Molecular Analysis of Soil Mesofaunal Diversity and the Effects of Natural Capital in Agroecosystems

Angeli Sahdra

Thesis submitted to the University of Ottawa

in partial Fulfillment of the requirements for the

Master of Science Biology

Department of Biology

Faculty of Science

University of Ottawa

Abstract

Soil biodiversity is key to maintaining ecosystem functions, but is threatened by anthropogenic changes. To help mitigate these risks, retaining small habitat patches within agroecosystems is a promising method, but it can be challenging to study due to the vast diversity of soil organisms. One important, but understudied, group that contributes to ecosystem functions, such as contributing to soil structure and cycling nutrients in the soil, is the soil mesofauna (~0.2 – 2 mm in size), mainly consisting of mites and springtails. DNA metabarcoding is a tool that is increasingly used to expand our knowledge of biodiversity, but it is still a work in progress with respect to investigating soil mesofaunal communities. A comparison of this technique performed on bulk mesofauna specimens versus bulk soil samples, collected in 2021 from an intensively farmed area of Eastern Ontario, found differences in species richness, but nonetheless, similarities were detected in community composition between both methods (Chapter 2). Identifying specimens morphologically gave different results compared to both molecular methods. During the summer of 2022, a survey of mesofauna in four types of non-field habitats (ditch margins with and without trees, small forest patches, and larger forest patches) was conducted in the same region (Chapter 3). Although the number of mesofaunal species was highly variable across all the sites surveyed, differences were detected in community composition between forested and unforested sites. The communities were highly variable within each habitat type. The lack of overlap in species presence between the different sites could be due to a variety of factors, such as the small-scale habitat features of the sites or differences in microhabitats between samples. As well, there was no clear separation between small and large forest patches with respect to mesofaunal community composition, which showcases the value that even small patches of natural habitat can have within agroecosystems.

Soil physicochemical properties, particularly organic matter, moisture, and cation exchange capacity (CEC), also influenced these communities, while there was little evidence that the mesofaunal community varied over the growing season. This thesis improves understanding of soil mesofaunal diversity and community composition within agroecosystems in Canada, which remain largely uncharacterized. DNA metabarcoding yields insights into species composition of soil mesofaunal communities that are elusive, at best, using traditional morphological techniques.

Acknowledgements

I would like to thank both of my thesis supervisors, Marla Schwarzfeld and Jeremy Kerr, for their support and guidance throughout my degree. I am very grateful to Marla for pushing me outside of my comfort zone and encouraging me to attend and present at the Canadian Society for Ecology and Evolution and Canadian Botanical Association conference, as I learned so much and was able to connect with my peers through this experience and opportunity. I am also very thankful for Monica Young, Tori Miller, and Jennifer Lafave for their help both within the field and the lab. I am also grateful towards David Lapen, Emilia Craiovan, Niloofar Alavi-Shoushtari, and Amanda Halstead for help with site selection and field logistics. I would also like to thank Lisa Koziol for helping me to troubleshoot my sequencing data, along with Lori Phillips for being so kind and welcoming and for providing me with soil physicochemical data for my project. I am also thankful for my committee members, Joe Bennett and Heather Kharouba, for their guidance and insight for the completion of this thesis. Along with this, I am also thankful to the University of Ottawa, the Research Affiliate Program, Environmental Change One Health Observatory (ECO2), and Agriculture and Agri-Food Canada (AAFC) for funding my research. Finally, I would like to thank my friends and family for being my support system throughout my degree.

Table of Contents

List of Tables

List of Figures

Chapter 1. General Introduction

Soils contain many species that influence key ecosystem functions, including nutrient cycling, soil formation, and carbon transformation (Barrios, 2007; Lavelle et al., 2006). These abundant soil organisms are incredibly diverse, with thousands of species per gram of soil (Roesch et al., 2007). Soil organisms can range from microorganisms (e.g. bacteria, fungi, and protists) and microfauna (e.g. nematodes and rotifers, < 200 μm in size) to mesofauna (e.g. mites and springtails, ~0.2 – 2 mm in size) and macrofauna (earthworms, snails, and arthropods, > 2 mm in size) (FAO, 2020; Lavelle et al., 2006). The soil biota has a large influence on soil structure and function. Microbiota within the soil, such as protists and nematodes, can release nutrients within the soil from microbial biomass (Potapov et al., 2022). These organisms feed on soil bacteria and fungi and aid in litter decomposition, influencing the amount of organic matter contained within the soil (Heděnec et al., 2022; Joly et al., 2018). This influences the physical and chemical properties of the soil, such as moisture and pH (Heděnec et al., 2022). Herbivorous nematodes can also influence aboveground vegetation and plant communities (Heděnec et al., 2022). Macrofauna, such as earthworms and ants, are responsible for changing the structure of soil through their movement and ingestion of soil (Potapov et al., 2022). They also aid in nutrient cycling, plant growth, and interact with the soil microbial community to help decompose organic matter through predation of soil microorganisms and increase nitrogen mineralization within the soil (Akhila et al., 2022). One of the most diverse and abundant group of soil organisms are the soil mesofauna (FAO, 2020). They play a large role in maintaining essential ecosystem functions, such as contributing to soil structure and providing nutrients to the soil through the breakdown of organic matter (Coulibaly et al., 2019; FAO, 2020; Potapov et al., 2022). However, while much research has focused on the soil microbiome (i.e. bacteria and fungi) (Barrios, 2007;

Girvan et al., 2005; van Bruggen & Semonov, 2000), less attention has been given to soil fauna, and especially to soil mesofauna.

The soil mesofauna consists of a highly diverse complement of small animals (~0.2 – 2 mm in size); however, the dominant groups are mites (Acari) and springtails (Collembola) (Behan-Pelletier, 2003). Springtails are small hexapod arthropods that graze on microorganisms, with some predatory taxa feeding on nematodes and rotifers, thus contributing to the breakdown of organic matter (Coulibaly et al., 2019; Ferguson & Joly, 2002; Hopkin, 1997). Such species are an important prey source for small arthropods, including spiders, beetles, and predatory mites (Bilde et al., 2000). Soil mites (Acari) are the most diverse and abundant taxon in many soil ecosystems, and greatly influence many ecosystem processes (Beaulieu et al., 2019). These arachnids include detritivores, such as oribatid mites, which feed on dead organic matter and microbial residues and contribute to soil structure in their production of faecal pellets rich in nutrients that can act as a natural fertilizer for the soil (FAO, 2020; Potapov et al., 2022). Other taxa, such as Mesostigmata, are predators that can affect prey populations, which in turn can have top-down effects on soil community structure (Koehler, 1999; Petrova et al., 2004). They have also been used as a biological control for insect pests (McMurtry et al., 2015; Schneider & Maraun, 2009). Along with various feeding strategies, mites have a wide range of life-histories. For example, some groups, such as Astigmata, can reproduce quickly and track resource pulses, whereas others, such as oribatid mites, have low reproductive rates and stable populations (Iatrou et al., 2010; Kheradmand et al., 2007; Lebrun & van Straalen, 1995). Finally, mites have sometimes been used as bioindicators of soil and freshwater health (van Straalen, 1998).

Although this complex network of soil organisms contribute to ecosystem functions, species-level surveys of mesofauna are rare due to the vast amount of diversity within the group.

In turn, this makes it difficult to understand how environmental factors, such as habitat type and soil properties, influence their diversity or community structure (Gardi et al., 2013; Lee, 1994; Tibbett et al., 2020). While there are several studies that focus on surveying the mesofauna (e.g. Arroyo et al., 2013; George et al., 2017; Kamczyc et al., 2022; Meehan et al., 2019; Porco et al., 2013; Rueda-Ramírez et al., 2022), these studies are still relatively scarce. Collecting baseline information on mesofauna can help to understand the effects of anthropogenic changes on the soil mesofauna and can inform sustainable practices for the conservation of soil biodiversity (George et al., 2017).

Anthropogenic changes, such as habitat disruption and fragmentation, agricultural intensification, and soil erosion, have been found to negatively impact diversity of the soil community and the essential ecosystem services that it provides (Gardi et al., 2013; Tibbett et al., 2020). Agricultural intensification, for example, can result in a large decline of the soil fauna community due to tillage and pesticide use, which causes the destruction of habitats and high rates of mortality for soil fauna (Decaëns et al., 2006). The soil fauna is also sensitive to changes within the soil, such as organic matter, moisture, and available nutrients that have been found to impact the soil communities of agricultural systems (de Graaff et al., 2019). A useful method that has been shown to benefit above-ground biodiversity includes retaining small habitat remnants, termed “natural capital” , especially within highly fragmented and disturbed agricultural matrices (Montgomery et al., 2020; Morris et al., 2010). For example, agricultural ditches and the adjacent vegetation, ranging from small, untreed ditch margins to larger forested areas, have been used on farms as refugia to preserve biodiversity (Needelman et al., 2007; Shaw et al., 2015). These environments have been shown to promote the biodiversity of butterfly, bird, earthworm, and insect species richness (Montgomery et al., 2020; Udawatta et al., 2019); however there is

limited research addressing the impact that retaining natural capital has on soil mesofaunal diversity. Within these small habitat patches, many environmental factors can influence the soil community across a growing season, such as the physical and chemical properties of the soil (i.e. moisture, organic matter, etc.) (Bardgett & Cook, 1998; Decaëns et al., 2006; de Graaff et al., 2019; Wu & Wang, 2019). For example, Wu & Wang (2019) found that seasonal changes in precipitation can shape the soil mesofaunal community due to changes in soil moisture. Crotty et al. (2023) found that forested habitats containing higher amounts of organic matter had greater soil arthropod abundance and richer soil communities. While seasonality and the physical properties of soils can affect forest soil communities (Crotty et al., 2023; Wu & Wang, 2019), there is little research that targets how environmental factors and seasonal changes impact soil mesofaunal diversity and communities within agricultural ecosystems.

Identifying specimens to the species level using morphological methods can limit our understanding of how these communities change through time and across ecological gradients (Lee, 1994; Semenov, 2021). Even if appropriate taxonomic expertise exists, morphological species identification can be time consuming even for a small number of samples, and constraints are likely to be prohibitive for larger projects. Moreover, many mesofaunal species are not yet formally described (Lumley et al., 2013; Turnbull & Stebaeva, 2019), so taxonomic keys for them do not exist (Beaulieu et al., 2019). Species also show marked morphological variation at different life cycle stages, increasing the challenge of their identification (Druciarek et al., 2019). For these reasons, most studies have only focused on a subset of taxa present (Baumann, 2021; Postma-Blaauw et al., 2012), or identify taxa to higher taxonomic levels (Meehan et al., 2019; Ruf & Beck, 2005). On the other hand, morphological identifications yield abundance measurements for species and can be relatively affordable.

DNA metabarcoding offers an alternative approach to assess species richness (Taberlet et al., 2012; Young & Hebert, 2022), based on DNA taxonomy and high-throughput sequencing (Boggs et al., 2019; Yang et al., 2014). It has been used on both bulk specimen samples and bulk soil samples (Boggs et al., 2019; Oliverio et al., 2018). However, there is little research on the comparison of these two methods in assessing the soil mesofaunal community (but see Arribas et al., 2016; Young et al., 2022). This tool has been used to assess species richness in a variety of habitats, such as in the Arctic, in temperate and tropical forest habitats, and in freshwater and grassland habitats (Arribas et al., 2016; Basset et al., 2022; Oliverio et al., 2018; Schenk et al., 2020; Young & Hebert, 2022). Agricultural mesofaunal communities remain relatively unexplored. DNA metabarcoding can be a costly tool, especially for a large number of samples (Bohmann et al., 2021), and it does not provide abundance data (Kennedy et al., 2020). Metabarcoding results depend on the sampling process, requiring both DNA extraction and PCR amplification steps, so techniques require caution to eliminate possible cross contamination and biases between steps (Liu et al., 2020; Zinger et al., 2019).

In this thesis, I explore using molecular methods to assess the species richness and community composition of the soil mesofaunal community. In Chapter 2, I do this by performing DNA metabarcoding on paired soil samples, with one of each pair being a bulk soil sample and the other being a bulk specimen sample. This chapter compares the two methods in their efficacy in detecting the mesofaunal community within the paired samples. This chapter also compares the morphological identifications in each sample with the metabarcoding data from the bulk specimen samples to determine if each method captures similar or different levels of soil mesofaunal diversity and community composition. In Chapter 3, I assess the role that natural capital plays on the soil mesofaunal community by performing specimen-based metabarcoding

on soil samples from different levels of natural capital within an agricultural matrix across the summer season. I also examine the soil physicochemical properties of each habitat type to see if these factors play a role in shaping the soil mesofaunal community. The metabarcoding data collected from all sites from Chapter 2 and Chapter 3 are combined to provide a comprehensive species list of the soil mesofauna present within these agricultural environments (Appendix 1). This thesis demonstrates the potential that DNA metabarcoding has for assessing species richness and community composition of the soil mesofaunal community, and it also addresses the importance of retaining natural capital within a highly disturbed and fragmented agricultural landscape in order to preserve soil biodiversity.

Chapter 2. A Comparison of Specimen- vs Soil-Based Metabarcoding for Mesofaunal Communities

Abstract

Soil mesofauna play critical roles in maintaining soil health and resilience, however are difficult to study due to their diversity and abundance. DNA metabarcoding is a powerful technique that has been used to broaden our understanding of biodiversity, but its potential for exploring soil mesofaunal communities is still under development. This study aims to compare species composition and diversity found from metabarcoding bulk soil versus mixed mesofaunal specimens extracted from the soil. Twelve paired samples were collected from agricultural sites in southeastern Ontario; one of each pair was used for soil DNA extraction, while the other was Berlese-extracted to obtain the mesofaunal specimens. These specimens were non-destructively DNA-extracted, and were then slide-mounted for morphological identification. Illumina sequencing of COI was conducted on both soil and bulk specimen samples. In total, over 2,000 operational taxonomic units (OTUs) were recovered, with the majority found in the bulk soil samples. However, many of these OTUs are unidentifiable with existing databases and many may be unreliably identified. Although species diversity differed between the molecular methods tested, similar ecological trends were detected when addressing mesofaunal community composition. Molecular and morphological identifications differed significantly, perhaps because a family level comparison was used or because of methodological issues. DNA metabarcoding of bulk soil samples detects far higher diversity than Berlese specimen samples, but methodological challenges of DNA metabarcoding require greater refinement and validation.

Introduction

Soil biodiversity is critically important for maintaining essential ecosystem functions, such as soil formation, nutrient cycling, and carbon transformation (Barrios, 2007; Lavelle et al., 2006). However, due to the vast diversity within the soil, it is often considered a "black box" with respect to understanding what organisms are present and in what ways they are interacting (Geisen et al., 2019; Lee, 1994). The advent of molecular methods has massively expanded knowledge of soil biodiversity; however, most work has focused on the soil microbiome (bacteria and fungi) (Barrios, 2007; Girvan et al., 2005; van Bruggen & Semonov, 2000), while significantly less attention has been given to the soil fauna, and particularly to the soil mesofauna.

Soil mesofauna are animals between approximately 0.2 to 2 mm in size (Lakshmi et al., 2020; Lavelle et al., 2006). The category includes a wide variety of taxa, but is dominated by small arthropods, in particular, mites (Acari) and springtails (Collembola) (Behan-Pelletier, 2003). Springtails are non-insect hexapods that are primarily fungivorous and contribute to the breakdown of organic material, though they also include predatory springtail species that typically feed on nematodes and rotifers (Ferguson & Joly, 2002; Hopkin, 1997). Mites are arachnids that are extremely abundant and diverse and fill a wide range of ecological roles (Beaulieu et al., 2019). For example, oribatid mites are particle-feeding detritivores, feeding on fungi, bacterial films, and sometimes scavenging dead or slow-moving animals (Behan-Pelletier, 1999; Potapov et al., 2022). In contrast, many mites, such as Mesostigmata, are predators that can exert top-down effects on soil community structure (Koehler, 1999; Petrova et al., 2004). Mites also have a wide range of reproductive strategies; some can reproduce rapidly and track resource pulses (e.g. Astigmata), whereas others have low reproductive rates and stable

populations (e.g. Oribatida) (Iatrou et al., 2010; Kheradmand et al., 2007; Lebrun & van Straalen, 1995). Mites have also been used as bioindicators of soil and freshwater condition (van Straalen, 1998).

