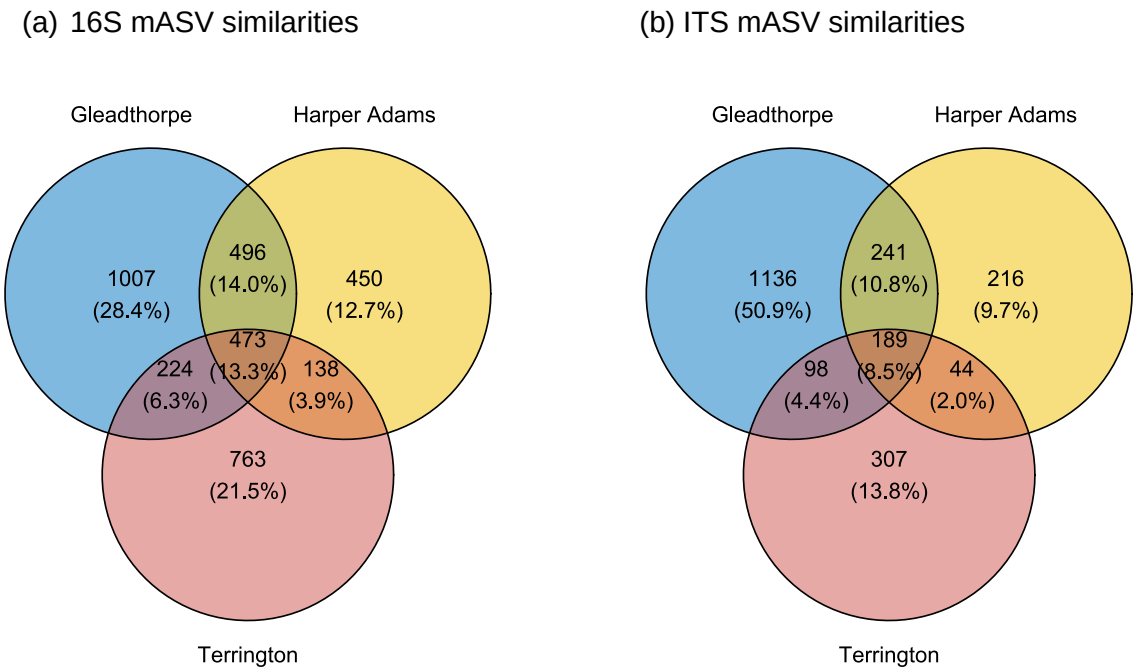


Figure 40: Commonality of (a) 16S and (b) ITS mASVs across three long-term organic amendment trials.

Numbers: Analyses revealed no significant effect of organic amendment application on 16S richness (P >0.2; Table 11) in any of the trials. However, a slightly significant effect was observed for ITS ASV richness only in the Gleadthorpe trial.

Table 11: Kruskal-Wallis results for Observed ASVs along with PERMANOVA results for Jaccard distances assessing the effect of the different organic amendment treatments in the different experimental sites separately over bacterial (16S) and fungal (ITS) communities. Values at P < 0.05 are shown in bold.

Site		**Gleadthorpe**		**Harper Adams**		**Terrington**	
Barcode		**16S**	**ITS**	**16S**	**ITS**	**16S**	**ITS**
Numbers	**P**	0.634	**0.045**	0.202	0.430	0.561	0.561
	E^2	-	**0.573**	-	-	-	-
Types	**P**	**0.007**	**0.001**	0.192	**0.015**	0.158	0.129
	R^2	**0.379**	**0.360**	-	**0.309**	-	-
Abundances	**P**	**0.004**	**0.001**	0.247	**0.035**	0.305	**0.015**
	R^2	**0.406**	**0.463**	-	**0.359**	-	**0.309**

When the ASV data from all three sites were merged and normalised, there was no effect of any of the organic amendments on numbers of the 16S mASVs compared with the unamended controls or each other (Figure 41). However, there were again indications that the cattle slurry treatment may have increased fungal ITS mASVs compared with unamended control, broiler litter and pig manure treatments.

By far the largest effect on both 16S and ITS mASV richness was due to the trial site location (Table 12). Pairwise analyses showed that 16S mASV richness was significantly higher at Terrington than at Gleadthorpe and Harper Adams. ITS mASV richness in Gleadthorpe was higher than at Harper Adams and Terrington (Figure 42).

Table 12: Kruskal-Wallis results for Observed ASVs along with PERMANOVA results for Jaccard distances assessing the effect of spatial over bacterial (16S) and fungal (ITS) communities. Values at P < 0.05 are shown in bold.

		Spatial Variation	
Barcode		**16S**	**ITS**
mASV richness	**P**	**0.002**	**7.4E-6**
	E^2	**0.355**	**0.746**
Types of mASVs	**P**	**0.001**	**0.001**
	R^2	**0.317**	**0.353**
Abundances of mASVs	**P**	**0.001**	**0.001**
	R^2	**0.389**	**0.421**

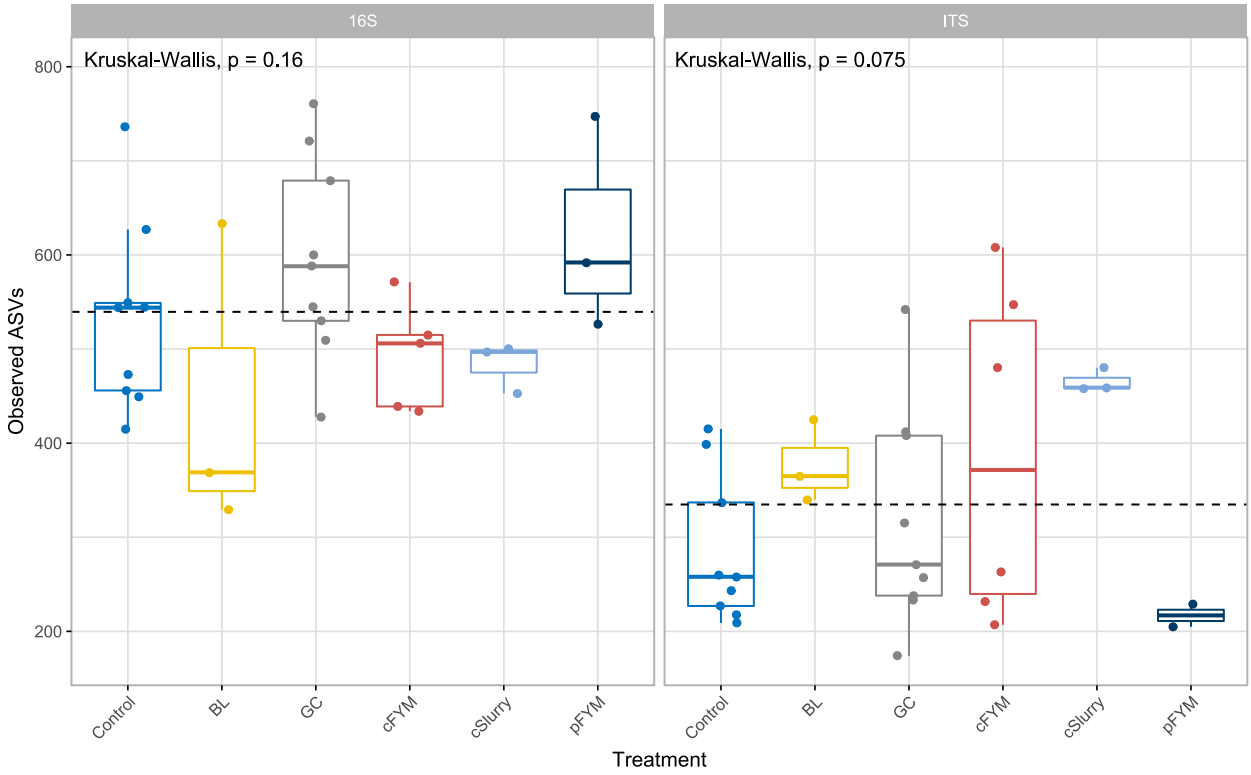


Figure 41: Absolute richness of 16S and ITS mASVs as affected by soil amendments with broiler litter (BL), cattle farmyard manure (cFYM), pig farmyard manure (pFYM), cattle slurry (cSlurry) or green compost (GC), compared with untreated controls.

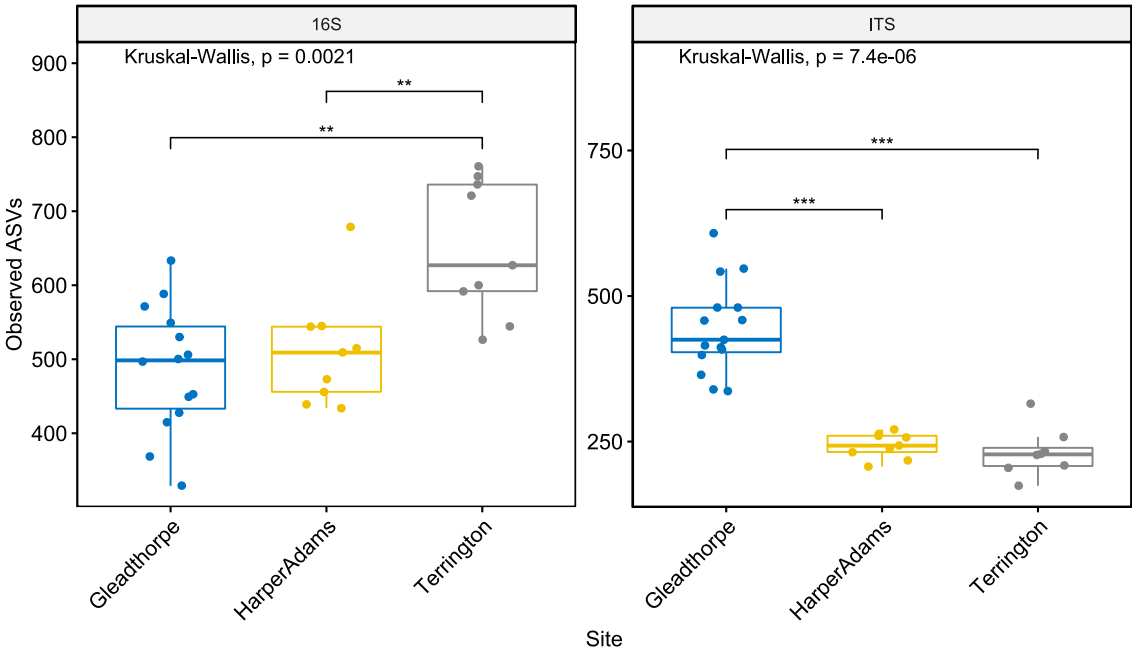


Figure 42: Absolute 16S and ITS mASV richness across all three trials. Kruskal-Wallis general and pairwise results are indicated where adjusted p-value was significant (0 '***', 0.001 '**', 0.01 '*').

Types: PERMANOVA analyses revealed significant differences between organic amendment treatments in the types of 16S and ITS ASVs present in Gleadthorpe, explaining approximately 37% of the variance in assigned bacterial types. Differences in the types of ITS mASV present were also significant between treatments at Harper Adams, explaining 31% of the variance in assigned fungal types ($p<0.015$; Table 11). No significant effects of organic amendments on types of 16S or ITS ASVs were observed at Terrington.