Identification of mesofaunal species has traditionally required morphological approaches. However, even specialists often encounter species that cannot be distinguished on the basis of their morphology, as well as species that are not yet described. It can be difficult to identify some species at all stages of their life cycles (Druciarek et al., 2019). Finally, due to the small size and high abundance of soil mesofauna, morphological identifications can be extremely time-consuming and impractical for a large number of samples (Geisen et al., 2019). For these reasons, most studies have focused on only a subset of the taxa present (Baumann, 2021; Postma-Blaauw et al., 2012), or identify taxa to higher taxonomic levels (Meehan et al., 2019; Ruf & Beck, 2005). The advantages of morphological analyses is that they can be less costly than molecular analysis, and they provide abundance data that may not be measurable with some molecular methods.

DNA metabarcoding has been used for identification of species within communities in a sample using high throughput sequencing and DNA taxonomy (Taberlet et al., 2012; Young & Hebert, 2022). Although DNA metabarcoding is costly, it is useful for capturing the total biodiversity contained within bulk soil samples and has been shown to identify species and ecological structure of invertebrate communities (Arribas et al., 2016; Dopheide et al., 2019; Watts et al., 2019). However, unlike morphological identification, it is not possible to obtain abundance data using metabarcoding since specimens are lost due to destructive DNA extraction steps. Although read counts can sometimes be used as a proxy for abundance, this is still an active area of research (Kennedy et al., 2020). Results depend on the sampling process, DNA

extraction, PCR amplification, and the reliability of mechanisms used to avoid cross contamination and biases among these steps and among samples (Liu et al., 2020; Zinger et al., 2019).

DNA metabarcoding can be used to sequence environmental DNA, such as from bulk soil samples, without first extracting individual specimens (Oliverio et al., 2018; Watts et al., 2019). An advantage of metabarcoding bulk soil samples is that a larger diversity of taxa contained within the sample can be identified, than would be obtained by first separating the specimens from the soil (Watts et al., 2019). However, a disadvantage of this method is that samples can include dead specimens that fell to the soil from above. Since species' DNA degrades at different rates, abundance data derived from the sample may not be consistent (Deiner et al., 2017).

DNA metabarcoding has also been used to identify bulk samples of mesofaunal specimens extracted from the soil with Berlese funnels (Arribas et al., 2016; Oliverio et al., 2018). However, a potential disadvantage to using Berlese-extracted specimens is that this method has been shown to produce biases, such as specimens drying out before they can escape into collection jars (Sabu et al., 2011), or the inability to recover certain life stages (e.g. resting stages; Dritsoulas & Duncan, 2020), resulting in undetection of some of the species contained within the sample. While a few studies have compared these two methods (Arribas et al., 2016; Oliverio et al., 2018; Young & Hebert, 2022), most previous work has been done in natural ecosystems (e.g. Boggs et al., 2019; Yang et al., 2014). Agroecosystems remain relatively unexplored in this respect.

The objective of this study was to compare the diversity and community composition detected by metabarcoding bulk Berlese-extracted specimens versus bulk soil samples from agricultural sites in southeastern Ontario. I also assessed whether morphologically identified

voucher specimens were successfully detected through metabarcoding samples of bulk mesofaunal specimens. This will provide a better understanding of the efficacy and biases of using molecular techniques to identify the soil mesofaunal community.

Methods

Sample collection

Soil from 12 sites was sampled in agricultural fields and ditch margins located in southeastern Ontario, September 202l (45.265, -75.172). From each site, paired samples were collected using a tulip bulb planter, resulting in cores approximately 6 cm in diameter and 11 cm in depth. One of each pair was used for bulk specimen metabarcoding and morphological identification, using Berlese funnels and non-destructive DNA extraction (=BERLESE). The other samples were used for bulk soil eDNA extraction and metabarcoding (=SOIL).

Molecular methods

For the BERLESE samples, the collected soil samples were stored in the fridge overnight and then placed in Berlese funnels lined with cheesecloth to extract specimens (Figure 2.1). The Berlese funnel extraction was performed over a period of four days, with heat provided by a 60W bulb. Specimens were collected into 95% ethanol. The specimens were then filtered through a 45-µm sieve, transferred to 20 ml scintillation vials with fresh 95% ethanol, and stored at -20ºC. These bulk specimens were then DNA-extracted non-destructively using the Qiagen DNeasy Blood and Tissue kit (Qiagen), with slight modifications to allow for voucher specimen recovery. Prior to the lysis step, the specimens were transferred to 2 ml centrifuge tubes. The majority of ethanol was carefully removed, and the remaining ethanol was evaporated using a vacuum centrifuge (Vacufuge, AL setting) for approximately 10 minutes. The amount of lysis buffer was increased over the protocol recipe to ensure enough would be obtained while

minimizing the loss of specimens; thus 270 µl of buffer ATL and 30 µl of proteinase K were added to the specimen tube, which was then incubated overnight at 56ºC. After incubation, 200 µl of the remaining solution was siphoned out of the tube using a pipette and put into a separate 1.5 mL tube for extraction, with caution taken to avoid pipetting any specimens. The tube with the remaining voucher specimens was refilled with 95% ethanol for storage. The remaining steps followed the kit protocol, with the elution step repeated to maximize DNA yield. After this step, the spin column filters used for extraction were rinsed with 95% ethanol and examined under a microscope to ensure no specimens were lost during DNA extraction. If specimens were found, they were added to the appropriate voucher specimen tube.

For the SOIL samples, DNA extraction followed the protocol of Taberlet et al. (2012). Soil samples were soaked in a phosphate buffer solution of equal volume to the weight of the soil samples for 20 minutes to allow for the extracellular DNA to be extracted from the soil (Taberlet et al., 2012). Three replicate aliquots of 50 mL of the buffer were then centrifuged at 10,000 x g for 10 minutes; following centrifugation, 500 µl of the supernatant was collected and DNA extraction was performed using Qiagen PowerSoil kits (Qiagen), modified to skip the cell lysis step, thus beginning at step 6 of the protocol.

For both DNA types (SOIL and BERLESE), a 313 bp region of COI was amplified following the recommendation of Krehenwinkel et al. (2016), using the following primers: mICOIintF (Brandon-Mong et al., 2015) and FoldR (Arribas et al., 2016) with modified adapters containing 0-6 “N” bases between the adapter sequence and the target region of COI to increase base diversity. The PCR was replicated three times for each DNA extract, for a total of 36 BERLESE samples (12 sites x 3 PCR replicates) and 108 SOIL samples (12 sites x 3 extraction replicates x 3 PCR replicates). In addition, I conducted three SOIL negative extractions (using

the phosphate buffer with no soil substrate) x 3 PCR replicates and one BERLESE extraction negative x 6 PCR replicates. Finally, 3 PCR negatives were included on each sequencing plate, for a total of 6 PCR negatives. Each PCR reaction contained 12.5 µl Qiagen Multiplex PCR Mix (Qiagen), 0.25 µl each of the forward and reverse primers, 10 µl nuclease-free water, and 2 µl of template DNA. The plates were then amplified in an Eppendorf thermocycler using the following program: 95ºC for 15 minutes, 40 cycles of 94ºC for 30 seconds, 45ºC for 90 seconds, 72ºC for 90 seconds, and a final extension of 72ºC for 10 minutes before being held at 4ºC. Subsequent steps followed standard protocols at the Molecular Technologies Laboratory of the Ottawa Research and Development Centre. In brief, samples were purified and normalized using the NGS Normalization 96-Well Kit (Norgen Biotek Corporation, Canada). Samples were then tagged using i3 and i5 indices from the Nextera Index Kit (Illumina, San Diego, CA, USA). A second round PCR clean-up and normalization was done using Norgen kits as before. Finally, libraries were quantified using the KAPA Library Quantification Kit Kit (KAPA Biosystems, Wilmington, MA, USA), pooled and sequenced on a MiSeq next-generation sequencer, using a single MiSeq V2 Nano Kit (Illumina, San Diego, CA, USA).

Specimen identification

Voucher specimens recovered from the BERLESE samples were either slide mounted in a polyvinyl alcohol (PVA) medium or stored in alcohol. Specimens stored in alcohol were identified using a cavity slide and lactic acid in order to be able to adjust the specimen for identification. Identification to family level was done using a microscope and a variety of identification keys (e.g. Bellinger et al., 2024; Krantz & Walter, 2009; Srivastava Lab, 2012; Walter et al., 2014), with mites being identified to family level by Monica Young and springtails

identified to family level by me. All voucher specimens were databased and submitted to the Canadian National Collection of Insects, Arachnids, and Nematodes.

Bioinformatics and taxon assignment

Raw sequences were analyzed and grouped into operational taxonomic units (OTUs) using the '*JAMP'* package available in R (Elbrecht, 2018). Paired end merging used functions from *USEARCH* (Edgar, 2013) to match the forward and reverse reads with a minimum of 75% alignment. Primer trimming and length filtering were done using *Cutadapt* (Martin, 2011) with a maximum error of 0.2 and a sequence length of 301-319 base pairs. The paired end merging and primer trimming settings were modified from the default *JAMP* settings used for the pipeline to accommodate for more reads to merge, with the error filtering step intended to remove reads with low overlap later on. Error filtering was done using *USEARCH*, with an emax of 1 (0.3% error rate). The error filtering step of the pipeline was chosen based on the recommendation of Edgar (2015), who found that a 0.3% error rate was the optimal setting to use to filter out erroneous reads. OTU clustering was conducted using *VSEARCH* (Rognes et al., 2016) to filter out singleton reads and cluster reads based on a 3% similarity threshold (Alberdi et al., 2018). I also tested different length filtering settings: only including reads with the target 313 bp length or including reads in sets of three. Choosing to keep reads only within a certain sequence length of 301-319 base pairs was done by analyzing the distribution of sequence lengths to determine where most of the data was found and accounts for potential sequencing errors or the presence of pseudogenes that could be present within the samples (Song et al., 2008). Along with this, I also tested a stricter error filtering setting (Emax=0.2, 0.06% error rate), and two OTU clustering settings: a more conservative setting of 2% similarity threshold and a more lenient setting of 4% similarity threshold. Finally, any replicates that contained less than 100 reads or less than 3

mesofauna OTUs were excluded from the final dataset, as these were considered to be failed replicates.

Taxonomy was assigned to the resulting OTUs by searching the Barcode of Life Database (BOLD; boldsystems.org) using the BOLDigger program (Ratnasingham & Hebert, 2007). Since some OTUs could not be assigned to any taxon, BOLDigger separated OTUs into "classified" (at least to kingdom level) versus "unclassified" (no taxon assigned) categories. I then categorized the classified OTUs into "existing OTUs" (those with less than a 2% divergence from any reference sequence in BOLD) versus "novel OTUs" (those that were more than 2% divergent) (Young & Hebert, 2022). Abundance filtering was conducted on all OTUs to filter out any OTUs with less than 0.01% abundance using the *BiodiversityR* package in R (Kindt & Coe, 2005). For each OTU in the negative controls, the maximum read count was deducted from all other instances of that OTU to reduce the impact of any tag switching or contamination.

Except where noted, graphs and figures were constructed in R (v4.2.3; R Core Team, 2023). OTU accumulation curves were calculated for each molecular method for all taxa combined and for mesofaunal taxa using *BiodiversityR* (Kindt & Coe, 2005). The abundance filtered OTU dataset was converted to a presence-absence matrix to be used for most analyses. To assess the consistency of the extraction and PCR replicates, non-metric multidimensional scaling (NMDS) graphs were constructed using Sorenson's dissimilarity with the metaMDS function from the *vegan* package (Oksanen et al., 2015). PCR replicates for each sample were then pooled based on the following criteria: if there were three successful replicates, OTUs were retained if they were present in two out of three replicates. However if there were any PCR failures (i.e. there were only one or two successful replicates for a sample), OTUs were kept if they were found in at least one replicate. The extraction replicates for the SOIL samples were

similarly combined, with the three PCR replicates for each extraction first combined using the above criteria, and the three extraction replicates then combined.

Comparison of soil and specimen-based metabarcoding

A comparison of the taxa found using each molecular method was visualized using Krona charts (Ondov et al., 2011). To compare the amount of overlap in mesofauna OTUs, the number of shared mesofaunal OTUs were calculated for each paired sample using EstimateS (Colwell, 2009). Novel versus existing OTUs were compared at the order and family levels for all taxa combined, and for mesofauna (defined as springtails and mites) using *BiodiversityR* (Kindt & Coe, 2005).

The remaining analyses were conducted using *vegan* (Oksanen et al., 2015). Species richness was compared between molecular methods (SOIL vs BERLESE) by conducting paired t-tests on the mean number of OTUs, for all taxa combined and for the mesofauna. Prior to these analyses, I confirmed that data were normally distributed using the Wilk-Shapiro test ($p=0.92$). The community composition of each sample type was compared using hierarchical clustering based on an agglomerative algorithm using a complete linkage approach. Incidence-based Sorensen's dissimilarity metric was used for all taxa and for mesofauna by grouping together samples with similar communities. Permutational analysis of variance (permANOVA, 1000 randomizations) was then used to test whether there were significant differences between communities for each sample type. These same analyses were conducted for each of the different settings used in the *JAMP* pipeline. To assess whether community dissimilarity (based on Sorensen matrix) was correlated with geographic distance between sites, I conducted Mantel's test (10000 randomizations).

To compare family richness observed through morphological identification versus bulk specimen metabarcoding, mean numbers of families for mesofaunal taxa were calculated, and a paired t-test was performed. Data was normally distributed, according to the Wilk-Shapiro test ($p=0.40$). Community composition was also analyzed through a permutational analysis of variance (permANOVA, 1000 randomizations) to test whether there were differences between communities found for each method tested. To compare morphological identifications with molecular results, the mesofaunal OTUs were grouped and compared by family for each method.

Results

Sequencing summary and comparison of pipeline parameters

From the abundance-filtered sequencing data, 5,070 OTUs were recovered from more than one million reads within all the samples. OTU accumulation curves created for each method for all taxa and for mesofaunal taxa neared a plateau for the BERLESE replicates, whereas the curves continued to rise steeply for the SOIL samples (Figure 2.2). Although there were a few exceptions, extraction replicates (SOIL samples only) and PCR replicates (both methods) generally clustered together for each sample, indicating that pooling the replicates was appropriate (Figure 2.3). After pooling the replicates, 2,249 OTUs were found within the 12 paired samples, with the majority found in the SOIL samples (Table 2.1). Of the 166 mesofaunal OTUs in the final dataset, the majority were found in the BERLESE samples (135 OTUs) compared to the SOIL samples (75 OTUs). From the negative controls, 8821 reads representing 223 OTUs were recovered. The bulk soil and specimen samples contained higher reads counts (127-59K reads) than 81% of negative controls (<100 reads), with some negative controls containing higher read counts than expected (>1000 reads), which may be due to tag-switching.

Of the 2,249 OTUs found in both sample types, only 687 were able to be assigned to a taxon (= classified); these consisted of 31 classes, 55 orders, and 98 families. In the SOIL samples, 55% of the OTUs detected were identified as animals, with 35% of these being mesofauna (Acari and Collembola) (Figure 2.4). Within the SOIL samples, insect OTUs made up 16% of detected OTUs, with the majority (77%) only identifiable to the order level. For the BERLESE samples, animal OTUs made up 98% of detected OTUs, with Acari and Collembola making up the vast majority of these (75%) (Figure 2.4). Higher numbers of Acari and Collembola OTUs were detected in the BERLESE samples (135 OTUs) compared to the SOIL samples (75 OTUs). The majority of taxa were unique to one sample type, though the amount of overlap varied between samples (Figure 2.5). While many mesofaunal taxa were categorized as novel, the majority were "existing" (i.e. less than 2% divergent from any sequence on BOLD; Figure 2.6).

Reducing the error filtering setting (emax) to 0.2 resulted in 7% fewer total OTUs and 5% fewer mesofaunal OTUs than the original settings (Table 2). For the length filtering setting that only included reads with the target 313 bp length, there was a 22% reduction in total OTUs and a 6% reduction in mesofaunal OTUs. The length filtering setting that only included reads in sets of three consisted of a 3% reduction in total OTUs and a 5% reduction in mesofaunal OTUs. For the OTU clustering settings, the more conservative setting of 2% similarity threshold resulted in a 7% increase in total OTUs and a 14% increase in mesofaunal OTUs, and the more lenient setting of 4% similarity threshold resulted in a 5% reduction in total OTUs and a 7% reduction in mesofaunal OTUs (155 OTUs).

SOIL vs BERLESE diversity

The SOIL samples contained significantly more OTUs per sample (72.9 OTUs, SD=33) than were seen in the BERLESE samples (23.9, SD=10) for all taxa ($p<0.001$, Figure 2.8; Table 2.3). In contrast, fewer mesofaunal taxa were detected in the SOIL samples (9.1, SD=6) than in the BERLESE samples (19, SD=6); however, this difference was not significant ($p=0.998$; Figure 2.8; Table 2.3). Hierarchical clustering of the all-taxa dataset showed that sites cluster based on the molecular method used, with a significant difference in the two methods (permANOVA $R^2=0.15$, $p<0.001$; Figure 2.9; Table 2.4). In comparison, for the mesofaunal community, the same site for both methods often clustered together, regardless of which molecular method was used (Figure 2.9), although there was still a statistically significant difference between the two methods (permANOVA $R^2=0.08$, $p=0.005$; Table 2.4). For example, the paired samples (SOIL and BERLESE from sites 5G, 1C, and 11C) each clustered together, indicating similar communities (ditch samples clustering and field samples clustering) were detected by both methods. As well, the clustering pattern for sites 7A and 4J was similar for both SOIL and BERLESE samples, along with 10A and 10B which were sampled close together. The geographic distance between samples was also significantly correlated with the community dissimilarity (Mantel statistic $r=0.22$, $p=0.008$). For both species richness and community composition metrics for the different *JAMP* settings used, the number of OTUs and reads differed for each method tested (Table 2.2); however, the ecological trends remained the same. This is reflected in the species richness patterns, which were significantly different for all taxa, but not for mesofaunal taxa, for all of the parameters tested (Table 2.3). Community composition patterns were also similar to the original *JAMP* parameters used, with significant differences found between all taxa and mesofaunal taxa for all parameters tested (Table 2.4).

Morphological vs molecular identification

A total of 215 voucher specimens from the BERLESE samples were transferred to ethanol, with 158 of these subsequently slide-mounted (Table 2.5). Ethanol-preserved specimens included large oribatid mites, larger springtails and other invertebrates (e.g. earthworms, centipedes) (Table 2.5). For the slide mounted specimens, 73 specimens were Acari and 78 were Collembola. An average of 7.4 mesofauna families (SD=3) were morphologically identified per sample, whereas the sequencing resulted in significantly more mesofaunal families, with an average of 11.2 (SD=3) per sample ($p<0.001$). Significant differences were also found in community composition between these two methods (permANOVA $R^2=0.08$, $p=0.005$). There was surprisingly little overlap in family identity between the two methods (Figure 7); for Acari, there was some overlap for certain families, such as Oppiidae and Scheloribatidae, and for Collembola, there was overlap found for Onychiuridae, but most families were not captured by both methods. Some mesofaunal families, such as Onychiuridae and Eupodidae, were detected at a wider range of sites using molecular data, and sporadically from morphologically identified voucher specimens.

Discussion

Bioinformatic parameters

It is important to validate bioinformatic results by choosing appropriate parameters to provide confidence in the settings used to process metabarcoding data since these parameters can have a strong impact on the outcome of bioinformatic data and analyses (Alberdi et al., 2018; Edgar, 2015). In this study, changing parameters, such as varying the length filtering, error filtering, and OTU clustering through the bioinformatic processing steps, did change the total number of OTUs and reads for all taxa and for mesofaunal taxa. Nonetheless, species richness and

community composition remained similar, regardless of which setting was used, suggesting the metabarcoding results obtained were robust. Tag-switching, which involves nucleotide tags being switched during amplicon sequencing, is another factor that can impact metabarcoding data (Bohmann et al., 2022; Young & Hebert, 2022). This may explain the higher than expected read counts within some of the negative control samples; dynamic abundance filtering based on total OTU abundance was thus implemented to filter out these samples (Young & Hebert, 2022). Although it is good practice to consider various parameters for metabarcoding bioinformatics, they may not have a significant effect on the ability to accurately assess mesofaunal species richness and community composition.

Soil eDNA vs specimen-based metabarcoding

Previous studies comparing bulk specimen versus bulk soil samples found similarities in species richness and community composition between the two methods (Donhauser et al., 2023; Oliverio et al., 2018; Young & Hebert, 2022). Similarly to the study done by Young & Hebert (2022), this study detected more animal taxa in SOIL samples compared to BERLESE samples. However, metabarcoding soil eDNA can complicate data analysis, as noise, such as spurious OTUs or OTUs with erroneous designations, can be introduced into the dataset. In this study, the SOIL samples contained an unusually large portion of insect OTUs that were only identifiable to the order level. This is unexpected since taxonomic data for BOLD's reference library are extensive for insects (Hebert et al., 2016); the vast majority of insect sequences should at least be identifiable to family. The source of these unidentified sequences is unknown; one possibility is that they are due to chimera formation (Bohmann et al., 2022; Keck et al., 2022). In any case, probable spurious sequences such as these are likely artificially causing overestimation of diversity in these samples (Bohmann et al., 2022). While more animal taxa overall were detected

in the SOIL samples, a greater number of mesofaunal taxa were detected in the BERLESE samples. This may be due to the read depth not being high enough to capture the extreme diversity of taxa found in soil eDNA. Separating out the bulk specimens in the BERLESE samples before DNA extraction is advantageous, as it allows sequencing depth to be focused on the soil mesofaunal community, while limiting the noise obtained from eDNA in the soil. Although species richness differed between the two methods, there were some similarities in soil mesofaunal community composition between similar sites, a pattern also seen in previous studies comparing these methods for soil arthropods (Porter et al., 2019, Young & Hebert, 2022). While metabarcoding soil eDNA captures a wider range of taxa, the resulting data can be more challenging to interpret, whereas the Berlese funnel method detects a smaller array of mesofaunal taxa but appears to do so with fewer errors.

Many studies seeking to assess species richness and community composition of the arthropod community through metabarcoding have also detected non-target DNA (Basset et al., 2022; Donhauser et al., 2023; Hermans et al., 2022; Sire et al., 2023). In this study, both molecular methods detect many soil taxa. However, measurements for the SOIL samples included non-animal DNA (for bacteria and fungi). Since COI is not a standard genetic marker for these non-target groups, OTU designations for these groups are less informative than for animal species. As well, these non-target taxa take up sequencing depth, which results in less coverage of animal DNA. To address this challenge, some studies develop and use species-specific primer sets (e.g. nematode specific primers) (Heidemann et al., 2014; Read et al., 2006), but little research has tested specific primers for mesofauna (i.e. mite- or Collembola-specific primers). Although the primer sets chosen in this study were effective at amplifying DNA broadly across taxa, it may be helpful to employ more specific primers for taxa of particular

interest. This approach would reduce detections of non-target OTUs. While BERLESE samples captured a narrower range of taxa than the SOIL samples, a large portion of the OTUs found were from the soil arthropod community, with little representation by other groups (e.g. bacteria and fungi). This supports other studies which also found that isolating bulk specimens was beneficial for limiting metabarcoding noise in the dataset compared to working with soil eDNA (Arribas et al., 2016; Donhauser et al., 2023; Young & Hebert, 2022).

The bulk soil and bulk specimen samples each had their own logistical challenges. One drawback to using the SOIL method was that the associated DNA extraction process was very time-consuming. The BERLESE method targeted mesofauna more specifically, but some specimens may have dried out before falling into collection jars (Sabu et al., 2011). Specimens at immature or resting life stages may also have been under-represented in DNA analyses (Dritsoulas & Duncan, 2020). Berlese funnel extractions need to be performed rapidly after sample collection to minimize specimen mortality, which can impose logistical challenges related to equipment or personnel time. The choice of method may ultimately depend on the goal of the study. The SOIL method is clearly better at detecting non-motile OTUs, and therefore would be the better option if a wide range of taxa is being studied for a limited amount of samples. Conversely, the BERLESE method is more effective at targeting particular soil mesofauna groups.

Effects of geographical distance and habitat on the soil mesofaunal community

Environmental differences and distances between samples are associated with variation in mesofaunal species composition in this study, as has been observed in other ecosystems (Andújar et al., 2022; Arribas et al., 2020; Lehmitz et al., 2012). According to Gao et al. (2014), changes in environmental factors across the landscape can result in decreased mesofaunal community

similarity with increasing distance. Since mesofauna are operating at very localized scales, environmental changes over even small distances can influence and shape the mesofaunal community found at a particular location.

The number of mesofaunal specimens recovered through morphological identifications was lower than expected, which could be attributed to the environment at the time of sampling (Bardgett & Cook, 1998; Kamczyc et al., 2022) or the method used for specimen extraction. Because many of these samples were taken in agricultural environments that undergo cultivation, this can damage and destroy microhabitats of these soil organisms, resulting in a reduction in the population. Samples were collected in September near the end of an extended dry period, which could have impacted the amount of specimens found, consistent with previous work that showed that mesofaunal diversity was sensitive to short term environmental variation (Bardgett & Cook, 1998; Kamczyc et al., 2022). As well, since the soil was already very dry, the Berlese funnel extraction, which works by desiccating the soil and driving specimens downwards in search of moisture, may have been less efficient. Repeating this study earlier in the growing season would permit assessment of whether the same patterns hold true throughout the growing season.

Morphological vs molecular characterization of samples

Although previous metabarcoding work comparing morphologically identified specimens and sequenced specimens revealed similarities between families found in both methods (Young & Hebert, 2022), there was surprisingly little overlap between the morphologically identified specimens and the sequenced specimens in the BERLESE samples. Differences between morphologically identified specimens and sequenced specimens may have been due to the limitations of comparing specimens at the family level. Some specimens were only identifiable to higher taxonomic levels morphologically (e.g. order only) due to their life cycle stage (Dritsoulas

& Duncan, 2020), so the family level comparison is not comparable for these specimens. More mesofaunal families were found through sequencing than through morphological identification. One factor that may have caused this is the loss of smaller voucher specimens, such as certain mites (e.g. Eupodidae), during the DNA extraction steps. The small amount of overlap found for Collembola families, with more OTUs detected by sequencing, might be due to the presence of prey species in the guts of predatory mites (Eitzinger et al., 2013). In contrast, some morphologically identified voucher specimens were not detected using molecular methods. This is likely due to primer bias - since primers will not bind equally efficiently to COI of all taxa (Elbrecht & Leese, 2015), the primers used in this study could be failing to amplify certain taxa due to primer mismatch. As well, some sequences on BOLD may not have been identified to family, limiting the ability to assign taxa to all OTUs. Because of this discrepancy between morphologically identified and sequenced specimens, it would be optimal to use morphological identifications if possible; however, this is a time-consuming process that requires taxonomic expertise. Since many mesofaunal species have not been described yet, or lack taxonomic keys (Beaulieu et al., 2019; Lumley et al., 2013; Turnbull & Stebaeva, 2019), molecular methods are an effective solution for studies with a large number of samples. Nonetheless, more efforts are needed to describe mesofaunal species and to provide taxonomic resources for morphological identifications.

Tables

Table 2.1. Summary of OTUs and reads found for each sample type for all OTUs and classified mesofaunal OTUs.

Sample	Samples/ replicates	All OTUs & Reads						Mesofauna OTUs & Reads	
		Total OTUs	Total reads	Identified OTUs	Identified reads	Unidentified OTUs	Unidentified reads	Total OTUs	Total reads
Soil & Berlese	12/106	2249	956280	687	545196	1562	411084	166	319880
Soil	12/77	2102	510387	550	136503	1552	373884	75	9031
Berlese	12/29	206	445893	187	408693	19	37200	135	310849
Negatives	-/21	223	8821	80	6435	143	2386	35	718

Table 2.2. Summary of total and mesofaunal OTUs found in all samples for each JAMP parameter tested.

JAMP Parameter	Total OTUs	Mesofaunal OTUs
Error Filtering (Emax=0.2)	2101	158
Length Filtering (313 bp only)	1760	157
Length Filtering (pairs of 3 between 301-319)	2202	158
OTU Clustering (2%)	2401	192
OTU Clustering (4%)	2146	155

Table 2.3. Paired t-test results comparing all taxa and mesofauna taxa for all JAMP parameters tested. Paired Wilcoxon-signed rank test was used for data that was not normally distributed (WSRT).

	JAMP Parameter	**Test Statistic**	**P-Value**
All Taxa	Original	$t_2 = -5.9$	<0.001*
	Error Filtering	$t_2 = -4.4$	<0.001*
	Length Filtering (313 bp only)	$t_2 = -4.1$	<0.001*
	Length Filtering (pairs of 3)	$t_2 = -4.7$	<0.001*
	OTU Clustering (2%)	$t_2 = -3.1$	0.004*
	OTU Clustering (4%)	$t_2 = -3.8$	0.002*
Mesofauna Taxa	Original	WSRT	0.998
	Error Filtering	$t_2 = 5.6$	0.999
	Length Filtering (313 bp only)	$t_2 = -1.3$	0.11
	Length Filtering (pairs of 3)	$t_2 = -1.3$	0.11
	OTU Clustering (2%)	$t_2 = 6.5$	1.00
	OTU Clustering (4%)	$t_2 = 5.4$	0.999

*Significant p-values

Table 2.4. Permutational analysis of variance (permANOVA) results between SOIL and BERLESE samples for all taxa and mesofauna taxa for all JAMP parameters tested.

	JAMP Parameter	R^2	P-Value
All Taxa	Original	0.15	<0.001*
	Error Filtering	0.15	<0.001*
	Length Filtering (313 bp only)	0.15	<0.001*
	Length Filtering (pairs of 3)	0.15	<0.001*
	OTU Clustering (2%)	0.12	<0.001*
	OTU Clustering (4%)	0.13	<0.001*
Mesofauna Taxa	Original	0.08	0.005*
	Error Filtering	0.08	0.007*
	Length Filtering (313 bp only)	0.08	0.007*
	Length Filtering (pairs of 3)	0.08	0.02*
	OTU Clustering (2%)	0.08	0.007*
	OTU Clustering (4%)	0.08	0.009*

*Significant p-values

Table 2.5. Summary of voucher specimens found in BERLESE samples.

Voucher Specimens		Acari	Collembola	Other invertebrates
Alcohol	57	31	4	22
Slide mounted	158	73	78	2
Total	215	104	82	24

Figures

Figure 2.1. Berlese funnel extraction set-up.