Further PERMANOVA analyses revealed larger significant differences ($P<0.001$) between the types of 16S and ITS mASVs present in soil from each trial site (Table 12, Figure 43). Approximately 32% of the variance in bacterial types and 35% in fungal types was explained overall by the site location.

Abundances: Significant differences were also observed between the treatments in relative abundances of types of ITS ASVs at all three sites and of types of 16S ASVs at Gleadthorpe only ($P<0.035$, Table 11). When the data were merged, it was clearly apparent that the location of the trial sites had a much greater effect on both types and relative abundances of both 16S and ITS mASVs than did the organic amendment treatments (Figure 43).

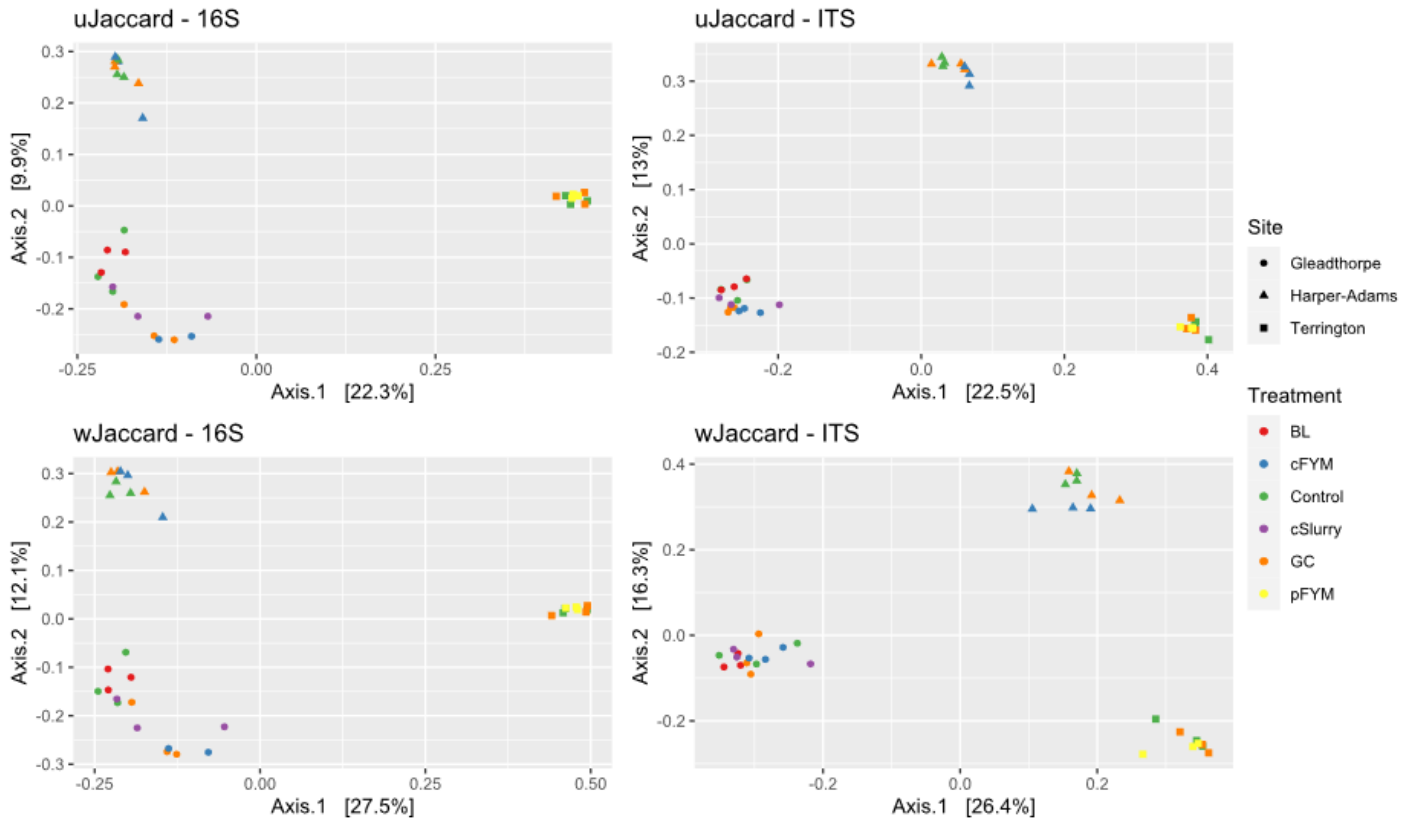


Figure 43: Principal Coordinate Analysis (PCoA) based on weighted and unweighted Jaccard distances matrices showing differences in relative abundances of 16S and ITS mASVs.

Long-term effects of re-ridging and inter-row companion crops or mulching on soil bacterial and fungal diversity in asparagus production

Estimated diversity of both bacteria and fungi were independently compared across the field trial (also described in Project 5), enabling comparison of any long-term effects on the microbiome of re-ridging between crops (Table 13) and any secondary effects of using inter-row companion crops or mulching with PAS-100 green compost compared with leaving bare soil between the asparagus beds (Table 14).

Table 13 Kruskal-Wallis results and PERMANOVA results for Jaccard distances comparing the effects of with and without re-ridging, for each secondary treatment and across all secondary treatments, on the diversity of bacterial (16S) and fungal (ITS) ASVs (values where P<0.05 shown in bold).

Treatment Groups		+ or - re-ridging Bare soil		+ or - re-ridging Companion crop		+ or - re-ridging Mulch		+ or - re-ridging All treatments	
Barcode		**16S**	**ITS**	**16S**	**ITS**	**16S**	**ITS**	**16S**	**ITS**
ASV richness	**P**	0.077	0.564	0.077	**0.043**	0.083	0.773	0.108	0.312
	E^2	-	-	-	**0.514**	-	-	-	-
Types of ASVs	**P**	0.21	0.259	0.119	0.304	0.205	0.343	**0.002**	**0.022**
	R^2	-	-	-	-	-	-	**0.067**	**0.055**
Abundances of ASVs	**P**	0.379	0.398	0.289	0.199	0.245	0.611	**0.012**	**0.031**
	R^2	-	-	-	-	-	-	**0.068**	**0.063**

Table 14: Kruskal-Wallis results and PERMANOVA results for Jaccard distances comparing the effects of with and without secondary soil treatments on diversity of bacterial (16S) and fungal (ITS) ASVs (values where P<0.05 shown in bold).

Treatment Groups		Bare soil (+ or - shallow disturbance)		+ or - Companion crop		+ or – Mulch		All treatments	
Barcode		**16S**	**ITS**	**16S**	**ITS**	**16S**	**ITS**	**16S**	**ITS**
ASV richness	**P**	0.157	**0.021**	0.289	0.564	0.564	0.773	0.494	0.203
	E^2	-	**0.722**	-	-	-	-	-	-
Types of ASVs	**P**	0.181	**0.036**	0.567	0.599	0.149	**0.049**	**0.009**	**0.001**
	R^2	-	**0.182**	-	-	-	**0.174**	**0.266**	**0.256**
Abundances of ASVs	**P**	0.117	0.061	0.572	0.329	0.203	0.362	**0.028**	**0.014**
	R^2	-	0.175	-	-	-	-	**0.267**	**0.261**

Numbers: No significant effect of re-ridging was observed on the total number of 16S ASVs (P >0.1; Table 13). However, re-ridging significantly increased the total number of ITS ASVs but only in samples where a companion crop had been grown between the rows (Figure 44). Shallow soil disturbance within the bare soil treatment also significantly reduced ITS ASV richness (Table 14). No effect of growing companion crops on the types of ITS ASV was observed.

Types: No significant effect of re-ridging on the types of 16S and ITS ASVs was observed in any specific treatment group (Table 13). However, when all sample were analysed together a significant effect was observed in types of both 16S and ITS ASVs with a marginal effect size of approximately 7% and 6%, respectively.

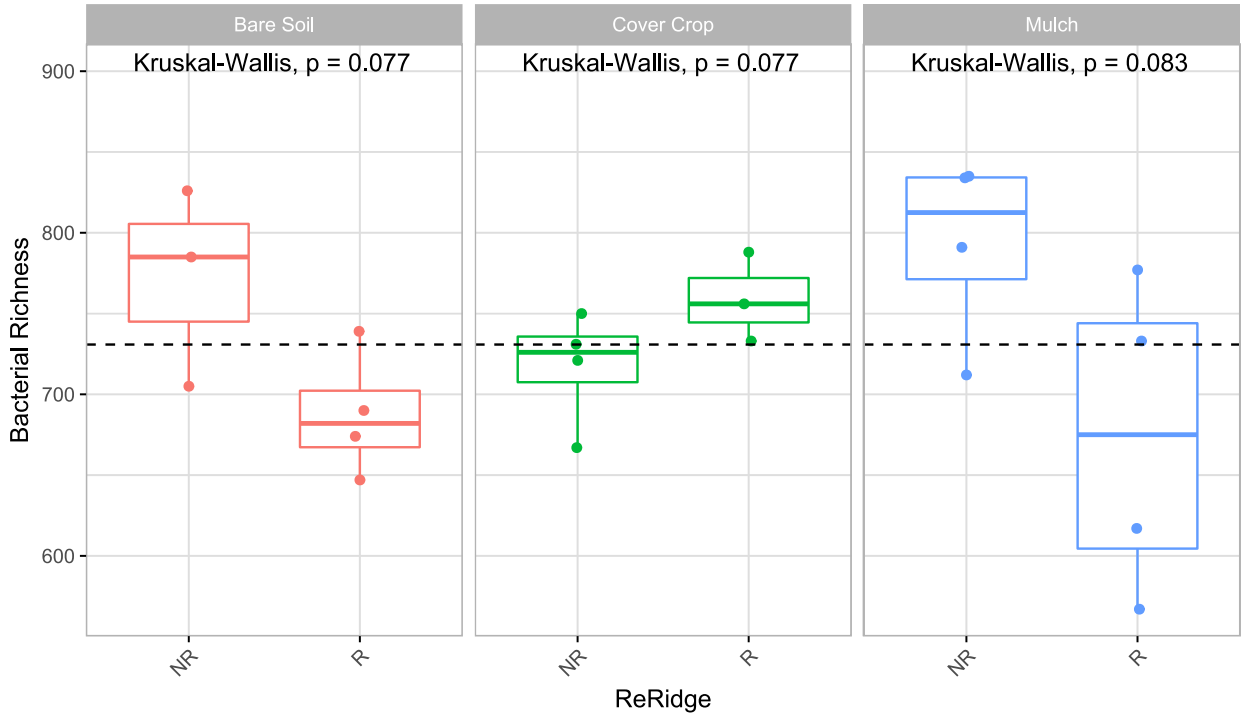
Abundances: Significant differences in relative abundances of both 16S and ITS ASV types, between both the re-ridging treatments and the secondary treatment groups, were also observed only when all samples were analysed together ($p<0.031$; Tables 13 and 14).

When the standardised approach was used to compare bacterial and fungal diversity, within equal numbers of sequences sub-sampled per treatment from the merged data set, the findings were similar (Table 15). There were no effects of treatments on the absolute richness of 16S or ITS mASVs, whereas the effects on types and abundances of both were highly significant. shows The community composition of bacterial and fungal microbiomes in these soils was roughly similar when mASVs were identified at phylum level (Figure 45). However, a high proportion of unique mASVs were associated with bare soil, mulch and companion crop treatments (Figure 46), with only around one quarter of bacteria and fungi unaffected by any of the three soil treatments.

Table 15: Kruskal-Wallis results and PERMANOVA results for Jaccard distances comparing the effects of all treatments on bacterial (16S) and fungal (ITS) mASV diversity (values at $P<0.05$ shown in bold).

		All treatments	
Barcode		**16S**	**ITS**
mASV richness	**P**	0.77	0.981
	E^2	-	-
Types of mASVs	**P**	**0.005**	**0.006**
	R^2	**0.119**	**0.108**
Abundances of mASVs	**P**	**0.005**	**0.018**
	R^2	**0.122**	**0.119**

(a)



(b)

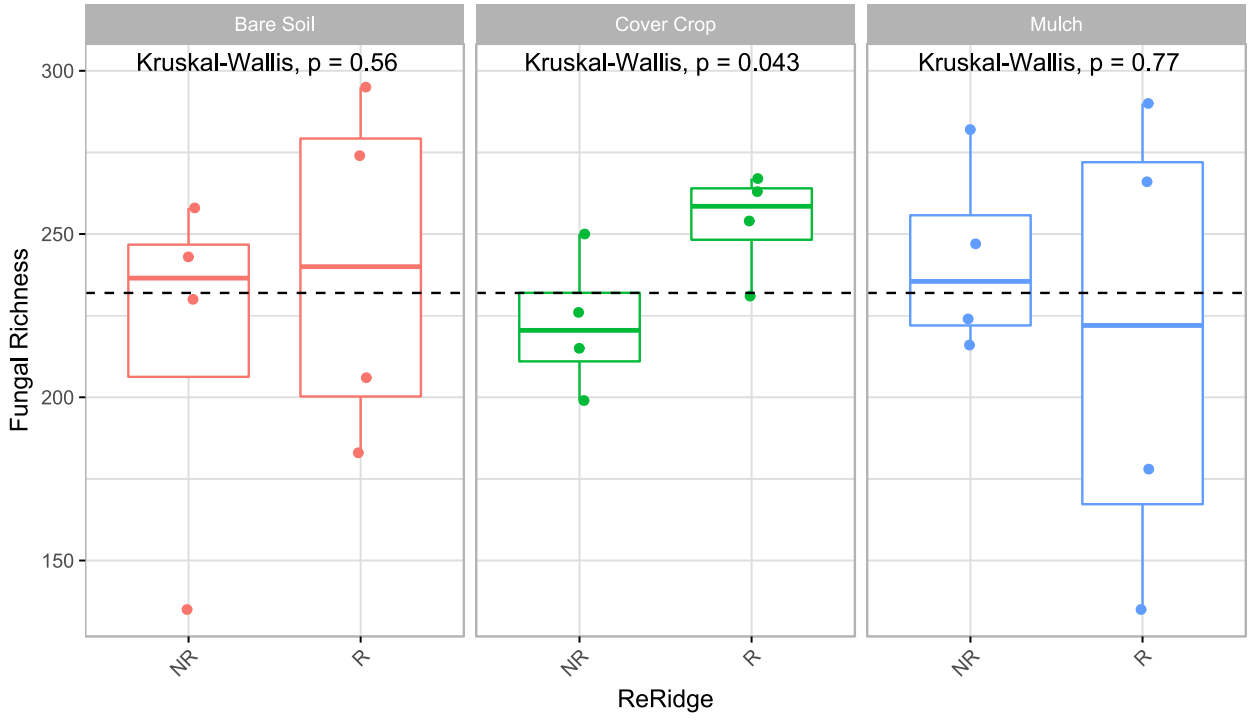


Figure 44: Numbers of (a) 16S and (b) ITS ASVs as affected by inter-row treatments of bare soil, companion/cover crops or mulch, with (R) or without (NR) annual re-ridging of the asparagus beds.

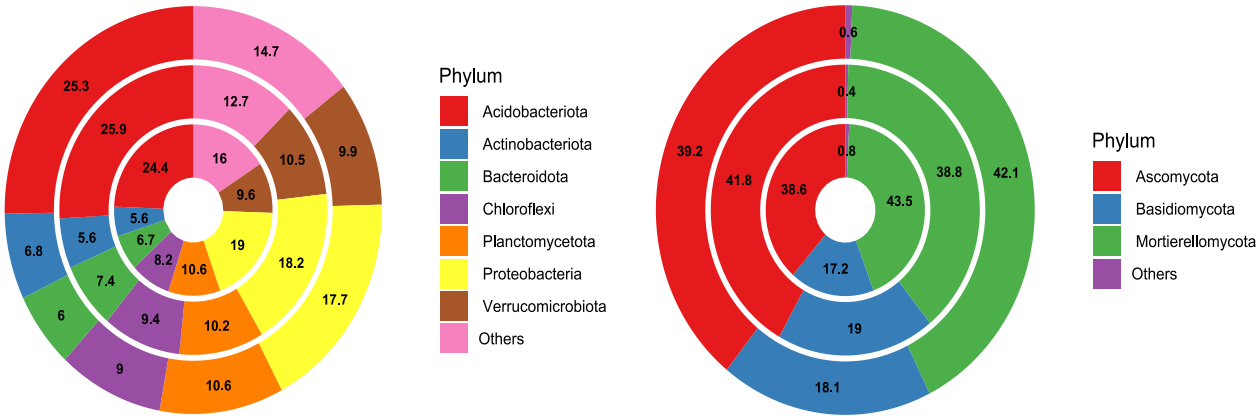


Figure 45: Comparison of bacterial and fungal soil community compositions across three treatment groups. Relative abundance of assigned phyla to (a)16S and (b) ITS mASVs across all samples from soils treated with mulch (outer), companion crops (middle) and bare soil (inner).

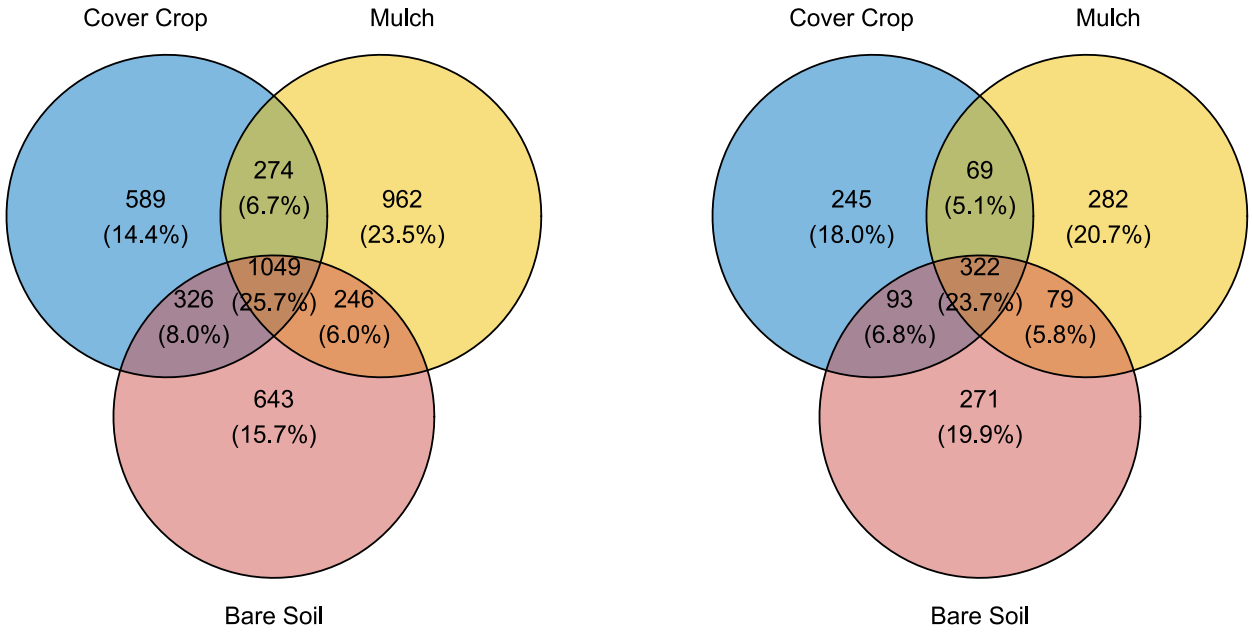


Figure 46: Commonality of (a) 16S and (b) ITS mASVs across three treatment groups.

4.4. Evaluating the outputs of the visual tool to show impacts of management on the soil biological community

Sense checking of the management impacts described was an important component of the development of the visual tool. Several data sets were used including the data from the longer-term experiments as part of the SBSH Partnership and data from a project collected from potato growers provided by AHDB. These data sets although not specifically collected to test the soil management tool gave good correlation with the predicted changes from the tool. The data for the incorporation of FYM into the Harper Adams experiment gave changes consistent with the soil management tool for High Carbon (C) Organic Matter (Figure 47). The more detailed data on the soil biological community highlights a small change in taxonomic diversity with organic amendments additions as already predicted by the descriptive model. Hence, one direct output is that no changes to the descriptive model are required from our findings.

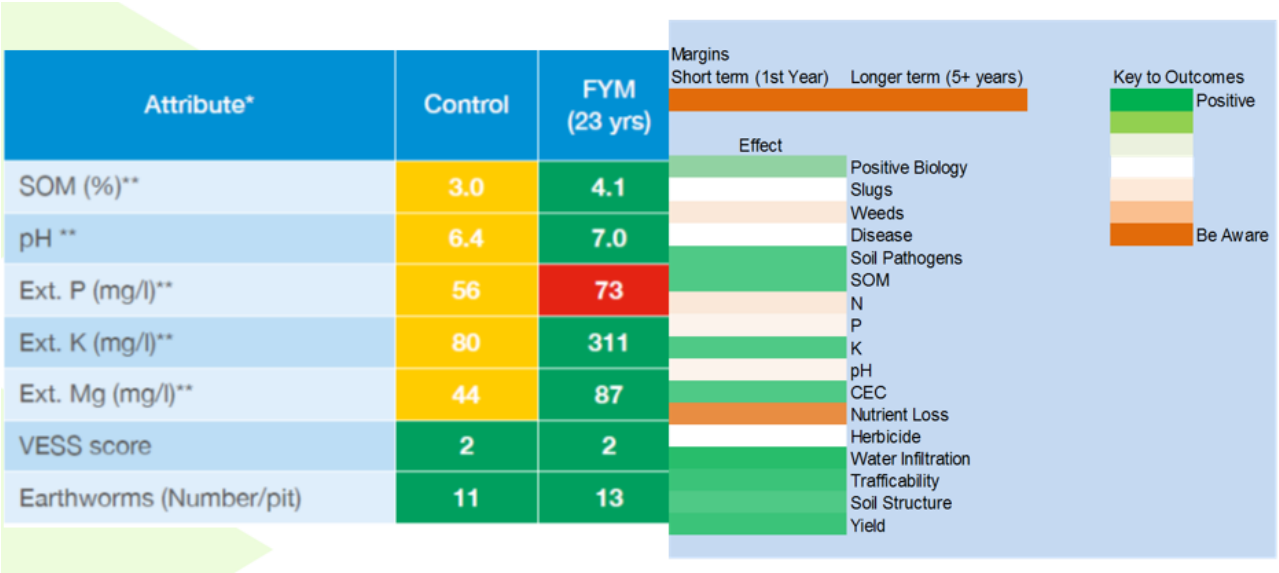


Figure 47. Harper Adams long-term FYM inputs score card compared with the outputs from the Soil Management Tool.