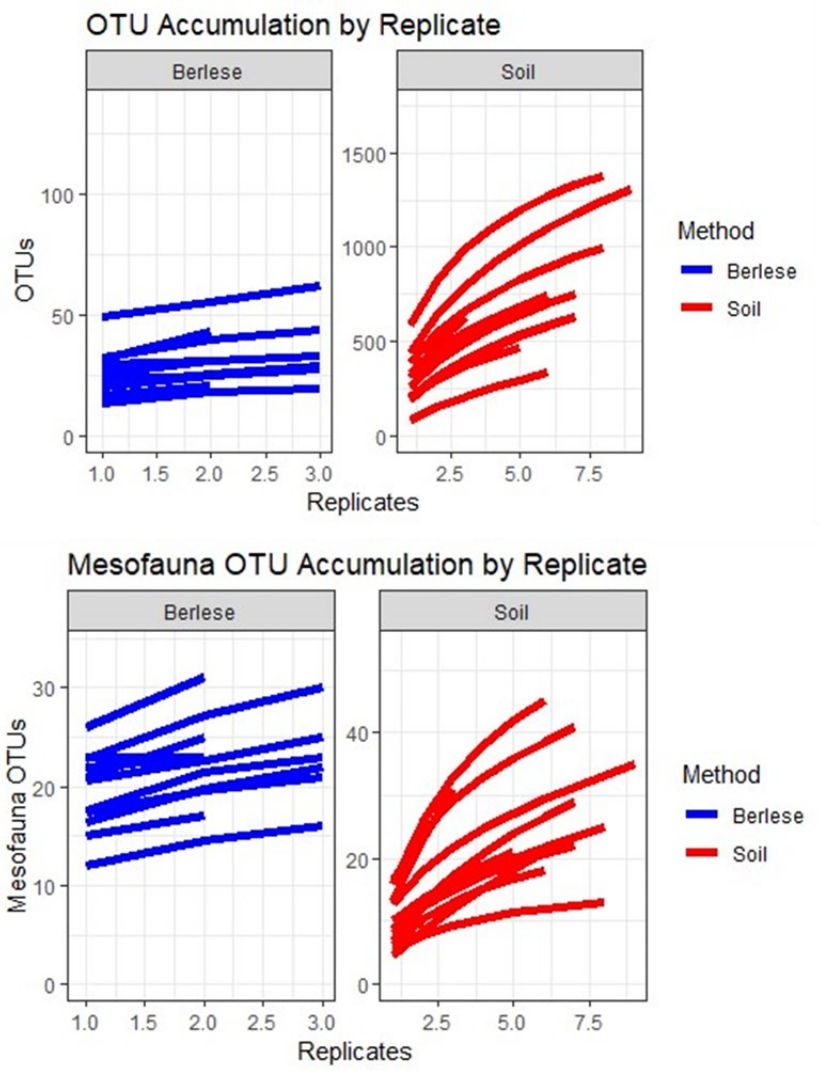


Figure 2.2. OTU accumulation curves (with 1000 randomizations) for 3 replicates for the BERLESE samples and 9 replicates for the SOIL samples summarized for all taxa and mesofaunal taxa detected for the 12 paired bulk soil and specimen samples.

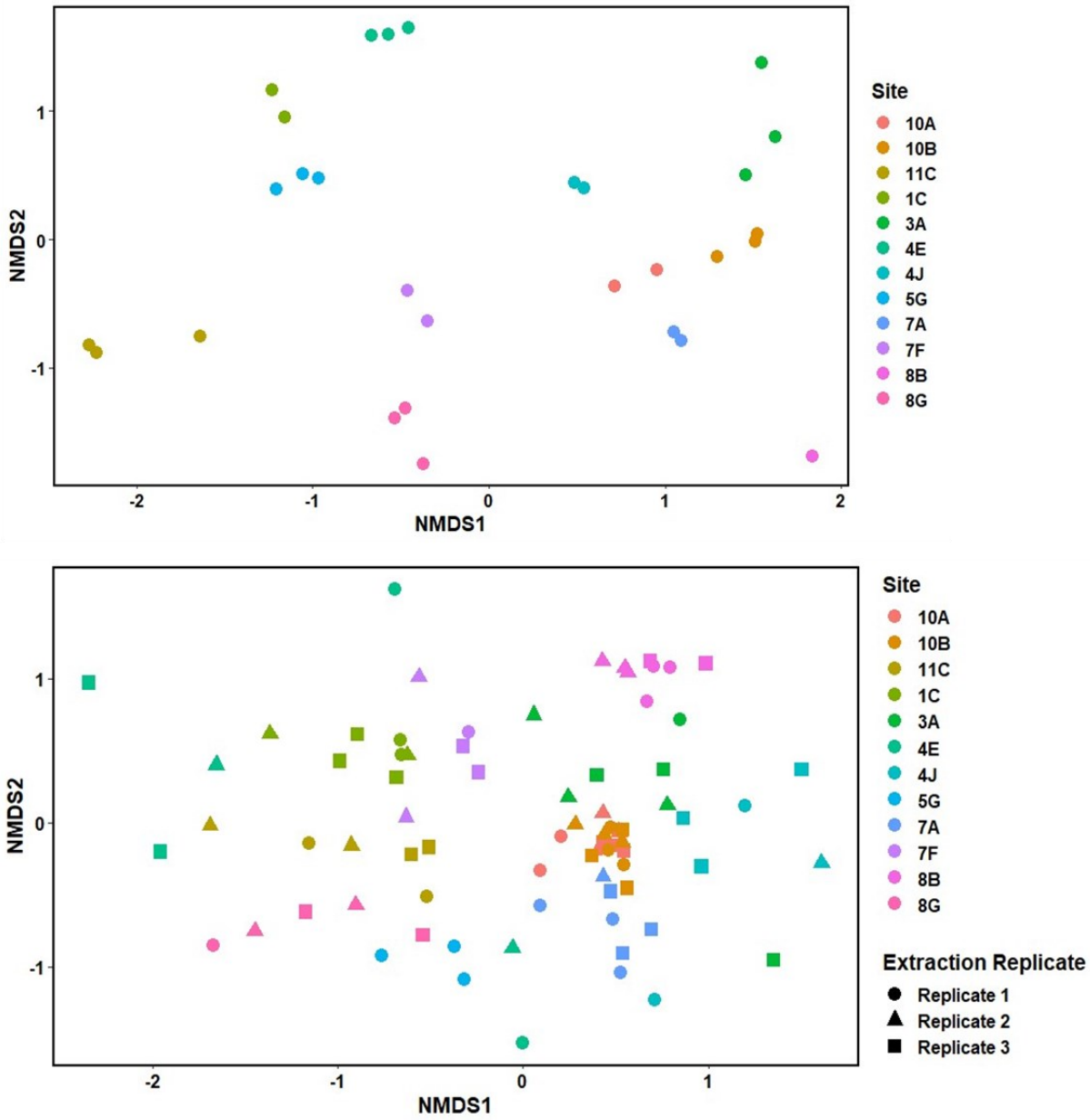


Figure 2.3. Differences in community composition (all taxa) for extraction and PCR replicates for BERLESE (top) and SOIL (bottom) samples using non-metric multidimensional scaling (NMDS) using Sorenson's dissimilarity values (stress=0.16).

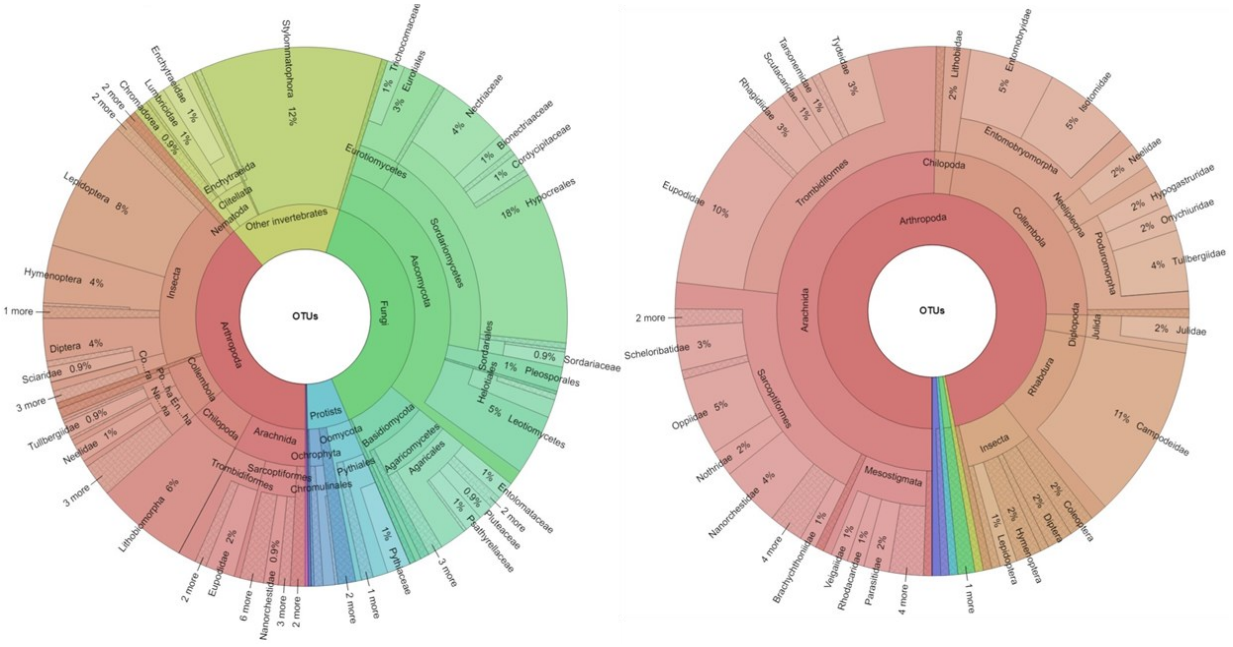


Figure 2.4. Classified OTUs with taxonomy assigned for SOIL (left) and BERLESE (right) samples.

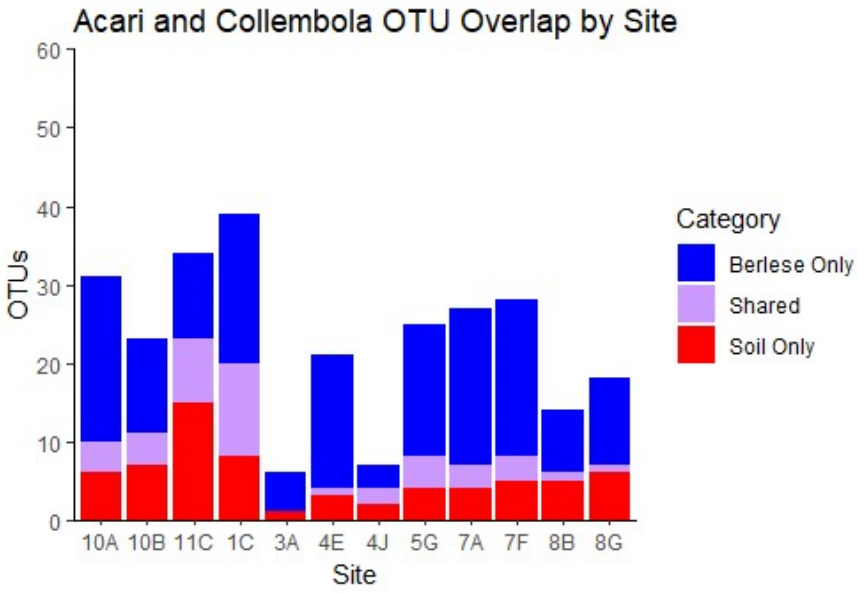


Figure 2.5. Overlap of Acari and Collembola OTUs per site between SOIL and BERLESE samples.

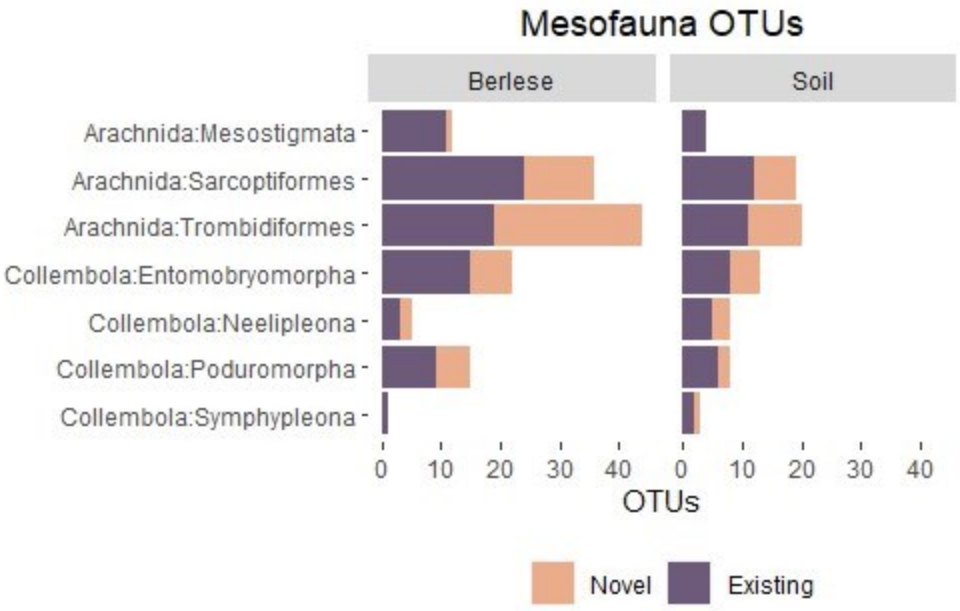


Figure 2.6. Comparison of novel and existing Acari and Collembola OTUs for SOIL and BERLESE samples.

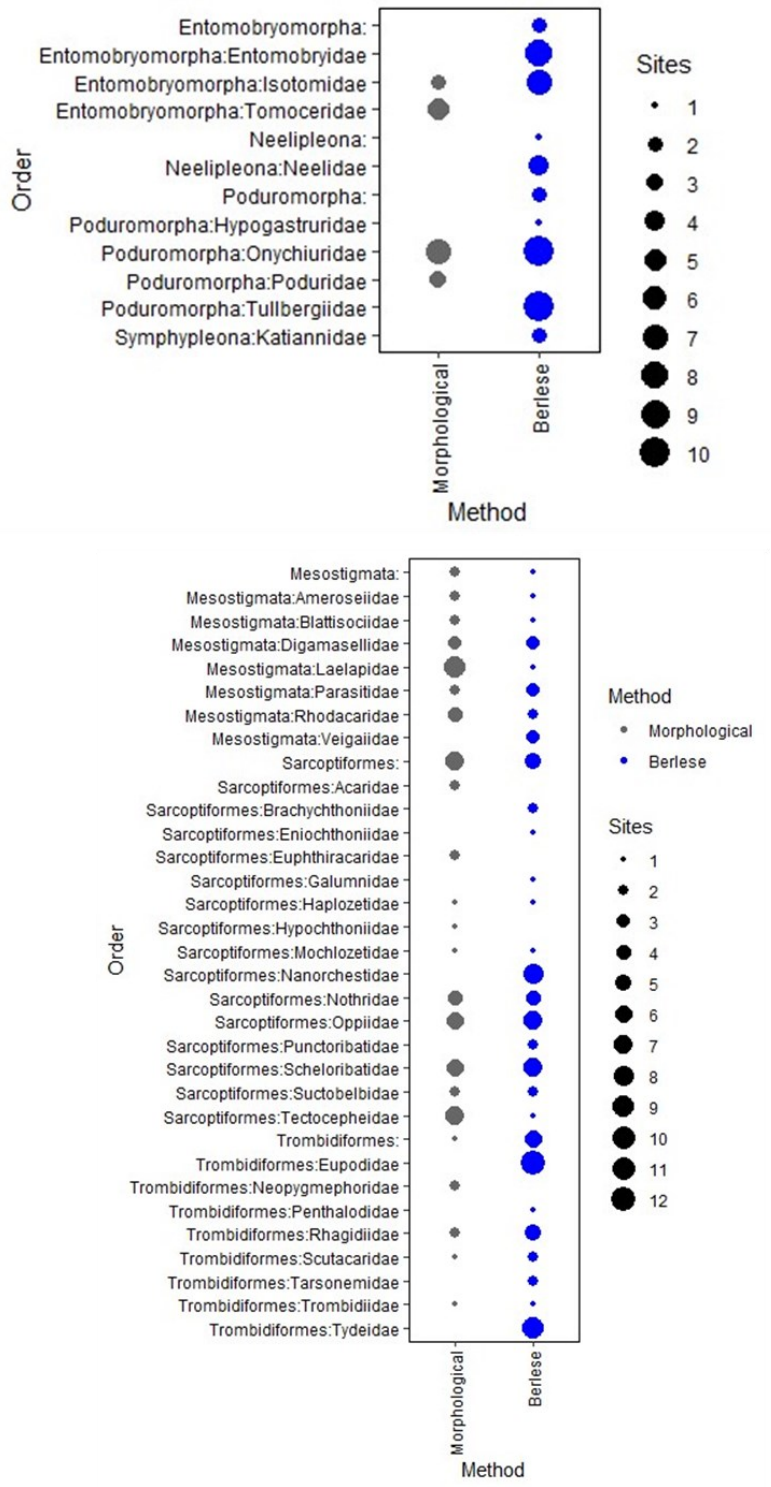


Figure 2.7. Comparison of morphological identifications of bulk mesofaunal specimens of Collembola (top) and Acari (bottom) with mesofaunal OTUs found in BERLESE samples. The size of each circle represents the number of sites at which each taxon was found.