Earthworm numbers were combined with other soil biology to give the score in the soil management outputs after discussion with stakeholders and expert knowledge. Unfortunately, no assessment was made for some of the other outputs, but it was agreed that these indicated a sensible outcome when discussed with stakeholders.

An average of the potato grower's changes to management was used to compare with the management tool predicted outcomes and again, even though a more limited set of outcomes than the tool those that could be compared revealed good agreement (Figure 48).

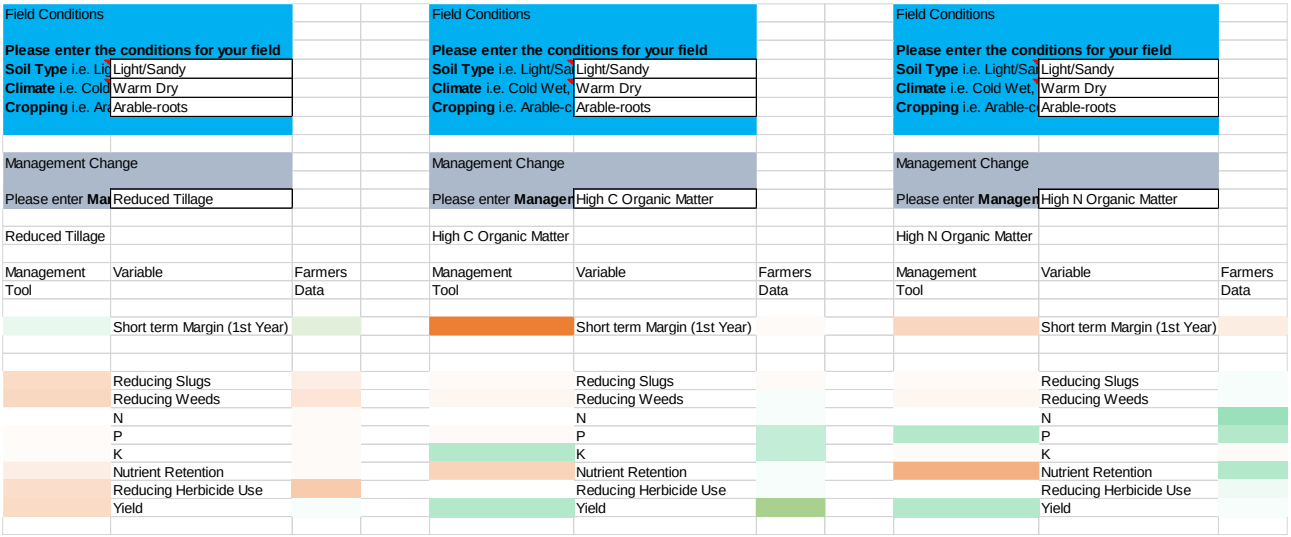


Figure 48. Comparison between average changes from farms with potatoes in the rotation and the outputs from the soil management tool.

Although the tool predicted a positive for phosphorus availability (P) with the addition of High C Organic Matter where the average for the farms was more neutral, again this was the case for N for the High N Organic matter in the management tool. This led to a greater weighting being placed on these nutrients in the tool.

Overall, the soil management tool performed well and provided useful indications for farmers and consultants considering a change to management with positives and challenges being highlighted. The tool would benefit from more options being added but these need to be balanced with the information available to provide the wide range of predicted outcomes from the published literature.

5. Discussion

5.1. Extraction of environmental DNA

Assessment of DNA holds much promise as a means to describe the microbial and mesofauna composition in complex matrices like soil. The use of eDNA approaches has been highlighted as a way to identify the habitats used by rarer animals that were otherwise unseen. In 2020 the technique was used to help track the COVID-19 virus (Randazzo et al. 2020) as a rapid alternative to more traditional DNA extraction. However, as shown in this project, for soil-based work the distinction between eDNA and total DNA extraction is confusing and semantic. Effective extraction of DNA from soils requires methods that also lead to cell lysis and hence the intra-cellular DNA of soil microorganisms is measured along with extracellular DNA. We suggest that reference to eDNA in association with soil is therefore only used in the context of studies of multicellular organisms in the environment, whose traces may be detectable even if the whole organism has moved on (Ruppert

et al. 2019). This approach is not yet fully developed with the science still in a discovery phase, and technological advances will continue to lead to new sequencing approaches and also to reduce the costs associated with sequencing and bioinformatics.

For soil-based studies, the major challenge in the analyses of the soil microbial community (and the analyses of microbiomes more generally) is the lack of standardised methods for DNA extraction and the analyses of data, especially as technology and data processing capability improve rapidly. This hinders comparisons between studies. Thus, care must be taken when comparing the effects of soil management across different studies. However, within the same studies, or where studies have used the same methods and analyses, it is possible to both infer and assess differences in soil communities and some of their functions directly, which can be used as a metric for soil health. This project has examined the approaches used to obtain DNA to determine what steps were required for optimisation of methods, and also considered the utility of DNA assessments as a part of soil health measurement. Here, we showed that it is possible to detect DNA from microbes and mesofauna from a range of soils of variable types and subject to different treatment regimes. Vestergaard et al. (2017) also noted that the DNA extraction method used determined the inherent bias involved and that no method comes without limitations. Nilsson et al. (2018) highlighted the need for reproducibility and public data availability in the study of fungal communities. Djemiel et al. (2022) highlight the importance of method standardisation (like the ISO 11063: 2020 standard, which has its own limitations) in order to provide consistent results across studies, more particularly when measurements are interpreted as bioindicators of soil health.

5.2. Effects of soil management on soil biological communities

5.2.1. Soil microbiome

Studies on the effects of soil management practices on bacterial and fungal communities have often reported contradictory results. This may be due to the wide spatial variation in richness and diversity of microbial communities in soils. In a recent study across 12 European long-term experiments, Hannula et al. (2021) reported that over two thirds of the fungal species described were unique for each of the countries involved. Moreover, some soil management practices had a variable effect on the diversity of fungal communities depending on the site location. In this project, the single most important factor driving the composition and structure of both bacterial and fungal communities was geographical location. However, it is also possible to draw out some smaller impacts of soil management.

Effect of long-term pH management on the soil microbiome

Much is already known of the effect of soil pH gradients on bacterial and fungal community structures (Fierer and Jackson, 2006, Rousk et al., 2010). The impacts of pH on bacteria are usually more marked than those on fungi, corresponding with the wider pH ranges observed for optimal growth of the fungal community. A recent metabarcoding analysis showed fluctuations of microbial diversity in 436 locations in 7 different temperate ecosystems across Wales, UK (George et al., 2019). In their analysis, pH was the best predictor for bacterial diversity and second best for fungal diversity. This work has confirmed those findings using the standardised metabarcoding approach at the controlled experimental site at Craibstone, which included pH levels across a range including a very low pH typical of degraded soils (Msimbira and Smith, 2020), as well as pH levels more typical of UK agricultural systems. In response to pH, strong divergent trends in all three measured diversity metrics (absolute richness, relative richness and relative abundance) were confirmed for both bacteria and fungi/oomycetes. These findings confirmed that changes in soil pH can produce strong changes in community composition, especially where soils become strongly acid.

The sequence data obtained was also successfully used to identify potential bacterial biomarkers within the bacterial and fungal communities that appeared to fluctuate most in response to pH levels. A larger number of bacterial phyla were associated with pH 4.5, but more fungal phyla biomarkers were found at pH 7.5. However, when we analysed the total number of biomarkers across all taxonomic levels, we found significantly more bacterial biomarkers at pH 7.5 than at the other pH levels, whereas similar numbers of fungal biomarkers were found at each pH level. This suggests that the total number of biomarkers is a good indicator of the strength of selection pressure exerted by a given treatment on biological communities. This may provide a quick way of screening a large set of samples subject to a range of treatments to identify those where more detailed microbial community analysis should be targeted.

The project also used the sequence data obtained to predict functional divergence within the distinct soil microbial communities found in different pH conditions. The distribution of different assigned fungal guilds was found to differ with soil pH, with symbiotrophs significantly more abundant at more neutral pH (6.0 and 6.5), while pathotrophs were more abundant at pH 7.5. Bacterial functionality predictions were still too inconclusive due to the lack of taxonomic annotation at species level amongst the available databases. This was one of the first attempts to use this kind of approach in assessment of soil health. Although it was only partially successful here, prediction of microbial function in soil communities will improve both with increasing availability of taxonomic annotation and with higher resolution sequencing approaches. Nonetheless, these data add to the evidence that changes in microbial community diversity will result in functional changes in agricultural soils.

Effects of long-term inorganic fertiliser application on the soil microbiome

There is a lack of precise information on the long-term impact of fertilisers on microbial communities. Previous studies have attempted to assess microbial fluctuation with different fertiliser treatments, but their conclusions are conflicting (Zhong et al., 2010). Positive effects are often related to increased microbial nutrient availability from fertilisers directly sustaining increased populations, and also indirect effects as a result of improved plant rooting, growth and increased root exudation (Lehman et al., 2015). There may also be other indirect effects of fertiliser applications such as soil acidification (Kidd et al. 2017), but management intervention at the Craibstone site has prevented such impacts. The long-term effect of fertiliser application over almost a century studied here showed that the relative abundance of the species within fungal community may have been affected rather than in the richness or types of species present. No particular phylum was significantly affected by the regular use of fertiliser, in contrast with observations in other studies (Peine *et al.*, 2019; Silva et al., 2017). Discrepancies between the findings of different studies may be due to management differences and may also arise due to methodological differences, including sampling depth, molecular and bioinformatic approaches, and sequence databases used in the analyses, highlighting the necessity for a standardised pipeline that can be used to compare diverse studies of soil microbial communities from different locations.

Effect of long-term organic amendments on the soil microbiome

Studies on the effects of organic soil amendments on bacterial and fungal communities have also reported contradictory results (Li *et al.*, 2019). Hannula *et al.* (2021) showed that some soil management practices had a variable effect on the diversity of fungal communities depending on the site location and their study found no significant effects of different sources of organic amendment. In this study, we evaluated the effect of organic amendments across three long-term trial sites on soil bacterial and fungal communities. As with our results of the effects of long-term inorganic fertiliser application, the organic amendments results appeared to affect the fungal rather than the bacterial community. These results suggest a higher sensitivity of fungal communities to organic/synthetic fertiliser application, at least when sampled after harvest. However, the most relevant factor driving the composition and structure of both bacterial and fungal communities was the geographical location. The three experimental sites are distantly located, and their soil types differ with loamy sand (6% clay), sandy loam (12% clay) and silty clay loam (28% clay) at Gleadthorpe, Harper Adams and Terrington, respectively. Interestingly, the largest effect of organic amendments on microbial diversity was observed at the Gleadthorpe site, suggesting that low clay soils may be more sensitive to these treatments. This may be due to the impact of clay content on water retention in soils; the Gleadthorpe site is known to be dry, and this is likely to have an impact on soil biology. In addition, a high clay content is thought to negatively affect DNA extraction from the soil matrix (Högfors-Rönnholm *et al.*, 2018). The average DNA yield obtained from Terrington

samples was the lowest (42 ng/µL), although average DNA yield from Harper Adams (51 ng/µL) and Gleadthope (48.6 ng/µL) were similar even though soils at Harper Adams have about double the clay content. Clay content did not, however, appear to affect the diversity metrics, for example, Terrington samples showed the highest bacterial richness.

Long-term effects of re-ridging and inter-row companion crops or mulching on soil bacterial and fungal diversity in asparagus production

Soil treatments in the asparagus trial were mainly designed to alleviate the effects of compaction, a common problem during highly mechanised asparagus production. Results from our study have also shown that these treatments also have some influence on the soil microbial community. Although the re-ridging treatment appeared to have no effect on absolute richness of bacterial or fungal ASVs, there was some evidence that shallow soil disturbance had some effect on reducing fungal richness at the time of sampling. Other interacting factors such as the use of companion crops and mulches appeared to have variable effects on the types and relative abundances of different bacterial and fungal taxa.

5.2.2. Analysis of management impacts on the soil microbiome – standardised approach allowing integration across different studies

A standardised approach was formulated and tested here for the first time to allow data from multiple studies to be merged and normalised in such a way that the effects of multiple factors can be evaluated on members of the soil microbial community that are common to each study. Many studies on soil biology have reported the effect of agricultural management practices, however the implementation of available diversity indexes has differed between studies. For that reason, a standardised analysis approach was proposed and implemented in this project. This has demonstrated that this enables a robust consideration of which bacterial and fungal diversities change in response to key soil managements across time and spatially dispersed datasets.

Merging ASV tables enables control for the effects of normalisation (i.e. all data are normalised together in the same way), and then more holistic analyses can be conducted, including for example identification of differences in the direction of effects and clusters of highly similar ASVs between studies. Together this allows comprehensive comparisons of the nature of effects and the types of taxa that are in common (or different) between different treatments in different studies. The identification of matching ASVs is possible because identical DNA sequences underlying ASVs produce the same MD5 identification tag between sequence runs with dada2 making ASVs from independent trials directly comparable, but only if sequences are in the same orientation and trimmed/truncated to the same length and sites. However, it is highly likely that independent studies will not have trimmed/truncated ASV reads to the same length and sites unless there has been previous coordination. To understand which highly similar ASVs are common across independent studies, a methodology was applied that clustered and analysed ASVs in the merged table into

mASVs, each of greater than 98% identity. Crucially, the analyses of mASVs then allowed robust hypothesis testing of the impact of management on soil microbial communities by comparing the differences and similarities of communities between studies.

In the first instance, the standardised DNA metabarcoding method was able to assess agricultural soil microbial diversity across two long-term agricultural trials that have consistently manipulated pH and fertiliser input for over 60 years. Comparisons of the analyses from both separate and then merged data confirm that long-term independent pH and fertiliser managements have applied differential sustained selection pressures that have resulted in diverged and differentially adapted soil microbial communities. Both analyses showed that pH but not fertiliser significantly affected fungal and bacterial taxa and both pH and fertiliser significantly affected the types of bacterial and fungal taxa and their relative abundances. The effect of fertilisers on the types of fungal ASVs was only found to be significant when the merged data were analysed. Importantly both individual and merged analyses showed the effect size due to pH to be approximately 6-times larger than the effect of fertiliser additions. Inferences of the sizes of effects between studies will be particularly useful to increase an understanding of the relative size of effects of specific agricultural and soil treatment on soil biological communities.

Analysis of 16S and ITS ASVs in this study showed that the composition of types of bacterial and fungal species identified as ASVs differed between the two trials. A majority of both 16S and ITS ASVs were unique to the pH trial and almost half of the ITS ASVs were unique to the fertiliser trial. There could be methodological reasons for these dissimilarities between trials including differences in sampling times and differences in performance between different sequencing runs. The latter was addressed to some extent by the implementation of the scaling with ranked subsampling (SRS) normalisation method. This performs a similar approach to traditional rarefaction (reducing all sample reads to a given sampling depth) but with a more conservative approach in terms of preserving ASVs richness (Beule & Karlovsky, 2020). Nonetheless, the ideal normalisation method with different ASVs library size continues to be a topic of discussion (McMurdie & Holmes, 2014; Weiss *et al.*, 2017). The approach taken in this study was to first merge datasets from both trials and normalise the merged dataset by SRS. Differences between the trials was then analysed using the standardised pipeline to compare the effects of the pH and fertiliser treatments with the results obtained from the separated datasets.

The same standardised approach was used to compare data from a further three trials designed to investigate the long-term effects of repeated soil amendment with organic matter from various sources that were more geographically separated (at ADAS Gleadthorpe, ADAS Terrington and Harper Adams University). In this case, analysis of the merged data clearly showed that, although significant differences were found in the types and relative abundances of 16S and ITS mASVs between amended and non-amended soils, these effects were much smaller than the differences

observed between locations of each trial. It was therefore concluded that whilst long-term independent organic amendments had applied differential selection pressures that resulted in diverged soil microbial communities, the magnitudes of these differences in soil biology created as a result of agronomic managements were much less than those observed between the natural biological communities adapted at each of the locations. It is assumed that differences in both soil types and conditions at sampling may have contributed to the differences observed between locations. These results indicate the true value of being able to merge datasets from multiple studies, rather than to extrapolate findings from any individual study, when forming general recommendations on the effects of soil management on soil microbiology.

Many individual studies on soil biology have reported the effect of agricultural management practices, however the implementation of available diversity indexes has differed between studies. The importance of the standardised approach to DNA analysis proposed here is multiplied where there are many different studies with greater expected variation than that observed between only the few sites studied here. For that reason, we have proposed a standardised combined metabarcoding approach and have demonstrated that it is viable to use it to test the extent to which bacterial and fungal diversities change in response to key soil managements and the need to assess the magnitude of those changes across multiple datasets.

5.2.3. Nematodes

Site differences (soil type, climate and cropping) dominated the measured differences in nematode communities. There were relatively few differences between treatments at any individual site and where differences occurred, most were associated with differences in the enrichment index, rather than the structure index (*e.g.,* pH at Craibstone). Differences did not always appear to relate directly to the changes in the underlying microbial community (Section 5.2.1) which may be a result of interacting tritrophic effects in the soil food web (Wardle et al. 1995).

The enrichment index reflects nematode species that have a rapid turnover rate, are able to respond quickly to changes in management, particularly in relation to nutrient inputs (Ferris et al., 2001). In these experiments, the enrichment index highlighted treatments dominated by bacterial feeders suggesting a surplus of nutrients. For example, although there was no consistent effect of organic amendment addition across the long-term sites, there was some evidence that repeated slurry additions at Harper Adams resulted in a more enriched community dominated by bacterial feeders, compared with a lower enrichment index on the control treatment (with both treatments having similar structure indices, measured in 2017). At Craibstone, particularly high enrichment indices were observed at the low pH levels under the oat and potato crops. The reasons for this are unclear but may be linked to lower nutrient uptake by poorly performing crops at low pH levels and the high extractable P concentrations measured in these soils. Tillage had no effect on nematode community

structure at Loddington, which had a relatively stable nematode community with no change in measured indices between samplings and which is close to the optimum for good biological health.

The highest structure indices were found in the grass leys on the Craibstone pH experiment and at Harper Adams (measured in 2017), but this was not the case for the grass ley in the old rotation experiment at Craibstone. Here the enrichment index was also exceptionally low, with very few species in this class observed. Exceptionally low enrichment indices were observed at Gleadthorpe, which may reflect the light texture and low organic matter content, with this site reporting the lowest cereal grain yields (Project 4 report). However, both the Craibstone rotation and Gleadthorpe organic amendment experiments were sampled in the same year (2019). During this time there was a change of personnel within the project delivery team, which led to a 4-5 month delay in extracting the samples (whilst training was undertaken). Although samples were appropriately stored (< 4°C), some deterioration could have occurred, which may have preferentially affected those species associated with the enrichment index.

Nematode communities were most influenced by site factors such as soil type and land use, with the most stable communities found under grassland. Management factors appear to have a smaller effect, particularly on those species contributing to the structure index; the enrichment index shows some variation associated with nutrient loading and pH. The project has confirmed previous studies that show that the nematode community composition is important and can indicate management impacts. However, the work has also shown that site has an overriding effect and hence it is not possible to compare measurements from different soils under different rotations and conditions. This may result in circumstances where the same management treatment implemented on different sites has different impacts on the nematode community.

5.2.4. Mesofauna

As with the nematode population there was a clear effect of site and sampling year on mesofauna communities, with most sites being reasonably well separated from each other in a NMDS analyses conducted on total numbers of major groups of soil mesofauna. The main axes produced by NMDS analysis were strongly correlated with the total population size, but in particular, associated with the population of Entomobryomorpha and Poduromorpha springtails. These two orders of springtails are almost universally found in soils and represent a wide range of species with various adaptations to different soil and litter environments. However, all are decomposers, feeding on decaying organic matter, and all are relatively soft-bodied (poorly sclerotised) animals.

Differences observed between soil mesofauna communities, and correlations between characteristics of these communities and other soil properties, can be interpreted with reference to the likely life-strategies of the organisms involved. The Universal Adaptive Strategy Theory (Grime

and Pierce, 2012) suggests all organisms must adopt life strategies that balance the need for growth, maintenance or regeneration and that the strategies of the organisms present will reflect the prevailing ecological conditions or resource availability. Although it has been primarily applied by developing the concept of C, S and R characterising the strategies of plant species (Grime, Hodgson and Hunt, 1988) the system can also broadly be applied to soil mesofauna. S-selected species put more effort into maintenance (including defence) and are stress-tolerators, adapted to low resource availability or harsh conditions by slow growth, long lives, and often investment in defensive structures (e.g., sclerotisation). They occupy niches that cannot support other more competitive species. R-selected species are ruderals adapted to frequent disturbance or temporarily available resources, through rapid breeding, high fecundity and high dispersal ability, at the expense of very little defensive investment, and occupy niches which more competitive species fail to access in time. C-selected species are competitors, adapted to outcompeting others in situations where resources are abundant by rapid increases in population, rapid growth, often larger body size, high fecundity and moderate defences. Species may have traits on a continuum between the extremes of these strategies and may change strategies through their life cycles. Perhaps more among animals than plants, it is possible for organisms with different strategies to co-exist in the same community, with their populations reflecting the relative availability of their preferred resources. A range of resources can develop in soils because decomposer communities, besides using organic matter as energy, also transform organic matter into different forms, for example turning plant material into springtail biomass and dung pellets or wood into fungal biomass. These different forms present new opportunities for different life strategies.

Most springtails are arguably C or R selected, in that they have reasonably high fecundity, short life cycles, and are fairly poorly defended against predation and have only moderate dispersal abilities. Their predators, the mesostigmata are C selected, but must outlive their prey and so are slightly better defended, with more stress tolerant traits, such as reasonably high levels of sclerotisation. Oribatida mites are classic stress tolerators, in that many species grow and mature slowly, in habitat where resources are limiting or harsh, and they invest a lot of energy in producing strong sclerotised exoskeletons for defence.