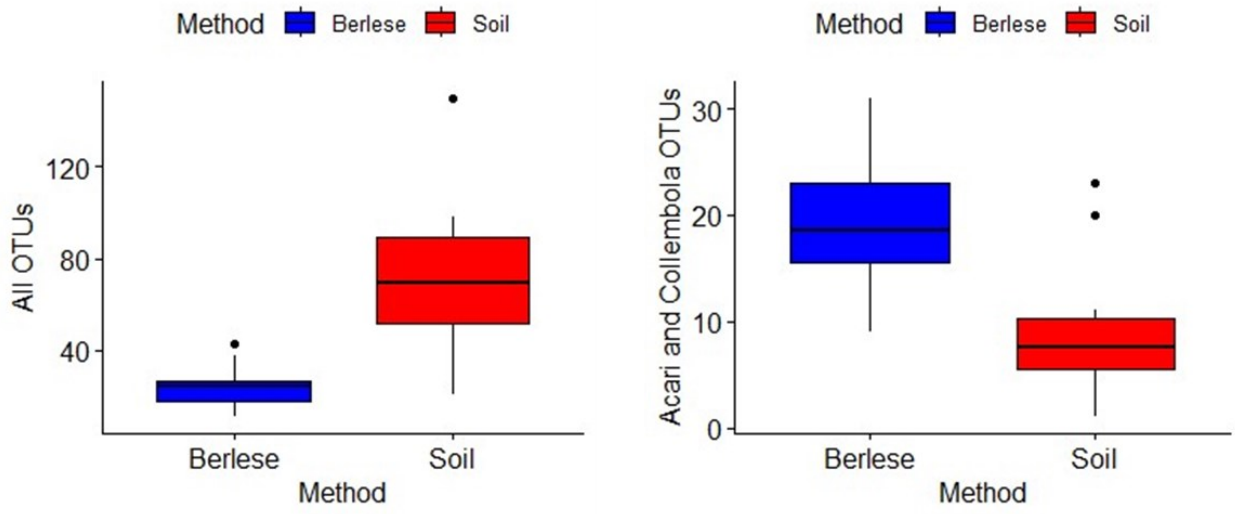


Figure 2.8. Comparison of total OTUs (left) and mesofaunal OTUs (right) between SOIL and BERLESE samples.

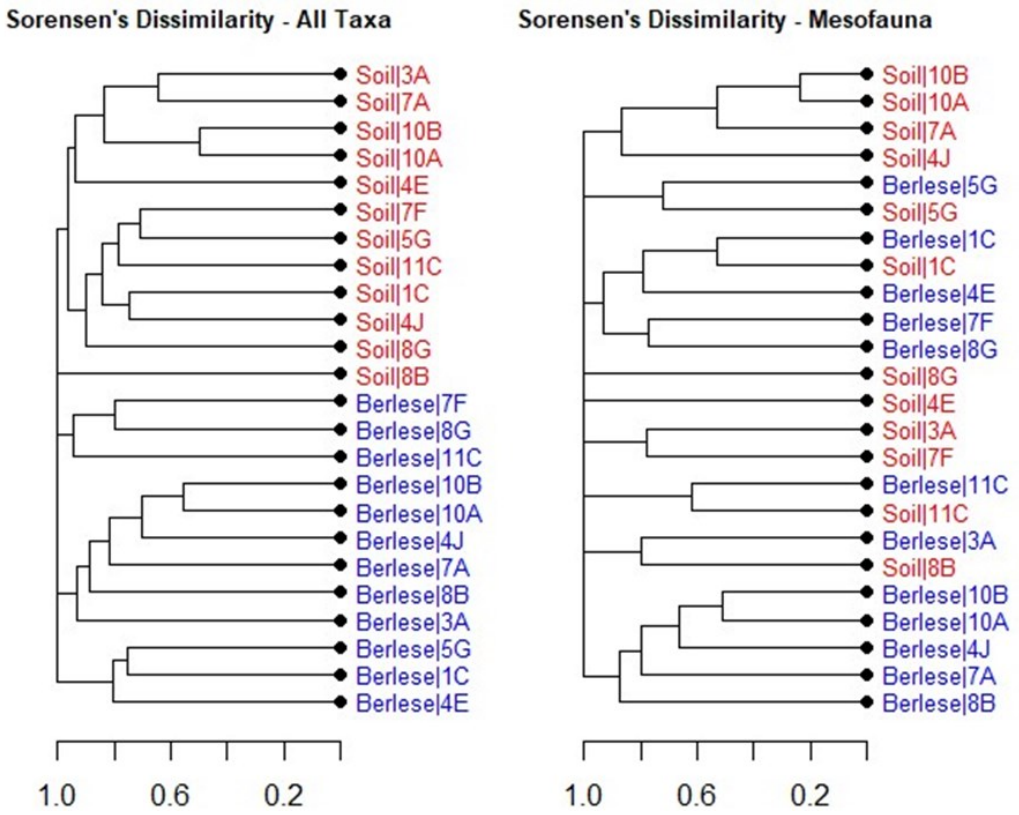


Figure 2.9. Hierarchical clustering for all taxa and mesofauna in SOIL and BERLESE samples.

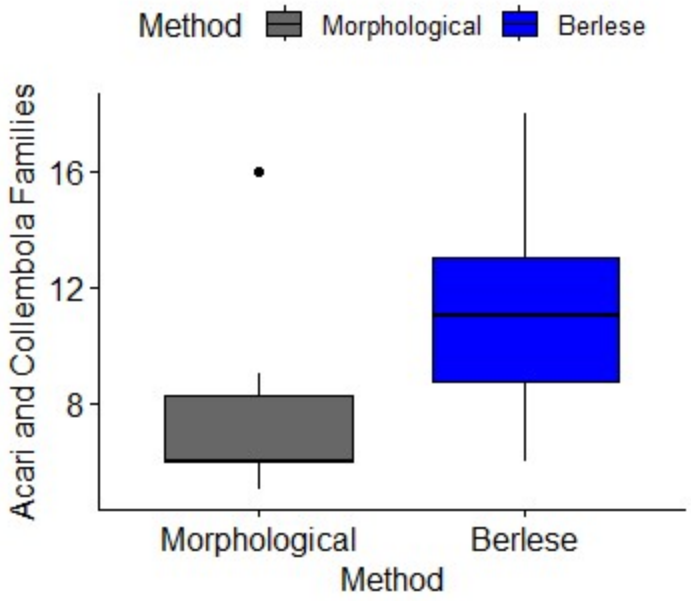


Figure 2.10. Mesofaunal families identified morphologically from voucher specimens and sequenced in BERLESE samples.

Chapter 3. Assessing the Role of Natural Capital in Maintaining Diverse Soil Mesofaunal Communities in an Agricultural Landscape

Abstract

Anthropogenic changes are a significant threat to the important ecosystem functions that soil biodiversity provides. Retaining natural capital within agricultural environments has been shown to improve biodiversity conservation. However, there are limited studies that address these effects on soil mesofauna due to their vast diversity and abundance. DNA metabarcoding is a practical tool that can be used to address this diversity. The main goals of this study are to measure the species diversity and community composition of the soil mesofauna and to compare soil physicochemical properties of sites within agroecosystems in Canada across the growing season. Six samples were collected from twelve sites of varying natural capital, ranging from grassy, untreed field margins to larger patches of forest, across three sampling periods in the summer from agricultural sites in southeastern Canada. The mesofaunal specimens were Berlese-extracted, and then non-destructively DNA-extracted. Next-generation sequencing of COI was conducted on the bulk specimen samples, resulting in over 1,000 OTUs being recovered. From this, a baseline survey of the soil mesofauna present within these agroecosystems was collected. Species diversity varied between habitat types and sites, whereas similar patterns were found in mesofaunal communities between forested sites, regardless of size. Soil physiochemical properties such as organic matter and moisture, were shown to influence mesofaunal communities, but seasonal differences did not impact mesofaunal communities. This study shows the value that retaining natural capital has on preserving biodiversity, and also highlights a need for more taxonomic efforts to be made to update and verify reference libraries.

Introduction

Soil is home to a vast number of organisms, including microorganisms (e.g. bacteria, fungi, protists); microfauna (< 200 μm in size, e.g. nematodes, rotifers); mesofauna (~0.2 – 2 mm, e.g. mites, springtails, enchytraeids); and macrofauna (> 2mm, e.g. earthworms, snails, many insects) (FAO, 2020; Lavelle et al., 2006). This enormous biodiversity is critically important in maintaining essential ecosystem functions, such as soil formation, nutrient cycling, and carbon transformation (Barrios, 2007; Lavelle et al., 2006). In both natural (e.g. forests, grasslands) and human-dominated landscapes (e.g. agricultural ecosystems), the complex interactions between these soil organisms provide the nutrients required for plant growth, as well as impacting global processes, for example, through their effects on carbon cycles in the soil (Barrios, 2007; Lavelle et al., 2006).

Soil mesofauna include a wide variety of taxa, but are dominated by small arthropods, in particular mites (Acari) and springtails (Collembola) (Behan-Pelletier, 2003). Springtails (Collembola) are small hexapod arthropods that impact the soil food-web by feeding on microorganisms and through the breakdown of organic matter (Coulibaly et al., 2019). As well, they are an important prey source for other small arthropods, such as beetles, mites, and spiders (Bilde et al., 2000). Mites (Acari) are one of the most diverse groups within the soil community (Beaulieu et al., 2019), representing a wide range of ecological roles. For example, predatory mites, such as Mesostigmata, influence prey populations, are used as biological control for insect pests, and can influence the soil community through top-down effects (Koehler, 1999; McMurtry et al., 2015; Petrova et al., 2004; Schneider & Maraun, 2009). In comparison, oribatid mites are responsible for nutrient cycling within the soil by feeding on dead organic matter and microbial

residues, and contribute strongly to the soil structure by producing nutrient rich faecal pellets that act as a natural fertilizer for the soil (FAO, 2020; Potapov et al., 2022).

Because of the complexity of soil biodiversity, it is essential to monitor and collect data about the effectiveness of soil conservation efforts in order to inform policy makers to create more sustainable practices for the conservation of soil biodiversity (FAO, 2022). National and international organizations are raising awareness about the importance of soil biodiversity, encouraging the incorporation of sustainable practices into national policies, and are working towards consistent data standards to assess whether management practices ensure adequate conservation of the soil community (FAO, 2022). The Global Soil Biodiversity Initiative (GSBI) brings together scientists worldwide to discuss the protection of soil biodiversity through management and policy. The Soil BON Foodweb Team is also a global network that focuses on collecting data for soil fauna all over the world to understand how soil ecosystems function and which conservation practices should be implemented (Potapov et al., 2022). However, inventory and monitoring of soil mesofauna is challenging due to their abundance and diversity (FAO, 2022). For this reason, it is important to collect baseline taxonomic information about understudied groups of soil organisms, such as the mesofauna. Surveys focusing on soil mesofauna are not very common (Arroyo et al., 2013; George et al., 2017; Kamczyc et al., 2022; Meehan et al., 2019; Porco et al., 2013; Rueda-Ramírez et al., 2022), so there is a need to fill in this taxonomic knowledge gap on this important group of organisms.

Threats to global biodiversity, such as habitat disruption and fragmentation, agricultural intensification, and soil erosion, can result in a reduction in the important ecosystem functions that the soil community provides (Gardi et al., 2013; Tibbett et al., 2020). Land use, such as agricultural intensification, has been found to result in a large reduction of soil fauna diversity

due to tillage and pesticide use (Decaëns et al., 2006; de Graaff et al., 2019). While there is increasing emphasis on within-field methods to preserve soil biodiversity (e.g. low tillage systems, cover cropping, etc.) (Carpio et al., 2019; van Capelle et al., 2012), retaining natural areas in the agricultural matrix is another management technique that has been found to benefit biodiversity (Montgomery et al., 2020; Morris et al., 2010). Agricultural landscapes often contain uncultivated ditches along the edges of fields, ranging from small grassy verges to larger forested habitats (Needelman et al., 2007), which have the potential to help preserve soil biodiversity (Shaw et al., 2015). Montgomery et al. (2020) found that hedgerows established on the edges of agricultural farms improve plant, butterfly, and bird species richness. Stroot et al. (2022) also found that wildflower strips promote plant species richness amongst agricultural landscapes. Udawatta et al. (2019) found that retaining natural capital in agroecosystems resulted in an increase in avian, earthworm, and insect species richness. Although many studies have investigated plant species richness and above-ground arthropod diversity within these ditches (Bennewicz & Barczak, 2020; Morris et al., 2010; Schirmel et al., 2016), there is limited research into whether varying degrees of natural capital, ranging from untreed field margins to larger forested areas, support mesofaunal diversity within agricultural landscapes.

Soil faunal diversity has been shown to change across the growing season (Bardgett & Cook, 1998; Wu & Wang, 2019). This is mainly due to changes in environmental factors across a season, such as amount of rainfall, which impact the physical and chemical properties of the soil (Decaëns et al., 2006; de Graaff et al., 2019). For example, seasonal changes in moisture and organic matter have been shown to impact soil community and abundance (Crotty et al., 2023; Wu & Wang, 2019). Some mesofaunal groups, such as oribatid mites, have also been shown to be sensitive to changes in soil moisture (Gergócs & Hufnagel, 2009). Although these seasonal

effects have been studied for soil faunal diversity within forest ecosystems (Wu & Wang, 2019), there is little research that exists for whether seasonal turnover impacts soil mesofaunal diversity and communities within agricultural ecosystems. It is important to understand whether soil mesofaunal communities change across the growing season for future studies that aim to survey these communities over a long period of time. If there are not strong seasonal patterns, then this could allow for sampling over a greater period of time without compromising the data.

Due to the abundance and diversity of soil mesofauna, assessing species richness and composition of the entire community is not feasible through morphological methods, with most studies focusing on selected taxa (Baumann, 2021; Postma-Blaauw et al., 2012), or else only identifying organisms to high taxonomic levels (Meehan et al., 2019; Ruf & Beck, 2005). In contrast, DNA metabarcoding is a method that uses DNA taxonomy and high-throughput sequencing in order to identify multiple species within a given sample (Coissac et al. 2012). It shows great promise for efficiently and accurately assessing soil communities (Boggs et al., 2019; Yang et al., 2014), including the soil mesofauna (e.g. mites; Arribas et al., 2016; Hoage, 2018; Young and Hebert, 2022; and springtails; Saitoh et al., 2016). Metabarcoding has been used to assess soil arthropod communities within a variety of different habitats, such as in the Arctic, and in temperate, tropical, freshwater, and grassland habitats (Arribas et al., 2016; Basset et al., 2022; Oliverio et al., 2018; Schenk et al., 2020; Young & Hebert, 2022). It can also be used on different types of samples, including bulk specimens and bulk soil samples (Arribas et al., 2016; Young & Hebert, 2022). DNA metabarcoding is a costly, but useful tool that can be used to assess the vast diversity of mesofauna to compare and contrast the variables that influence their biodiversity.

The objectives of this study are a) to conduct a baseline survey of soil mesofauna (Acari and Collembola) in ditch edges and forest patches in southeastern Ontario; b) to compare soil mesofaunal diversity and community composition under various levels of natural capital (ranging from grassy un-treed field margins to large forested sites) in an agricultural matrix, using specimen-based metabarcoding; c) to test for relationships between soil physicochemical parameters (e.g. organic matter, nutrients) and the mesofaunal community at each site; and d) to determine whether there are seasonal differences in the soil mesofaunal community.