The agricultural systems examined during this study were disturbed systems, so all organisms identified are likely to show ruderal traits, favouring those organisms capable of rapidly occupying and breeding within the soil. For example, very few oribatid mites were found compared to less disturbed habitats such as forest soils and leaf litter. In this context, differences between sites and samples in the total numbers of poorly-sclerotised Entomobryomorph and Poduromorph springtails (NEPS), can be interpreted in terms of the availability of readily decomposable organic matter resources that these animals need to breed. High scores on NMDS axis 1 in the full dataset, the reduced data set and high values of NEPS show a high abundance of resources in the form of

decomposable organic matter, suitable for growing biomass. This relationship is supported by the very strong significant positive relationships for NMDS axis 1 (for both full and reduced data sets) with NEPS and CO_2-C burst. Symphypleona springtails, which may be aboveground foliage dwellers and were present in low numbers, were less useful as soil indicators. Neelipleona springtails were more associated with more S-selected communities and responded in similar ways as did Oribatida mites at Harper Adams.

The variability in total population between years makes indicator metrics based on total numbers of mesofauna difficult to interpret. However, at the one site where 2 years of comparable data are available, numerous other parameters relating to soil health changed between the 2017 and 2020, as a result of rotational management, with the samples in 2017 taken at the end of a 2 year grass/clover ley, prior to cultivation and drilling of a winter wheat crop, followed by 3 years of plough-based arable cropping, including potatoes in 2019. This is discussed further in the Project 4 report and clearly demonstrates the benefit of having grass leys in the rotation. However, more data comparing between years for sites with ongoing similar management is needed to determine the impacts of weather conditions on soil mesofauna populations, to enable better understanding of their relationship with soil health.

These data indicate that total population density of C-selected springtails (NEPS) and percent of collembola that are springtails may be useful simple metrics to indicate the extent to which soils have received additions of decomposable organic matter, with associated higher microbial and earthworm activity, better soil structure. These indicators (total NEPS and % Collembola) are necessarily based on a low level of taxonomic understanding, since for most sites identification was carried out only to the level of major taxon group. The apparent focus on Collembola should not suggest that they are more important than acari in the community (which are usually numerically dominant), since at least one parameter (%Collembola) could equally well be expressed as %Acari, showing equally powerful, but reversed, statistical relationships. Furthermore, acari such as oribatida, which were present only in very low numbers in this study, are likely to prove better indicators of soil health and function in less disturbed habitats.

Setting the experimental plot data from the experiments included in this study into the context of the national community variability maintained many of the relationships observed within the experimental studies alone, but to some extent combined the axes to represent a more general trend linking mesofauna communities to organic matter content. It is interesting to note that, as for the microbial results, Gleadthorpe exhibited some of the strongest differences observed in response to the different organic amendment treatments (control, manure and compost). Soils at Gleadthorpe had the lowest clay content and were significantly drier soils compared with Terrington and Harper Adams. Therefore, in the absence of organic amendments, competitive springtails are likely to be

unable to flourish as they require near 100% humidity conditions. Addition of organic matter improved water content in all these sites, but the impact was greatest at Gleadthorpe, as a result of the strongest amelioration of conditions.

The study at Craibstone however, suggests that more sensitive effects of management may be picked up where mesofauna are characterised at higher taxonomic level, with these indicating the shorter-term impacts of different crops on soil conditions. It is interesting that pH appeared to be less important in controlling mesofauna communities, than the current crop or stage in the rotation, which would both reflect the quality of organic matter and the recent history of disturbance.

Identification of mesofauna to family level, as with nematodes, requires specialised identification skills. The development of DNA-based methodologies potentially provides a solution to this problem (e.g. George *et al.* 2019). For this study, however, at the sites where co-investigation of mesofauna by sequencing and microscopy took place, there was a poor correlation of the population metrics. The poor correlation may be the result of community heterogeneity, in that metabarcoded samples may have had different communities to those identified under the microscope. It might also result from microscopic misidentification of specimens, or misidentification of specimens in the barcode libraries used to match to the metabarcoding results. The DNA extraction was carried out on bulk soil, which is known to provide or poor amplification of soil mesofauna, when compared to metabarcoding carried out on Tullgren extracts. Finally, metabarcoding of bulk soil captures both living and dead organisms or parts or organisms. This could result in the detection of more diversity in less biologically active soils, where whole or parts of dead organisms may persist for longer in the soil without decomposition. Approaches using COI barcoding for soil mesofauna communities continue to be developed, and are likely to provide a useable, if imperfect alternative to microscopy. Kirse et al. (2021) recently presented an evaluation of methods for use of metabarcoding approaches for mesofauna biodiversity assessment from soil samples and many unanswered questions remain about how best to deploy DNA-based methodologies, not least the higher proportion of unassigned OTUs currently associated with COI than with 18S (annelids, nematodes) markers.

5.3. Evaluating new biological indicators for soil health

Soil communities are inherently complex, variable and resilient, and this project suggests that they only appear to change community structure in response to extreme treatments. The project applied DNA-based approaches to examine difference in microbial communities and has strongly validated this method as an approach for research studies seeking to disentangle how management drivers lead to changes in soil function. The findings in this study showed similar overall patterns to those determined in other studies for the impact of pH, and also revealed more of the detail showing how changes in the composition of the microbiome may affect decomposition and soil-plant interaction.

Using a new standardised approach to merging taxonomic datasets (ASV tables) allowed the nature of effects to be compared between different treatments in different studies and showed that that there is less impact on the composition of the soil microbiome from management practices, like organic amendment or tillage compared with the impact of long-term differences in pH. However, use of the on-line databases highlighted the large gaps in the records for soil organisms, and hence this approach cannot currently be deployed to link diversity and function and thus to support the development of soil-improving practices. Data from this project have enabled some updates to these shared databases and we expect the underpinning science to continue its rapid advancement, so that DNA-based approaches measuring the soil microbiome may be integrated into soil health monitoring for on-farm use in the future.

Whilst the community analysis approach for nematodes and mesofauna offers some insights into the biological health of a soil, these are not yet available on a routine commercial basis, as the methods used to extract and identify nematodes and other mesofauna are laborious requiring specialist identification skills. However, DNA-based methodologies for identification and quantification of the soil nematode community are developing rapidly and this may offer a more practical and affordable way for this kind of analysis in the future. Griffiths et al. (2018) compared morphological identification approaches to molecular approaches using high throughput sequencing for extracted nematodes directly and also with DNA extracted from soil. They showed that nematode community analysis differed with extraction methods as also shown by Quist et al. (2016). The profound differences in nematode community structure may result from identification skills and/or PCR biases; the soil extraction approach is also limited by the sample size (10 g); a sample size of 200 g is recommended to reliably reveal a soil nematode community (Wiesel et al., 2015). Recent work (e.g. Kawanobe et al. 2021) has focussed on the development of molecular methods, including identification of nematode specific primer sets (e.g. F548_A/R1912), to identify nematodes both from soil extracted DNA and extracted nematodes. Similar DNA-based approaches for arthropods are less well developed but are being established rapidly (Kirse et al. 2021) and are also very likely to become available for research and monitoring in the near future.

DNA methods will be increasingly used because of reducing analysis costs, high throughput, greater taxonomic resolution and compatibility with available technical skills. There is a need now to further understand the methodological discrepancies identified here (e.g. sample size; extraction and PCR biases; primer specificity; read number and taxonomic resolution). Du Preez et al. (2022) continue to highlight the importance of both improvements in methods and standardisation; in particular that calibration of molecular to morphological information is still urgently needed. For all the soil ecosystem-focussed approaches assessed in the project, the main driver of community composition was site rather than management. Since molecular data are highly variable from site to site, molecular analyses of the soil ecosystem are not suitable to assess the general effects of soil

management across a wide range of soils and locations, and cross-site benchmarking seems unlikely to be of value. The work carried out within the project has enabled significant progress in increasing standardisation; ease of use; and ease of interpretation of analysis of the soil microbiome. Whilst this project did not address seasonal variation, Hannula et al. (2019) observed the highest bacterial and fungal abundances in November, just after the end of the growing season. Hence the recommended timing for sampling for the Soil Health scorecard (autumn as the soil wets up) will coincide with these maximal populations. However, the costs of measurement are still very high for each sample and no UK-wide benchmarking framework can currently be established, hence we do not currently recommend the use of soil ecosystem-focussed indicators for routine soil health monitoring on farm.

As a trial venture, these approaches have already been used to assess samples of arable soils submitted from growers (including samples from the field sites selected in this project) through the Big Soil Community initiative at Fera Science Ltd. (https://www.fera.co.uk/news/big-soil-community). This involves a community effort between growers, agronomists and Fera scientists to investigate the diversity of UK soils and understand how it can affect crop production and long-term soil health. Initially costed at £250 per sample, participating growers are aware that they are contributing to the development and interpretation of the tests as well as understanding the diversity of their soils. It is anticipated that the cost per sample will fall as interest in the scheme increases. Sustainability of the scheme is dependent on the usefulness of the information to each grower. Furthermore, standardised analytical procedures developed in this project, that allow analysis of merged DNA sequencing data, will permit comparisons of field assessments from multiple sources, even when there is variation in the methods used to collect the data. This will open the investigation of factors contributing to biological soil health to a much wider breadth of soils and agroecological systems both across the UK and internationally.

Automated methods that identify key taxa and their relative abundances from the soil DNA, developed in this project, are key to increasing the value of information that can be fed back to growers. For example, the methods that can predict functions associated with taxa identified within the soil microbial communities help to add practical value to the results by not only identifying the most abundant organisms but also estimating whether they are providing key ecological services and whether they may be harmful or beneficial to crops. It is expected that this approach will facilitate the identification of bioindicators common to all samples, allowing the development of more targeted field tests that could monitor changes in the behaviour of these indicators as influenced by factors such as cropping practices, soil management, soil types and climate. More research is also needed to link soil function (e.g. nutrient supply, structural stability, GHG emissions) with ecosystem-based measures of the soil biological community. Therefore, the best way to use these indicators to provide additional detail about the soil ecosystem, and the factors affecting soil functions, would be to set up

benchmarks at any specific sites where soil function is under study and then monitor in a consistent and carefully structured way over time alongside soil health monitoring.

The detailed morphological analysis and molecular methods (DNA extraction and metabarcoding) used within this project have delivered an increased understanding of soil biological health, in particular the interacting impacts of both site and soil management. The simple visual tool was developed at the start of the Partnership to combine the state-of-the-art knowledge about the impacts of management of soil health and productivity from literature with that held informally by experts Reflection on the data collated on the data on soil biological health together with the broader soil health data sets (presented in Project 4) has supported the embedded qualitative relationships, the descriptive visual model is robust for use to explore the impacts of management on soil biology and soil health.

6. Acknowledgements

The authors wish to thank all the experimental site field teams at ADAS, GWCT, SRUC and Harper Adams and laboratory technical support teams at ADAS, Fera, University of Lincoln and SRUC.

7. References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) 'Basic local alignment search tool', *J Mol Biol,* 215(3), pp. 403-10.