Methods

Sample collection

I collected soil samples for this study in agricultural ditch margins and forested areas following waterways in southeastern Ontario during the summer of 2022 (Figure 3.1; Table 3.1). To analyse the role of small habitat patches, referred to as "natural capital" (NC), in structuring the mesofauna community, I selected four levels of NC, with three replicates of each type, for a total of 12 sampling sites (= "habitat analysis"). The categories were based on vegetation type (e.g. forested vs non-forested) and the width of the forested area. The unforested sites consisted of: "No trees" = agricultural ditches with no shrubs and trees and less than 11 m in width between adjacent fields and "Scattered trees" = agricultural ditches of a similarly narrow width but with a few scattered shrubs and trees (< less than 20% tree canopy cover). Although vegetation surveys were not conducted for each site, observationally, the sites were dominated by grasses, with varying amounts of other herbaceous plants (e.g. *Pastinaca sativa* (wild parsnip), *Solidago* spp. (goldenrod), etc.). The forested sites consisted of: "Small forest" = small patches of forest ranging from 30-300 m in width, and "Large forest" = large patches of contiguous forest ranging from 400-2000 m in width (Figure 3.2). These mixed-wood sites were

somewhat variable in terms of slope, tree composition, and understory vegetation due to the availability of suitable sites. The soil from the majority of sites consisted of clay or loam texture, with some unforested sites containing more sandy or silty soil texture (Table 3.1). Six points were randomly selected along each of the twelve sites, with three on each side of the ditch/waterway (Figure 3.3). Samples were collected in the month of June, for a total of 72 paired samples (12 locations x 6 sites). To assess whether the mesofauna community changes over the growing season, one representative of each habitat type was selected to account for a range of habitats, and three points along each ditch/waterway were sampled in June, July, and August, for a total of 36 paired samples (4 locations x 3 sites x 3 sampling periods) (= "seasonal analysis").

At each sampling site, two soil samples were collected immediately adjacent to one other, with one of each set used for mesofaunal metabarcoding, and the second used for physicochemical analysis. Each of these samples was collected using a tulip bulb planter to extract two soil cores (each approximately 6 cm in diameter and 11 cm deep) that were combined in a Ziploc bag. Tree canopy cover was measured using a convex spherical crown densiometer. Four measurements were taken, one at each cardinal direction, and the mean of these measurements was then calculated. The mesofaunal samples were kept in the fridge overnight, and then specimens were extracted on Berlese funnels lined with cheesecloth for 4 days using heat from a 60W bulb, where specimens were collected into 95% ethanol. Specimens were filtered through a 45-µm sieve into 2 mL tubes filled with 95% ethanol and stored at -20^{o}C until DNA extraction. Meanwhile, the physicochemical samples were stored at -20^{o}C prior to being processed at the Harrow Research and Development Centre in Dr. Lori Phillips' lab. The samples were first thawed at room temperature for an hour before being sieved through a 2 mm

mesh. After sieving, gravimetric moisture was measured by drying the samples at 105ºC for 48 hours. A sub-sample was then air dried for one week and sent to A&L Canada Laboratories Inc. for physicochemical analysis, which included measurements of organic matter, cation exchange capacity (CEC), phosphorus, and percentages of potassium, magnesium, and calcium.

Molecular methods

DNA was extracted non-destructively from the mesofaunal specimens using a Qiagen DNeasy Blood and Tissue kit (Qiagen), with slight modifications to allow for voucher recovery. In particular, the ethanol was first evaporated from the tubes containing specimens using a vacuum centrifuge (Vacufuge, V-AL, 10 mins). 270 µl of buffer ATL and 30 µl of proteinase K were then added to the tubes without grinding the specimens, and tubes were incubated overnight at 56ºC. Following incubation, 200 µl of solution was withdrawn from the top and placed into a 1.5 mL tube, being careful not to disturb or pipette up any specimens. The tubes containing the voucher specimens were then refilled with 95% ethanol. The rest of the DNA extraction steps followed the standard kit protocol, with the final elution step repeated once to maximize DNA yield, for a total elution volume of 200 µl. Following the extraction, the spin column filters were rinsed with 95% ethanol and inspected under a microscope to check for any specimens that were drawn up with the extraction buffer; any specimens found during this step were added to the voucher specimen tube. Initial PCR tests had high failure rates, possibly due to PCR inhibitors from soil residue that was extracted with the specimens. Diluting the DNA improved PCR results, and it was determined that a 20% dilution was most successful across samples; therefore, all DNA extracts were diluted with water to 20% prior to sequencing.

A 313 bp region of COI was targeted, based on the recommendation of Krehenwinkel et al. (2016), and using the following primers: mICOIintF (Brandon-Mong et al., 2015) and FoldR

(Arribas et al., 2016), with modified adapters that had 0-6 "N" bases between the target region of COI and the adapter sequence to increase base diversity. PCR was replicated 3 times for each DNA extract; therefore, a total of 216 samples were sequenced for the habitat analysis (12 locations x 6 sites x 3 PCR replicates), while 108 samples were sequenced for the seasonal analysis (4 locations x 3 sites x 3 months x 3 PCR replicates). Two extraction negatives x 3 PCR replicates plus 10 PCR negatives were also sequenced. Each PCR reaction consisted of 16.75 µl nuclease-free water, 2.3 µl 10X Titanium buffer, 2.3 µl of 2 mM dNTPs, 0.575 µl each of the forward and reverse primers, 0.5 µl Titanium Taq, and 2 µl of DNA. Cycling was conducted in an Eppendorf thermocycler, with the following program: 95ºC for 3 minutes, 40 cycles of 94ºC for 30 seconds, 45ºC for 90 seconds, 72ºC for 90 seconds, and a final extension of 72ºC for 3 minutes, before being held at 4ºC. The remaining steps in the protocol were conducted at the Molecular Technologies Laboratory of the Ottawa Research and Development Centre. In brief, samples were purified and normalized using an NGS Normalization 96-Well Kit (Norgen Biotek Corporation, Canada). A second round of PCR was then performed wherein samples were tagged using i3 and i5 indices (Illumina, San Diego, CA, USA) of the Nextera Index Kit (Illumina, San Diego, CA, USA), followed by PCR clean-up and normalization as above. The samples were then pooled, quality was assessed using a Tapestation (Agilent, Santa Clara, CA, USA), and quantification was done using a KAPA Library Quantification Kit (KAPA Biosystems, Wilmington, MA, USA). Finally, next-generation sequencing was performed using a single MiSeq v3 600 Kit (Illumina, San Diego, CA, USA).

Bioinformatics and taxon assignment

The raw sequence data were analyzed and grouped into operational taxonomic units (OTUs) using the '*JAMP'* package available in R (Elbrecht, 2018). The parameters of the

pipeline were chosen based on the results of the Chapter 2 analyses. In brief, paired end merging was conducted with *USEARCH* (Edgar, 2013), with a minimum 75% alignment match. Primer trimming (maximum error of 0.2) and length filtering (301-319 bp) were done using *Cutadapt* (Martin, 2011). Error filtering was conducted with *USEARCH* (Emax=1). Finally, singleton reads were discarded and OTUs were clustered with 3% similarity using *VSEARCH* (Rognes et al., 2016). In addition, any replicates that contained less than 5 mesofaunal OTUs were considered a failed replicate and were excluded from the dataset. Two sampling sites (Sites 21-3 and 21-6) were also excluded, as they were outliers with far more than the expected number of OTUs; examination of the voucher specimens confirmed that the sequencing data greatly overestimated the number of mesofaunal OTUs. The identified voucher specimens from these samples were databased and submitted to the Canadian National Collection of Insects, Arachnids, and Nematodes.

Following OTU clustering, taxonomy was assigned using the BOLDigger program (Ratnasingham & Hebert, 2007) and the Barcode of Life Database (BOLD). OTUs were designated by BOLDigger as classified (at least to kingdom level) or unclassified (no taxon assigned). Abundance filtering was performed on all OTUs, filtering out any OTUs with less than 0.01% abundance using the *BiodiversityR* package in R (Kindt & Coe, 2005). To reduce the possible impact of tag switching or other sources of contamination, the maximum read count for each OTU in the negative controls was subtracted from all other instances of that OTU. PCR replicates were then pooled for each sample, only including OTUs that were present in at least two out of three PCR replicates; however, in instances where one or two PCR replicates failed, OTUs were kept if they were in a single replicate. OTU read counts were then converted to

presence-absence data, and R (v4.2.3; R Core Team, 2023) was used to construct figures and for the majority of statistical analyses.

Habitat analysis

Mesofaunal species richness of the different habitats was compared using nested analysis of variance (ANOVA), after first confirming the data were normally distributed using the Wilk-Shapiro test (p=0.99). The main factor was habitat, with sites nested within habitat. If significant differences were found, I performed post-hoc testing using Tukey's HSD test to determine which sites differed from each other. Mean mesofaunal OTUs at the family level to assess differences in species richness for each habitat type were graphed using *BiodiversityR* (Kindt & Coe, 2005). I tested whether the mesofaunal community composition differed between habitat types with a nested permutational analysis of variance (permANOVA, 1000 randomizations) using the *vegan* package (Oksanen et al., 2015), based on the incidence-based Sorensen's dissimilarity metric. I conducted post-hoc testing using a nested pairwise permANOVA test with the *pairwiseAdonis* package (Martinez, 2020). Overlap in mesofaunal OTUs between the three sites for each habitat type were compared using Venn diagrams constructed with the *BioVenn* package (Hulsen, 2021). In order to visualize differences in mesofaunal community composition and the soil physicochemical properties for each site, non-metric multidimensional scaling (NMDS) using Sorenson's dissimilarity calculations for community indices and environmental variables were presented based on correlation vectors with the metaMDS function using *vegan* (Oksanen et al., 2015), with the number of dimensions for the NMDS decided based on testing the stress value for the NMDS by graphing the increasing the number of dimensions by stress and determining where a plateau was reached.

Soil physicochemical properties

To determine whether soil physicochemical properties differed across habitats, I conducted a nested analysis of variance (ANOVA), after first testing for normality by graphing Q-Q plots (quantile-quantile plot) in R. If significant differences were found, Tukey's HSD test was performed to determine which habitat types and sites within a habitat type were significantly different from each other. Mantel's test (10000 randomizations) was used to assess for correlations between geographical distance matrices and soil physicochemical properties matrices, with community composition using Sorenson's dissimilarity based on mesofaunal OTU presence-absence, using *vegan* (Oksanen et al., 2015). These traits were also included in the NMDS graphs using correlation vectors to map out how strongly each property correlates with the community composition at each site.

Seasonal analysis

I compared the mean mesofaunal OTUs at the family level for each site across the three months sampled using *BiodiversityR* (Kindt & Coe, 2005) as well. To compare the mesofaunal community composition across sampling periods, I conducted a repeated measures nested permutational analysis of variance (permANOVA, 1000 randomizations) using *vegan* (Oksanen et al., 2015). To visually assess any patterns in seasonality, I conducted an NMDS analysis based on Sorenson's dissimilarity matrix for community indices with the metaMDS function using *vegan* (Oksanen et al., 2015), using the same process as described above to decide on the number of dimensions used.

Results

Overall diversity and community composition

After abundance-filtering, a total of 1,218 OTUs were found from over 800,000 reads. Once negative reads were removed and replicates were pooled, a total of 1,024 OTUs were recovered from all 12 sites; however, only 695 of these could be assigned to any taxon (Table 3.2). These classified OTUs consisted of 23 classes, 42 orders, and 119 families. Of the animal taxa found, 75% were not assigned to species level, with 44% of families having no OTUs assigned to the genus level, and 69% of families having no species-level OTU assignments (Table 3.2). A total of 443 mesofauna OTUs were detected, of which 70% were Acari and 30% were Collembola (Table 3.2). Within Acari, Trombidiformes (primarily Prostigmata) and Sarcoptiformes (primarily Oribatida) were each far more OTU-rich than Mesostigmata (Figure 3.4). The most OTU-rich families were Eupodidae (71 OTUs), Tydeidae (30 OTUs), and Nanorchestidae (26 OTUs), while 13 families were only represented by a single OTU. Fifty-two mite OTUs (12%) were assigned to species, 87 (20%) were assigned to the genus level, while the majority, 206 OTUs (47%), could only be assigned to family (Table 3.2). For the Collembola, the orders Entomobryomorpha and Poduromorpha contained the majority of OTUs in comparison with the other orders (Figure 3.5). The most OTU-rich families were Isotomidae (49 OTUs) and Tullbergiidae (20 OTUs), while 3 families were only represented by a single OTU. Forty springtail OTUs (22%) were assigned to species, 45 OTUs (25%) were assigned to the genus level, and 70 OTUs (39%) could only be assigned to family (Table 3.2).

Habitat analysis

From this dataset, 58 replicates constituted failed replicates and thus were removed; this resulted in the complete removal of some samples (two samples from Site 58 and one sample

from each of Sites 333, 334, and 347). There was no significant difference in the number of mesofaunal OTUs between each of the habitat types (p=0.64; Figure 3.4a), but significant differences were found between sites within a habitat type (p=0.03) (Table 3.4). However, post-hoc testing revealed the only significant difference was between two of the no-trees sites (Site 21 and 1, p=0.04), with none of the other sites differing significantly. Comparing the mean OTU richness at the order level also did not reveal any pattern with respect to habitat, with all taxa having highly variable species counts across sites (Figure 3.4b, 4c).

The mesofaunal community composition did, however, differ significantly between habitats (R^2=0.10, p <0.001), as well as between sites within a habitat type (R^2=0.21, p <0.001). Post-hoc testing indicated that all habitat types and all sites within a habitat type had significantly different mesofaunal community composition (p <0.001). This is supported by how few taxa overlap between different sites, even within the same habitat type (Figure 3.5). Despite these results, the NMDS analysis showed extensive overlap between the sites and habitat types; there was however a distinction between unforested (no trees and scattered trees) and forested (large forest and small forest) sites along the NMDS2 axis (Figure 3.6).

Samples within each site mostly clustered together, however there were several exceptions. For example, the samples from Sites 21 (no trees), 115 (small forest), and 333 (large forest) were widely dispersed across the ordination. The samples from Site 1 (no trees) form two distinct clusters, with four samples in one cluster among the unforested sites and two samples clustering towards the forested sites. The samples from Site 7 (scattered trees) form three clusters, two amongst the unforested sites, and one sample among the forested sites. Samples from Site 347 (large forest) also formed two distinct clusters, with three samples in one cluster

and two samples in the other. One sample from Site 334 was also highly distinct from the rest of the site, and the six samples from Site 343 formed 3 distinct clusters.

Environmental factors

There were significant differences in tree cover ($p=0.001$) and percentage of potassium ($p=0.02$) between habitat types (Table 3.5), with post-hoc tests indicating most differences were between forested and unforested sites. For all other soil parameters, however, soil properties did not vary with habitat. In comparison, all soil parameters differed significantly between sites within a habitat (Table 3.5). There was a significant correlation between community composition and organic matter (Mantel statistic $r=0.17$, $p=0.002$), cation exchange capacity (CEC) (Mantel statistic $r=0.14$, $p=0.008$), percentage of magnesium (Mantel statistic $r=0.14$, $p=0.004$), and calcium (Mantel statistic $r=0.14$, $p=0.005$) (Table 3.6). As well, community composition was significantly correlated with the geographic distance between samples (Mantel statistic $r=0.16$, $p<0.001$) and tree cover (Mantel statistic $r=0.08$, $p=0.001$; Table 3.6). These patterns were apparent in the NMDS graph, which depicts tree cover, gravimetric moisture, and organic matter correlation vectors being associated with NMDS2, with all three increasing towards the forested sites (Figure 3.6). Tests of correlation indicated gravimetric moisture and cation exchange capacity (CEC) were positively correlated with organic matter, and percentage of magnesium was negatively correlated with organic matter (Figure 3.7; Appendix 2); this can also be seen in the NMDS graph, as the percentage of magnesium and organic matter correlation vectors point in opposite directions. Gravimetric moisture and CEC correlation vectors were pointing in the same direction, however, on opposite sides of the graph. Percentage of potassium was also negatively correlated with gravimetric moisture (Figure 3.7; Appendix 2), which can be seen in the NMDS graph, as these correlation vectors are pointing in opposite directions. Although tree

cover and gravimetric moisture were not found to be correlated, the correlation vectors pointing in the same direction in the NMDS graph suggest that they may be correlated. The strong phosphorus correlation vector is due to a high amount of phosphorus found in one sample from Site 343. The strong percentage of potassium correlation vector is due to the high amounts of potassium found in samples from Site 1 and 58. The percentage of calcium correlation vector is due to the high amounts of calcium found in two samples from Site 21.