Anderson, M. J. and Walsh, D. C. I. (2013) 'PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing?', *Ecological Monographs,* 83(4), pp. 557-574.

Andre, H., Ducarme, X., Anderson, J., Crossley Jr, D., Koehler, H., Paoletti, M., . . . Lebrun, P. (2001) 'Skilled eyes are needed to go on studying the richness of the soil', *Nature,* 409(6822), pp. 761-761.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Barrios, E. (2007) 'Soil biota, ecosystem services and land productivity', *Ecological Economics,* 64(2), pp. 269-285.

Beule, L. and Karlovsky, P. (2020) 'Improved normalisation of species count data in ecology by scaling with ranked subsampling (SRS): application to microbial communities', *Peerj,* 8.

Biswal, D. (2022) 'Nematodes as ghosts of land use past: elucidating the roles of soil nematode community studies as indicators of soil health and land management practices', *Applied Biochemistry and Biotechnology*. 194, pp. 2357–2417.

Blair, J. M., Bohlen, P. J. and Freckman, D. W. (1997) 'Soil invertebrates as indicators of soil quality', *Methods for Assessing Soil Quality*, pp. 273-291. (eds J.W. Doran and A.J. Jones). Soil Science Society of America, Inc. https://doi.org/10.2136/sssaspecpub49.c16

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., . . . de Bruyn, M. (2014) 'Environmental DNA for wildlife biology and biodiversity monitoring', *Trends in Ecology & Evolution,* 29(6), pp. 358-367.

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., . . . Gregory Caporaso, J. (2018) 'Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin', *Microbiome,* 6(1), pp. 90.

Bolyen, E. and Rideout, J. R. and Dillon, M. R. and Bokulich, N. A. and Abnet, C. C. and Al-Ghalith, G. A., . . . Caporaso, J. G. (2019) 'Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2', *Nature Biotechnology,* 37(8), pp. 852-857.

Bru, D., Ramette, A., Saby, N. P. A., Dequiedt, S., Ranjard, L., Jolivet, C., . . . Philippot, L. (2011) 'Determinants of the distribution of nitrogen-cycling microbial communities at the landscape scale', *The ISME Journal,* 5(3), pp. 532-542.

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

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., . . . Knight, R. (2010) 'QIIME allows analysis of high-throughput community sequencing data', *Nature Methods,* 7(5), pp. 335-336.

Crotty, F. V., Adl, S. M., Blackshaw, R. P. and Murray, P. J. (2012) 'Protozoan pulses unveil their pivotal position within the soil food web', *Microbial Ecology,* 63(4), pp. 905-918.

Darby, B.J., Todd, T.C., Herman, M.A., (2013) 'High-throughput amplicon sequencing of rRNA genes requires a copy number correction to accurately reflect the effects of management practices on soil nematode community structure'. Molecular Ecology 22, pp. 5456-5471.

De Cáceres, M., Sol, D., Lapiedra, O. and Legendre, P. (2011) 'A framework for estimating niche metrics using the resemblance between qualitative resources', *Oikos,* 120(9), pp. 1341-1350.

Decaëns, T., Porco, D., Rougerie, R., Brown, G. G. and James, S. W. (2013) 'Potential of DNA barcoding for earthworm research in taxonomy and ecology', *Applied Soil Ecology,* 65, pp. 35-42.

Djemiel C., Dequiedt, S., Karimi, B., Cottin, A., Horrigue, W., Bailly, A., Boutaleb, A., Sadet-Bourgeteau, S., Maron, P. A., Prévost-Bouré, N. C., Ranjard, L. and Terrat, S. (2022) 'Potential of Meta-Omics to Provide Modern Microbial Indicators for Monitoring Soil Quality and Securing Food Production'. *Frontiers Microbiology*, 13, 889788. doi: 10.3389/fmicb.2022.889788.

Donn, S., Neilson, R., Griffiths, B. S. and Daniell, T. J. (2012) 'A novel molecular approach for rapid assessment of soil nematode assemblages – variation, validation and potential applications', *Methods in Ecology and Evolution,* 3(1), pp. 12-23.

Dufrêne, M. and Legendre, P. (1997) 'Species assemblages and indicator species: The need for a flexible asymmetrical approach.' *Ecological Monographs* 67(3): 345-366.

Du Preez, G., Daneel, M., De Goede, R., Du Toit, M. J., Ferris, H., Fourie, H., Geisen, S., Kakouli-Duarte, T., Korthals, G., Sánchez-Moreno, S., and J Schmidt, J. H. (2022) 'Nematode-based indices in soil ecology: Application, utility, and future directions' *Soil Biology and Biochemistry, 169,* 108640. doi.org: 10.1016/j.soilbio.2022.108640.

Elphinstone J.G, Griffiths B, M, G. and E, S. (2017) *Soil Biology and Soil Health Partnership Project 3: Molecular approaches for routine soil-borne disease and soil health assessment –establishing the scope. Accessed at: https://projectblue.blob.core.windows.net/media/Default/Research%20Papers/Cereals%20and%20Oilseed/2019%20(and%20earlier)/91140002%20final%20report%2003.pdf* .

Faith, D. P. (1992) 'Conservation evaluation and phylogenetic diversity', *Biological Conservation,* 61(1), pp. 1-10.

Ferris, H., Bongers, T. and de Goede, R. G. M. (2001) 'A framework for soil food web diagnostics: extension of the nematode faunal analysis concept', *Applied Soil Ecology,* 18(1), pp. 13-29.

Fierer, N. and 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), pp. 626-631.

Flegg, J. J. M. (1967) 'Extraction of *Xiphinema* and *Longidorus* species from soil by a modification of Cobb's decanting and sieving technique*', *Annals of Applied Biology,* 60(3), pp. 429-437.

Gao, Z. and Guo, G. (2020) 'Command filtered path tracking control of saturated ASVs based on time-varying disturbance observer', *Asian Journal of Control,* 22(3), pp. 1197-1210.

George, P. B. L., Keith, A. M., Creer, S., Barrett, G. L., Lebron, I., Emmett, B. A., . . . Jones, D. L. (2017) 'Evaluation of mesofauna communities as soil quality indicators in a national-level monitoring programme', *Soil Biology and Biochemistry,* 115, pp. 537-546.

George, P. B. L., Lallias, D., Creer, S., Seaton, F. M., Kenny, J. G., Eccles, R. M., . . . Jones, D. L. (2019) 'Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems', *Nature Communications,* 10(1), pp. 1107.

Griffiths, B.S., de Groot, G.A., Laros, I., Stone, D.G., and Geisen, S. (2018). 'The need for standardisation: exemplified by a description of the diversity, community structure and ecological indices of soil nematodes'. *Ecological Indicators*, 87, pp. 43 - 46. https://doi.org/10.1016/j.ecolind.2017.12.002

Grime JP and Pierce S. 2012. The Evolutionary Strategies that Shape Ecosystems. Wiley-Blackwell, Chichester, UK

Giraldo-Perez, P., Raw, V., Greven, M. and Goddard, M. R. (2021a) 'A small effect of conservation agriculture on soil biodiversity that differs between biological kingdoms and geographic locations', *iScience,* 24(4), pp. 102280.

Giraldo-Perez, P., Raw, V., Greven, M. and Goddard, M. R. (2021b) 'A small effect of conservation agriculture on soil biodiversity that differs between biological kingdoms and geographic locations', *iScience,* 24(4).

Griffiths, B. S., Bengough, A. G., Neilson, R. and Trudgill, D. L. (2002) 'The extent to which nematode communities are affected by soil factors - a pot experiment', *Nematology,* 4(8), pp. 943-952.

Hannula, S. E., Di Lonardo, D. P., Christensen, B. T., Crotty, F. V., Elsen, A., van Erp, P. J., . . . Toth, Z. (2021) 'Inconsistent effects of agricultural practices on soil fungal communities across 12 European long-term experiments', *European Journal of Soil Science*. 12(1), pp. 5686

Hannula, S.E., Kielak, A. M., Steinauer, K., Huberty, M., Jongen, R., De Long, J.R., Heinen, R. and Martijn Bezemer, T. (2019) 'Time after time: Temporal Variation in the Effects of Grass and Forb Species on Soil Bacterial and Fungal Communities'. mBIO 10(6), e0235-19 https://doi.org/10.1128/mBio.02635-19

Harkes, P., Suleiman, A. K. A., van den Elsen, S. J. J., de Haan, J. J., Holterman, M., Kuramae, E. E. and Helder, J. (2019) 'Conventional and organic soil management as divergent drivers of resident and active fractions of major soil food web constituents', *Scientific Reports,* 9(1), pp. 13521.

Hartmann, M., Frey, B., Mayer, J., Mäder, P. and Widmer, F. (2015) 'Distinct soil microbial diversity under long-term organic and conventional farming', *The ISME Journal,* 9(5), pp. 1177-1194.

Hendgen, M., Hoppe, B., Döring, J., Friedel, M., Kauer, R., Frisch, M., . . . Kellner, H. (2018) 'Effects of different management regimes on microbial biodiversity in vineyard soils', *Scientific Reports,* 8(1), pp. 9393.

Högfors-Rönnholm, E., Christel, S., Engblom, S. and Dopson, M. (2018) 'Indirect DNA extraction method suitable for acidic soil with high clay content', *MethodsX,* 5, pp. 136-140.

Kidd, J., Manning, P., Simkin, J., Peacock, S. and Stockdale, E (2017) 'Impacts of 120 years of fertilizer addition on a temperate grassland ecosystem'. PLoS ONE 12(3) e0174632. doi.org: 10.1371/journal.pone.0174632

Kirse, A., Bourlat, S.J.,Langen, Kand Fonseca, V. G. (2021) 'Unearthing the potential of soil eDNA metabarcoding - towards best practice advice for invertebrate biodiversity assessment. Frontiers in Ecology and Evolution, Article 630560 doi: 10.3389/fevo.2021.630560

Lauber, C. L., Hamady, M., Knight, R. and Fierer, N. (2009) 'Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale', *Applied and Environmental Microbiology,* 75(15), pp. 5111-5120.

Lawton, J. H., Bignell, D. E., Bolton, B., Bloemers, G. F., Eggleton, P., Hammond, P. M., . . . Watt, A. D. (1998) 'Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest', *Nature,* 391(6662), pp. 72-76.

Lecroq, B., Lejzerowicz, F., Bachar, D., Christen, R., Esling, P., Baerlocher, L., Østerås, M., Farinelli, L. and Pawlowski, J. (2011) Ultra-deep sequencing of foraminiferal microbarcodes unveils hidden richness of early monothalamous lineages in deep-sea sediments. *Proceedings of the National Academy of Sciences U S A.* 108(32):13177-82. doi: 10.1073/pnas.1018426108.

Lehman, R. M., Cambardella, C. A., Stott, D. E., Acosta-Martinez, V., Manter, D. K., Buyer, J. S., . . . Karlen, D. L. (2015) 'Understanding and enhancing soil biological health: the solution for reversing soil degradation', *Sustainability,* 7(1), pp. 988-1027.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., . . . Machida, R. J. (2013) 'A new versatile primer set targeting a short fragment of the mitochondrial COI

region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents', *Frontiers in zoology,* 10(1), pp. 1-14.

Li, Z., Qiu, K., Schneider, R. L., Morreale, S. J. and Xie, Y. (2019) 'Comparison of microbial community structures in soils with woody organic amendments and soils with traditional local organic amendments in Ningxia of Northern China', *PeerJ,* 7, pp. e6854.

Liu, T.-H., Yaghmour, M. A., Lee, M.-H., Gradziel, T. M., Leveau, J. H. J. and Bostock, R. M. (2020) 'An roGFP2-Based Bacterial Bioreporter for Redox Sensing of Plant Surfaces', *Phytopathology®,* 110(2), pp. 297-308.

Louca, S., Parfrey, L. W. and Doebeli, M. (2016) 'Decoupling function and taxonomy in the global ocean microbiome', *Science,* 353(6305), pp. 1272-1277.

Lozupone, C. and Knight, R. (2005) 'UniFrac: a New Phylogenetic Method for Comparing Microbial Communities', *Applied and Environmental Microbiology,* 71(12), pp. 8228-8235.

Mantel, N. (1967) 'The detection of disease clustering and a generalized regression approach', *Cancer research,* 27(2 Part 1), pp. 209-220.

McMurdie, P. J. and Holmes, S. (2013) 'phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data', *PLOS ONE,* 8(4), pp. e61217.

Morrison-Whittle, P. and Goddard, M. R. (2015) 'Quantifying the relative roles of selective and neutral processes in defining eukaryotic microbial communities', *The ISME Journal,* 9(9), pp. 2003-2011; https://doi.org/10.1016/j.isci.2021.102280.

Morrison-Whittle, P. and Goddard, M. R. (2018) 'From vineyard to winery: a source map of microbial diversity driving wine fermentation', *Environmental Microbiology,* 20(1), pp. 75-84.

Morrison-Whittle, P., Lee, S. A. and Goddard, M. R. (2017) 'Fungal communities are differentially affected by conventional and biodynamic agricultural management approaches in vineyard ecosystems', *Agriculture, Ecosystems & Environment,* 246, pp. 306-313.

Msimbira, L. A. and Smith, D. L. (2020) 'The roles of plant growth promoting microbes in enhancing plant tolerance to acidity and alkalinity stresses', *Frontiers in Sustainable Food Systems,* 4. https://doi.org/10.3389/fsufs.2020.00106

Muletz Wolz, C. R., Yarwood, S. A., Campbell Grant, E. H., Fleischer, R. C. and Lips, K. R. (2018) 'Effects of host species and environment on the skin microbiome of Plethodontid salamanders'. *Journal of Animal Ecology* 87(2) pp.341-353.

Murphy, D. V., Stockdale, E. A., Brookes, P.C. and Goulding, K. W. T. (2003). Impact of micro-organisms on chemical transformations in soil. In Soil Biological Fertility – A Key to Sustainable Land Use in Agriculture. Eds. L. K. Abbott & D. V. Murphy. pp. 225-240. Dordrecht, Kluwer.

Nguyen, N. H., Song, Z., Bates, S. T., Branco, S., Tedersoo, L., Menke, J., . . . Kennedy, P. G. (2016) 'FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild', *Fungal Ecology,* 20, pp. 241-248.

Nilsson, R. H., Taylor, A. F. S., Adams, R. I., Baschien, C., Bengtsson-Palme, J., Cangren, P., . . . Abarenkov, K. (2018) 'Taxonomic annotation of public fungal ITS sequences from the built environment – a report from an April 10–11, 2017 workshop (Aberdeen, UK)', *MycoKeys,* 28, pp. 65-82.

Nilsson, R.H., Anslan, S., Bahram, M., Wurzbacher, C., Baldrian, P and Tedersoo, L. (2019) ' Mycobiome diversity: high-throughput sequencing and identification of fungi' *Nature Reviews Microbiology* 17, 95-109.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P. and McGlinn, D. et al. (2018). vegan: Community Ecology. Package. R package version 2.5-2. https://CRAN.R-project.org/package=vegan

Pan, H., Chen, M., Feng, H., Wei, M., Song, F., Lou, Y., . . . Zhuge, Y. (2020) 'Organic and inorganic fertilisers respectively drive bacterial and fungal community compositions in a fluvo-aquic soil in northern China', *Soil and Tillage Research,* 198, pp. 104540.

Paulson, J. N., Stine, O. C., Bravo, H. C. and Pop, M. (2013) 'Differential abundance analysis for microbial marker-gene surveys', *Nature Methods,* 10(12), pp. 1200-1202.

Peine, M., Vitow, N., Grafe, M., Baum, C., Zicker, T., Eichler-Löbermann, B., . . . Leinweber, P. (2019) 'Effect of triple superphosphate and biowaste compost on mycorrhizal colonization and enzymatic P mobilization under maize in a long-term field experiment', *Journal of Plant Nutrition and Soil Science,* 182(2), pp. 167-174.

Plassart, P., Terrat, S., Thomson, B., Griffiths, R., Dequiedt, S., Lelievre, M., . . . Ranjard, L. (2012) 'Evaluation of the ISO Standard 11063 DNA Extraction Procedure for Assessing Soil Microbial Abundance and Community Structure', *PLOS ONE,* 7(9), pp. e44279.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., . . . Glöckner, F. O. (2012) 'The SILVA ribosomal RNA gene database project: improved data processing and web-based tools', *Nucleic Acids Research,* 41(D1), pp. D590-D596.

Quist, C.W., Schrama, M., de Haan, J.J., Smant, G., Bakker, J., van der Putten, W.H. and Helder, J. (2016). Organic farming practices result in compositional shifts in nematode communities that exceed crop-related changes. *Applied Soil Ecology* 98, pp. 254-260.

Randazzo, W., Truchado, P., Cuevas-Ferrando, E., Simón, P., Allende, A., and Sánchez, G. (2020). 'SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area' *Water Research*, 181: article115942. https://doi.org/10.1016/j.watres.2020.115942

Ritz, K. and Trudgill, D. L. (1999) 'Utility of nematode community analysis as an integrated measure of the functional state of soils: perspectives and challenges', *Plant and Soil,* 212(1), pp. 1-11.

Rognes, T., Flouri, T., Nichols, B., Quince, C. and Mahé, F. (2016) 'VSEARCH: a versatile open source tool for metagenomics', *PeerJ,* 4, pp. e2584.

Rousk, J., Baath, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., . . . Fierer, N. (2010) 'Soil bacterial and fungal communities across a pH gradient in an arable soil', *ISME Journal,* 4(10), pp. 1340-1351.

Ruppert, K. M., Kline, R. J. and 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*. **17**: e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S. and Huttenhower, C. (2011) 'Metagenomic biomarker discovery and explanation', *Genome Biology,* 12(6), pp. R60.

Seinhorst, J. W. (1955) 'Een eenvoudige methode voor het afscheiden van aaltjes uit grond', *Tijdschrift Over Plantenziekten,* 61(1), pp. 188-190.

Sieriebriennikov, B., Ferris, H. and de Goede, R. G. M. (2014) NINJA: An automated calculation system for nematode-based biological monitoring. *European Journal of Soil Biology* 61, 90-93.

Silva, U. C., Medeiros, J. D., Leite, L. R., Morais, D. K., Cuadros-Orellana, S., Oliveira, C. A., . . . Dos Santos, V. L. (2017) 'Long-term rock phosphate fertilisation impacts the microbial communities of maize rhizosphere', *Frontiers in Microbiology,* 8. https://doi.org/10.3389/fmicb.2017.01266

Stork, N. E. and Eggleton, P. (1992) 'Invertebrates as determinants and indicators of soil quality', *American Journal of Alternative Agriculture,* 7(1-2), pp. 38-47.

Toju, H., Tanabe, A. S., Yamamoto, S. and Sato, H. (2012) 'High-coverage ITS primers for the DNA-based identification of Ascomycetes and Basidiomycetes in environmental samples', *PLOS ONE,* 7(7), pp. e40863.

Tomczak, M. and Tomczak, E. (2014) 'The need to report effect size estimates revisited. An overview of some recommended measures of effect size', *Trends in sport sciences,* 1(21), pp. 19-25.

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G.H., Geniez, P., Pont, D., Argillier, C., Baudoin, J.M., Peroux, T., Crivelli, A. J., Olivier, A., Acqueberge, M., Le Brun, M., Møller, P. R., Willerslev, E., and Dejean, T. (2016) 'Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding'. *Molecular Ecology* 25(4) pp. 929-42. doi: 10.1111/mec.13428.

van Straalen, N. M. (1985). 'Comparative demosgraphy of forest floor Collembola populations'. *Oikos,* 45:2, pp. 253-265

Vervoort, M. T. W., Vonk, J. A., Mooijman, P. J. W., Van den Elsen, S. J. J., Van Megen, H. H. B., Veenhuizen, P., . . . Helder, J. (2012) 'SSU ribosomal DNA-based monitoring of nematode

assemblages reveals distinct seasonal fluctuations within evolutionary heterogeneous feeding guilds', *PLOS ONE,* 7(10), pp. e47555.

Vestergaard, G., Schulz, S., Schöler, A. and Schloter, M. (2017). Making big data smart—how to use metagenomics to understand soil quality. *Biology and Fertility of Soils* 53, 479–484. doi: 10.1007/s00374-017-1191-3

Vrijenhoek, R. (1994) 'DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates', *Mol Mar Biol Biotechnol,* 3(5), pp. 294-9.

Wallwork J.A. (1976). The distribution and diversity of soil fauna. London: Academic Press 355 pages.

Walters, W., Hyde, E. R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., . . . Bik, H. (2016) 'Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal Internal Transcribed Spacer marker gene primers for microbial community surveys', *mSystems,* 1(1), pp. e00009-15.

Wang, J., Song, Y., Ma, T., Raza, W., Jing, L., Howland, J. G., . . . Shen, Q. (2017) 'Impacts of inorganic and organic fertilisation treatments on bacterial and fungal communities in a paddy soil', *Applied Soil Ecology,* 112, pp. 42-50.

Wardle, D.A., Yeates, G. W., Watson, R.N. and Nicholson, K. S. (1995) The detritus food-web and the diversity of soil fauna as indicators of disturbance regimes in agro-ecosystems. Plant and Soil 170, 35-43.

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A. et al. (2017) Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5(1) pp.27

Wiesel, L., Daniell, T. J., King, D. and Neilson, R. (2015) 'Determination of the optimal soil sample size to accurately characterise nematode communities in soil', *Soil Biology and Biochemistry,* 80, pp. 89-91.

Yao, L., Wang, D., Kang, L., Wang, D., Zhang, Y., Hou, X. and Guo, Y. (2018) 'Effects of fertilisations on soil bacteria and fungi communities in a degraded arid steppe revealed by high through-put sequencing', *PeerJ,* 6, pp. e4623.

Yeates, G. W. and Bongers, T. (1999) 'Nematode diversity in agroecosystems', in Paoletti, M.G. (ed.) *Invertebrate Biodiversity as Bioindicators of Sustainable Landscapes*. Amsterdam: Elsevier, pp. 113-135.

Zhong, W., Gu, T., Wang, W., Zhang, B., Lin, X., Huang, Q. and Shen, W. (2010) 'The effects of mineral fertiliser and organic manure on soil microbial community and diversity', *Plant and Soil,* 326(1), pp. 511-522.