Seasonal analysis

Of the 108 total PCR replicates conducted for the seasonal analysis, ten were removed from the dataset as failed replicates; this resulted in the complete removal of two samples (one from Site 58 and one from Site 333). The mean number of Acari and Collembola OTUs did not show any trends at the ordinal level between sampling periods (Figure 3.8b,c). In terms of community composition, however, there were significant differences between June and August communities (permANOVA R^2=0.07, p=0.003; Table 3.7), while no significant differences were found between sites within a sampling period (permANOVA R^2=0.36, p=0.08; Table 3.7). An NMDS graph supported this, as samples clustered by site more than they did by month (Figure 3.9). For example, all samples from Site 57 clustered together, regardless of which month they were sampled in.

Discussion

Soil mesofaunal diversity in an agricultural landscape

Although agricultural ecosystems are considered highly disturbed environments, they can still be home to a large diversity of soil organisms, particularly in field edges or other semi-natural habitats (Bennewicz & Barczak, 2020; Morris et al., 2010). Gathering baseline information on these soil organisms is essential because it can provide insight into how threats to soil biodiversity,

such as agricultural intensification or climate change, can be measured and mitigated (FAO, 2020). Since surveys of soil mesofauna are still relatively rare (Arroyo et al., 2013; George et al., 2017; Meehan et al., 2019; Porco et al., 2013; Rueda-Ramírez et al., 2022), surveys such as this study can help to improve conservation efforts (Smith et al., 2006).

For the agricultural habitats sampled in this study, an incredible amount of diversity was detected. Metabarcoding enabled the detection of over 1000 OTUs (as a proxy for species) belonging to a wide range of taxa, including both the target mesofauna, as well as other soil dwelling fauna, such as ants and worms. Similarly high levels of soil diversity were found using metabarcoding methods in Arctic environments (Young & Hebert, 2022), tropical habitats (Basset et al., 2022), grassland habitats (Arribas et al., 2016), and forested habitats (Kirse et al., 2021). While it is difficult to compare numbers of OTUs, since studies focus on different taxonomic groups, a study analyzing tropical soil arthropods using metabarcoding methods detected close to 400 species of arthropods within their study (Basset et al., 2022). Within grassland habitats, almost 500 OTUs of mesofauna were detected molecularly (Arribas et al., 2016), and more than 400 species of arthropods were detected through metabarcoding within forested habitats (Kirse et al., 2021). Another metabarcoding study looking at arthropod diversity in the Arctic detected a mean of 47 OTUs per sample within their bulk specimen samples (Young & Hebert, 2022). In comparison to this study, an average of 26 OTUs were found per sample. While this could indicate lower diversity in disturbed agricultural environments, even compared to harsh arctic environments, differences in sample size, extraction efficiency, and selected habitats can also play a role. Many more studies from a variety of habitats are needed in order to draw more broad conclusions about the relative diversity of soil fauna in different regions and habitat types.

With respect to the mesofauna, many of these taxa (75%) could not be assigned to the species level, and 44% could not be assigned past family. This lack of taxonomic resolution is unsurprising considering the number of undescribed mesofaunal species within these ecosystems. For example, the majority of mite species in Canada have yet to be described (Beaulieu et al., 2019; Lumley et al., 2013) and only 13% of Collembola have been described within Canada (Turnbull & Stebaeva, 2019). A lack of species designation for an OTU can also stem from the incomplete BOLD reference library, particularly for mesofauna (Recuero et al., 2024). For example, it is estimated that only 10% of described Acari species have assigned BINs in BOLD's reference library (Beaulieu et al., 2019). There is also taxonomic bias in the species that are present in the database; for example, Collembola records in BOLD are skewed towards larger, surface-dwelling taxa, leaving the soil-dwelling taxa incomplete (Turnbull & Stebaeva, 2019).

These taxonomic issues are further exacerbated by the possibility of incongruence between operational taxonomic units (=OTUs, or BINs in the BOLD database) and true biological species. While COI has been used extensively to distinguish animal taxa species through metabarcoding (Hebert et al., 2003; Virgilio et al., 2010; Wilson et al., 2017), issues can arise when using COI for species delineation. For example, COI may not be able to distinguish between closely related species for some animal taxa (Carew & Hoffman, 2015; Elias et al., 2007; Kaila & Ståhls, 2006; Keck et al., 2022; Meier et al., 2006), or alternatively can over-split species (Lukic et al., 2021). While in some cases these may represent cryptic species that have not been identified (Beaulieau et al., 2019; Zhang et al., 2019), in others, it may cause an artificially high estimate of the number of species (Song et al., 2008). Disentangling these alternatives can be challenging without detailed taxonomic work. For example, the mite family Eupodidae contained the greatest number of OTUs in this study, yet none of these OTUs were assigned to species level. Based on the number of BINs

in BOLD, it is estimated that Eupodidae may be one of the most species-rich Acari families; however, taxonomic efforts are needed to document and describe this highly diverse group of species (Beaulieu et al., 2019). On the other hand, the molecular diversity of some families, such as Nanorchestidae, is possibly caused by high rates of genetic variation, thus overestimating the biological diversity of this family (Beaulieu et al., 2019). Similarly, high genetic variation amongst some groups of Collembola can result in the same species being assigned to different taxa (Potapov et al., 2020). Therefore, more taxonomic efforts are needed to update reference libraries for mesofauna, such as through increased efforts to survey and describe mesofaunal species, to assess concordance between BINs and biological species, and to morphologically identify BINs that are already present in BOLD.

Nonetheless, despite the current lack of species or even genus-level taxonomic designation, surveys such as this provide insight into mesofaunal communities within agroecosystems in Canada and can be useful resource for researchers who wish to survey mesofauna in agricultural landscapes within Canada, assess how they may be changing over time due to anthropogenic factors, and determine how best to conserve these soil communities (George et al., 2017). Some studies indicate that there may not be a need for species level identifications (Bacci et al., 2009; George et al., 2017; Herman & Heip, 1988). For example, Bacci et al. (2009) did an assessment of taxonomic sufficiency and found higher taxonomic levels were often ecologically informative, with family level being the best option to analyze benthic assemblages if species level identifications were not available. However, there is not much research on the efficiency of using higher taxonomic levels to understand mesofaunal communities. In the end, considering the current extinction crisis and the many threats to soil biodiversity, monitoring taxa at the species

level is the best way to track the effectiveness of conservation efforts; this will require both molecular tools and the training and support of taxonomists.

Natural capital and mesofaunal diversity

Maintaining natural capital in agricultural landscapes, such as hedgerows, treed or untreed field margins and drainage ditches, shelterbelts, etc., has been shown to benefit soil biodiversity (Bennewicz & Barczak, 2020). It has further been demonstrated that forested habitats contain more soil mesofaunal biodiversity than do unforested areas (Tomlin & Miller, 1987). Increased soil mesofaunal diversity has been attributed to increased soil moisture and organic matter found in forested habitats (Arroyo et al., 2013; George et al., 2017; Szigeti et al., 2022), as well as forests being more stable environments (Kuznetsova, 2006). Similarly, Ferguson (2001) found that increasing levels of vegetation from grassland to shrub habitats was advantageous for soil arthropod diversity, as it increased soil moisture and provided microclimates more suitable for these organisms. It was therefore surprising that there were no significant differences found in species richness between the unforested and forested sites sampled. This may have been because diversity patterns across all mesofauna were masked by patterns at finer taxonomic levels. However, for all the main groups of taxa, this was not observed. The three orders of mites collected in this study (Sarcoptiformes (primarily Oribatida), Trombidiformes (primarily Prostigmata) and Mesostigmata (the only representatives collected from Parasitiformes)) have widely varying life strategies; I thus expected them to respond differently to the habitats studied here. For example, oribatids are "k-selected" soil organisms; they reproduce relatively slowly, have low fecundity, and have longer lifespans (Behan-Pelletier, 2003). They are thus highly associated with stable, undisturbed environments (Maraun, 2000). However, the results of this study indicate that there were no differences in diversity between seemingly very different habitat types (un-treed field

ditches vs undisturbed forested streams). Conversely, Prostigmata and Mesostigmata are primarily made up of "r-selected" species, with high fecundity and shorter lifespans (Behan-Pelletier, 2003); because of this adaptability, it is less surprising that these groups had no differences in diversity between habitat types. Kokořová & Starý (2017) found that there was an increase in more tolerant and resilient oribatid species in disturbed ecosystems, which could be what is dominating the unforested sites that were right next to highly disturbed agricultural fields. A similar lack of association of any Collembola order with forested or unforested sites likely relates to the wide range of life histories within each order (Potapov et al., 2020).

Despite these observations, it is also possible that this pattern was simply due to high sample variability and insufficient sampling. There were only three sites sampled per habitat type, and significant variability in the number of OTUs detected was observed in the samples from each site. Since sample size is correlated with statistical power (Valdez et al., 2023), the limited number of samples taken for this study may have resulted in species richness patterns going undetected.

While no significant differences were found in mesofaunal species richness between habitat types, there were significant differences found between two of the no trees sites (relatively high diversity at Site 1, and very low diversity at Site 21). This may be due simply to random variation within the sites sampled; however, it is also of note that Site 21 had high levels of organic matter, while Site 1 had relatively low levels of organic matter. High organic matter content may contain higher levels of humic acid, which has been shown to hinder PCR success (Matheson et al., 2010), and may have led to a decrease in OTU detection for Site 21. Along with this, two samples from Site 21 were removed from the dataset due to PCR failures, which likely reduced the number of OTUs found for this site.

Natural capital and community composition

Although species diversity did not differ significantly, there were nonetheless significant differences in mesofaunal community composition, both between habitat types and between sites within a habitat type, as determined through the permANOVA. However, the ordination still showed considerable overlap in the communities, with the most distinct pattern being between the two forested habitats (large and small forests) and the two unforested habitats (scattered trees and no trees). This discrepancy could be because the permANOVA considers both centroid and spread of samples (Anderson, 2017), thus reflecting changes in variability within groups, rather than the overall community similarity.

As with the species diversity, high variability of the community data characterized these samples. More heterogeneous habitats have been shown to contain more heterogeneous groups of mesofauna (Kuznetsova, 2006). While it was expected that the forested sites would be highly heterogeneous at the microhabitat level, unforested sites appeared to be more homogeneous, having similar widths and vegetation. However, there was little overlap in species presence between sites in any of the habitats, and no distinct clustering of the communities. This indicates that the scale at which the communities are being structured is more fine-grained than the large-scale habitats defined here. Differences in community composition between samples could be attributed to the small-scale habitat features of each site, such as whether samples were collected underneath trees versus shrubs (Heydari et al., 2020), or in differences in the underlying soil layer (George et al., 2017). Geographical distances can also have an impact on mesofaunal species composition (Andújar et al., 2022; Arribas et al., 2020; Lehmitz et al., 2012). According to Gao et al. (2014), as spatial distance increases, the similarity of soil mesofauna communities decreases because of differences in environmental factors, such as changes in organic matter. This was

supported by the results in this study, as soil mesofaunal communities varied greatly amongst samples within sites, with samples that were taken further away containing different soil mesofaunal communities. These differences can be attributed to how changing environmental factors between samples within sites can impact microhabitats, thereby influencing the mesofaunal community composition of each site. Variability in soil properties between samples within sites, especially between unforested sites, could also explain large differences in the mesofaunal communities between sites. For example, an increase in organic matter or soil moisture for a microhabitat can increase the diversity and alter the soil mesofaunal community composition (Koudji et al., 2023; Kardol et al., 2011). Although high variability was detected between sites and samples, other factors may have confounded the ability to detect patterns in mesofaunal community composition.

One factor that may have weakened the ability to detect differences in community composition is the lack of abundance data for the collected taxa. Using read counts has not yet been shown to provide conclusive measures of abundance, due to issues arising from primer bias, as the primers used in this study may be amplifying certain taxonomic groups more than others (Elbrecht & Leese, 2015). Due to time constraints, it was not possible to identify and count all of the collected specimens. Identifying and counting these specimens would have revealed details about richness and evenness at each site, as well as which taxa were dominant. Using presence-absence data treats all taxa equally, failing to consider whether they are abundant and ecologically associated with a habitat or transient (Ashcroft et al., 2017; Kirichenko-Babko et al., 2021). As well, it is possible to pick up gut contents of predatory mite taxa, potentially resulting in higher number of OTUs for mite prey, such as nematodes or Collembola (Eitzinger et al., 2013). Another issue could be that DNA from dead organisms or their remains could have been

picked up in the soil (Deiner et al., 2017; van der Heyde et al., 2022) that are not actively living in that habitat. These factors could have impacted the patterns of community composition found, resulting in more diffuse patterns within and between sites.

Despite this underlying pattern of high sample to sample, and site to site variability, there was nonetheless an indication that forested sites (regardless of size) contained different mesofaunal communities than did the unforested sites (whether with or without scattered trees). These communities are strongly associated with increasing tree cover, unsurprisingly associated with increased moisture, and to a lesser extent, organic matter. Studies have shown that even small patches of forests within agricultural landscapes can increase biodiversity (Varela et al., 2018), along with allowing for different communities to form due to the formation of microclimates (Decocq et al., 2016). The results of this study support this, as although there were between site community differences found within habitat types, the communities found in both small and large forest habitat types clustered together. This indicates that small forests appear to be maintaining a forest-like community, which implies the value that forested sites have on an agricultural landscape, as even small patches of forests can contain their own mesofaunal communities. For example, patches of forest within an agricultural environment can allow for both forest specialist species and habitat generalist species to flourish (Decocq et al., 2016). Hendrickx et al. (2007) found that natural forested habitats within an agricultural matrix allows for a community of both field inhabiting spider species, along with species that are more adapted to natural habitats to occupy that area. Vanbergen et al. (2007) also found that increasing the landscape heterogeneity, regardless of patch size, within an agricultural area can impact taxa differently, with forested patches benefitting Collembola populations more than lumbricid worm populations. Although connectedness of sites were not addressed within this study, they could

also be a factor impacting mesofaunal communities. For instance, Rantalainen et al. (2005) found that corridors connecting habitat patches together can result in the colonisation of soil organisms into previously uninhabited areas. This indicates the significance of having multiple patches of natural area of varying sizes within an agricultural landscape, as this could promote a more diverse mesofaunal community.

Soil physicochemical properties also impact the soil mesofaunal community (Crotty et al., 2023; Wu & Wang, 2019). Gravimetric moisture and CEC can be influenced by tree cover, as more forested ditches tend to have wetter soils with higher levels organic matter content and CEC (Blazka & Fischer, 2014), whereas field ditches undergoing intense land-use and being exposed to direct sunlight without any tree cover tend to have drier soils (Czigány et al., 2023). Therefore, it was surprising that there were no significant differences in these soil properties between habitat types. Reasons for this could be that sites have variable amounts of organic matter and CEC, indicated by the weak correlation vector present in the NMDS graph. Higher than expected amounts of organic matter present in the unforested sites could be due to fertilizers used in adjacent fields or inputs of nutrients from field runoff (Manninen et al., 2018), thereby resulting in some unforested sites clustering with forested sites. For example, Wissuwa et al. (2013) found that soil moisture and organic matter content determined the oribatid mite species that were present within agricultural fields. A negative correlation was found between magnesium and organic matter, which may be attributed to high amounts of magnesium causing soil aggregates to form, which may block water from entering the soil, thereby reducing the organic matter present (Dontsova & Norton, 2002), which could be a reason why the unforested sites with less organic matter content and increased magnesium tend to cluster. Although Mantel's test has been critiqued for lacking statistical power, resulting in the potential for Type II errors to occur (false negatives) (Guillot &

Rousset, 2013; Somers & Jackson, 2022), the conclusions pertaining to the correlations of soil parameters on mesofaunal community composition observed here would not have changed if a more powerful test was used instead. Therefore, these soil properties likely played a role in the clustering pattern of forested versus unforested sites and how their communities are shaped.

Mesofaunal diversity and seasonal turnover

Seasonal differences have been shown to have an impact on the diversity and community composition of soil fauna species (Bardgett & Cook, 1998; Wu & Wang, 2019). Although species richness did not seem to change across the season, significant differences in community composition were found between the June and August samples; however, these differences were not evident in the NMDS graph. For instance, sites tended to cluster together, regardless of which month they were sampled in, despite large amounts of variation in community composition within sites. This indicates that inter-sample variability is influencing the mesofaunal community more than inter-seasonal variability. Contrastingly, Wu & Wang (2019) found that the soil mesofaunal community was impacted by seasonal changes in temperature and precipitation. However, soils tend to be more stable habitats throughout the growing season and may not be as seasonally driven as other habitats, such as bushes or plants (Smith et al., 2015). Since mesofaunal communities do not appear to vary greatly throughout the growing season, studies that aim to survey these communities can do so over a longer period of time, without potentially compromising the data.

Methodological issues

Although DNA metabarcoding is an effective tool in its ability to capture the diversity of the soil community within different types of samples (Frézal et al., 2008; Oliverio et al., 2018; Young & Hebert, 2022), the method has advantages and disadvantages. Metabarcoding bulk

specimens was an effective tool in this study, as it was able to capture a large portion of the mesofaunal community, recovering more than a thousand unique OTUs for this study. However, the results obtained from metabarcoding are sensitive to the molecular methods used, particularly with respect to DNA extraction and PCR amplification. In this study, the Berlese funnel extraction method resulted in some soil sediment from the samples being extracted along with the specimens. This may have caused issues with PCR amplification, as some soils, particularly from forested sites containing high organic matter content, may also contain high levels of humic acid, which has been shown to hinder PCR success (Matheson et al., 2010). The DNA extraction protocol used works effectively on bulk specimens, but it did not contain steps to remove PCR inhibiting factors common in bulk soil samples. This may explain why species richness between Site 21 and Site 1 varied significantly, as the high amount of organic matter present at Site 21 could have resulted in PCR inhibition and thus, fewer OTUs detected. Due to time constraints, it was not feasible to pick out all specimens individually prior to DNA extraction, and soil-specific DNA extraction kits are expensive to use for a large number of samples. However, future studies should consider this potential source of error and plan accordingly, to either ensure specimen samples are clean, or to consider the presence of potential inhibitors when planning extraction methods. While DNA metabarcoding is an expensive tool, conducting morphological identifications are also costly in terms of the time and expertise needed to identify individual specimens within a large number of samples. In terms of data, morphological analyses provide abundance data, which may provide additional insight into community composition and ecological processes; however, to obtain species-level (or OTU-level) detail across all taxonomic groups, a metabarcoding approach is much more efficient. Therefore, the choice of method depends on the goals of the project and the funds, expertise, and personnel hours available.

Conclusion

This study surveyed and examined the soil mesofaunal community and differences in species richness and community composition in various types of natural capital within an agricultural matrix across the growing season. Gathering this baseline information helps fill a knowledge gap in our understanding of mesofauna in agroecosystems within Canada. This study also highlighted a need for more robust taxonomic efforts to be made in order to update reference libraries for mesofauna for future metabarcoding work. Differences in species richness were not detected between the four habitat types or between most sites within habitats. However, community composition differences were found to be significant, which may be attributed to site differences, such as differences in vegetation or soil properties found at each site. Since the soil mesofaunal community is operating at a much smaller scale, these features can create microhabitats for soil mesofauna to thrive in, with the microhabitat features influencing the soil fauna, rather than the macrohabitat features. Sites did not vary much across the sampling season, which could be explained by microhabitats providing stable ecological conditions for soil mesofaunal communities. Significant differences were found between sites within a habitat type for all soil parameters tested, indicating that soil properties ultimately impact the soil mesofaunal community composition at these sites. This study demonstrates the importance of retaining natural capital in order to preserve soil biodiversity, seeing as even small agricultural ditches can be home to a vast number of soil organisms. This study also shows that even small forested habitat patches can contain similar mesofaunal communities to those of larger forested patches, indicating that it is beneficial for biodiversity to preserve these small forest patches within an agricultural matrix.

Tables

Table 3.1. Location and soil type of each sampling site.

Habitat Type	Site	Latitude	Longitude	Soil Type
No Trees	1	45.2615°	-75.1672°	Clay loam
	21	45.3034°	-74.9690°	Clay
	58	45.4140°	-75.2332°	Sandy loam
Scattered Trees	7	45.2656°	-75.1700°	Clay loam
	53	45.4395°	-75.3285°	Clay loam
	57	45.4060°	-75.3990°	Silty clay
Small Forest	8	45.2725°	-75.1648°	Loam
	115	45.4582°	-75.1755°	Loam
	334	45.4207°	-75.4359°	Clay loam
Large Forest	333	45.3387°	-75.1658°	Silt loam
	343	45.4054°	-75.1017°	Silt loam
	347	45.4142°	-75.3299°	Loam

Table 3.2. Summary of OTUs and reads found for each sample type for all OTUs and classified mesofaunal OTUs for each habitat type.

Habitat	Samples/ PCR replicates	All Taxa						Mesofauna	
		Total OTUs	Total reads	Identified OTUs	Identified reads	Unidentified OTUs	Unidentified reads	Total OTUs	Total reads
All site samples	64/192	844	634195	585	606496	259	27699	390	224085
No Trees	14/34	220	135490	162	134426	58	1064	120	41496
Scattered Trees	17/49	276	174828	236	172365	40	2463	175	40212
Small Forest	17/47	363	179843	241	174264	122	5579	167	77337
Large Forest	16/46	294	144034	212	125441	82	18593	152	65040
Negatives	-/16	55	191	48	154	7	37	31	92

Table 3.3. Number of OTUs assigned to each animal family for samples collected in the habitat and seasonal analysis datasets.

Phylum	Class	Order	Family	# of OTUs
Annelida	Clitellata	Crassiclitellata	Lumbricidae	11
		Enchytraeida	Enchytraeidae	32
		Haplotaxida		1
Arthropoda	Arachnida	Araneae	Araneidae	1
			Linyphiidae	4
			Lycosidae	1
			Theridiidae	1
		Mesostigmata	Ameroseiidae	1
			Ascidae	5
			Digamasellidae	2
			Laelapidae	7
			Macrochelidae	1
			Pachylaelapidae	2
			Parasitidae	14
			Parholaspididae	3
			Phytoseiidae	2
			Rhodacaridae	6
			Veigaiidae	8
		Sarcoptiformes	Acaridae	2
			Achipteriidae	2
			Astegistidae	4
			Brachychthoniidae	6
			Ceratozetidae	9
			Chamobatidae	3
			Eniochthoniidae	2
			Eremobelbidae	1
			Euphthiracaridae	1
			Euzetidae	1
			Galumnidae	11
			Gustaviidae	3
			Hypochthoniidae	4
			Mochlozetidae	1
			Nanhermanniidae	1
			Nanorchestidae	26
			Nothridae	8
			Oppiidae	22
			Oribatulidae	3
			Parakalummidae	3

		Phenopelopidae	1
		Phthiracaridae	1
		Punctoribatidae	1
		Scheloribatidae	22
		Suctobelbidae	5
		Tectocepheidae	11
		Tegoribatidae	1
	Trombidiformes	Ereynetidae	2
		Eriophyidae	1
		Erythraeidae	1
		Eupodidae	71
		Pygmephoridae	5
		Rhagidiidae	17
		Scutacaridae	4
		Tarsonemidae	5
		Trombidiidae	3
		Tydeidae	30
Chilopoda	Geophilomorpha	Geophilidae	11
		Schendylidae	3
	Lithobiomorpha	Henicopidae	3
		Lithobiidae	8
Collembola	Entomobryomorpha	Entomobryidae	17
		Isotomidae	49
		Tomoceridae	1
	Neelipleona	Neelidae	19
	Poduromorpha	Hypogastruridae	3
		Neanuridae	7
		Onychiuridae	13
		Tullbergiidae	20
	Symphypleona	Bourletiellidae	5
		Dicyrtomidae	1
		Katiannidae	11
		Sminthuridae	1
		Sminthurididae	8
Diplopoda	Chordeumatida		1
	Julida	Blaniulidae	2
		Julidae	18
Diplura	Rhabdura	Campodeidae	7
Insecta	Coleoptera	Cantharidae	1
		Carabidae	14
		Cryptophagidae	2
		Curculionidae	4
		Elateridae	6
		Ptiliidae	1

			Ptinidae	1
			Staphylinidae	13
		Diptera	Cecidomyiidae	1
			Chironomidae	1
			Phoridae	2
		Hemiptera	Aphididae	4
			Pseudococcidae	1
			Rhyparochromidae	2
		Hymenoptera	Formicidae	27
		Psocodea	Psyllipsocidae	1
		Thysanoptera	Phlaeothripidae	1
			Thripidae	3
	Malacostraca	Isopoda	Trichoniscidae	2
	Symphyla		Scolopendrellidae	10
Mollusca	Gastropoda	Stylommatophora	Arionidae	1
Nematoda	Chromadorea	Rhabditida	Diplogastridae	2
			Rhabditidae	2
		Tylenchida	Aphelenchidae	2
			Parasitylenchidae	1
	Enoplea	Dorylaimida		1
Rotifera	Bdelloidea	Adinetida	Adinetidae	2
		Philodinida	Habrotrochidae	3
			Philodinidae	6

Table 3.4. Nested ANOVA results for mean mesofaunal OTUs for each habitat type and sites within a habitat type.

	df	F	P
Between Habitat Types	3	0.585	0.64
Between Sites	11	2.09	0.03*

*Significant p-values

Table 3.5. Comparison of tree cover and soil physicochemical parameters for each habitat type. Tests of significance are done using a nested ANOVA test.

	Habitat type		Sites within habitat types	
Environmental parameter	**Test Statistic**	**P-value**	**Test Statistic**	**P-value**
Tree Cover	F = 30.4	<0.001*	F = 38.4	<0.001*
Organic Matter	F = 0.3	0.81	F = 11.2	<0.001*
Gravimetric Moisture	F = 0.3	0.81	F = 4.2	0.007*
Cation Exchange Capacity (CEC)	F = 1.5	0.30	F = 17.6	<0.001*
% of K	F = 5.5	0.02*	F = 8.1	0.002*
% of Mg	F = 1.8	0.22	F = 4.2	0.003*
% of Ca	F = 2.2	0.17	F = 6.2	<0.001*
Phosphorus	F = 0.6	0.65	F = 3.4	0.001*

*Significant p-values

Table 3.6. Mantel's test results to determine correlation between community similarity and distance, tree cover, and soil physicochemical properties.

Property	**Test Statistic**	**P-value**
Distance	r = 0.16	<0.001*
Tree Cover	r = 0.08	0.001*
Organic Matter	r = 0.17	0.002*
Gravimetric Moisture	r = 0.07	0.08
Cation Exchange Capacity (CEC)	r = 0.14	0.008*
% of K	r = -0.003	0.50
% of Mg	r = 0.14	0.004*
% of Ca	r = 0.14	0.005*
Phosphorus	r = 0.06	0.13

*Significant p-values

Table 3.7. Repeated measures nested permANOVA results for each month and sites within a month.

	Months				Sites within a Month			
	Df	**F**	**R^2**	**P-value**	**Df**	**F**	**R^2**	**P-value**
All Months	2	1.2	0.06	0.008*	9	1.5	0.36	0.08
June-July	1	1.1	0.04	0.11	6	1.7	0.41	0.25
July-August	1	1.0	0.04	0.17	6	1.5	0.35	0.32
June-August	1	1.6	0.07	0.003*	6	1.4	0.35	0.07

*Significant p-values

Figures

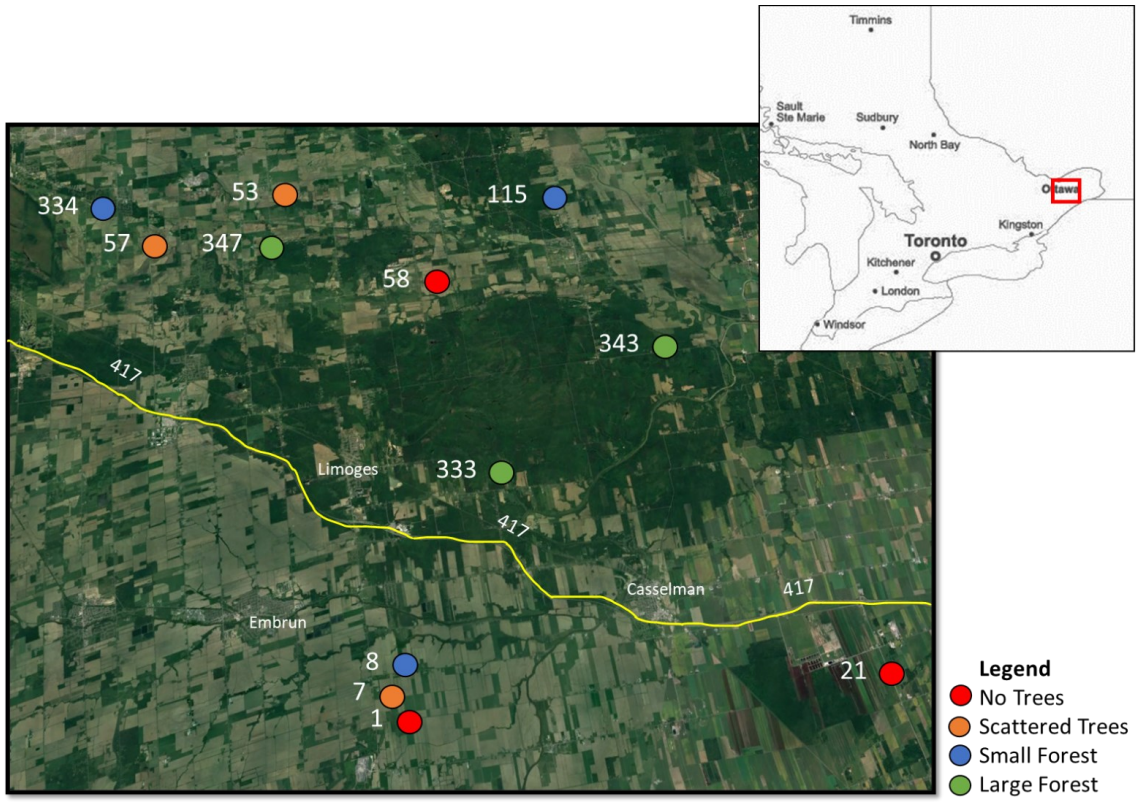


Figure 3.1. Map showing the location of the twelve sites that were sampled in this study.

Figure 3.2. Images depicting examples of each habitat type chosen for this study.

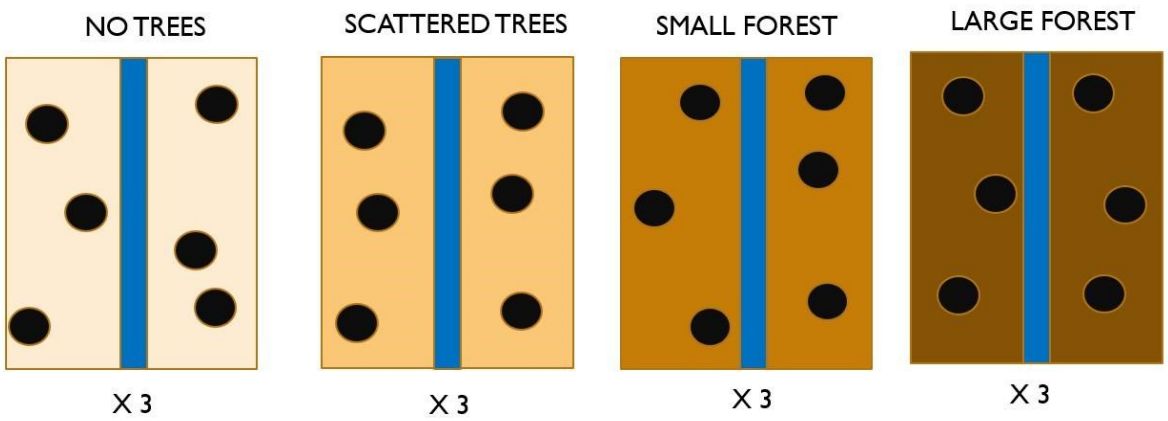


Figure 3.3. Experimental design schematic for 12 sample sites.

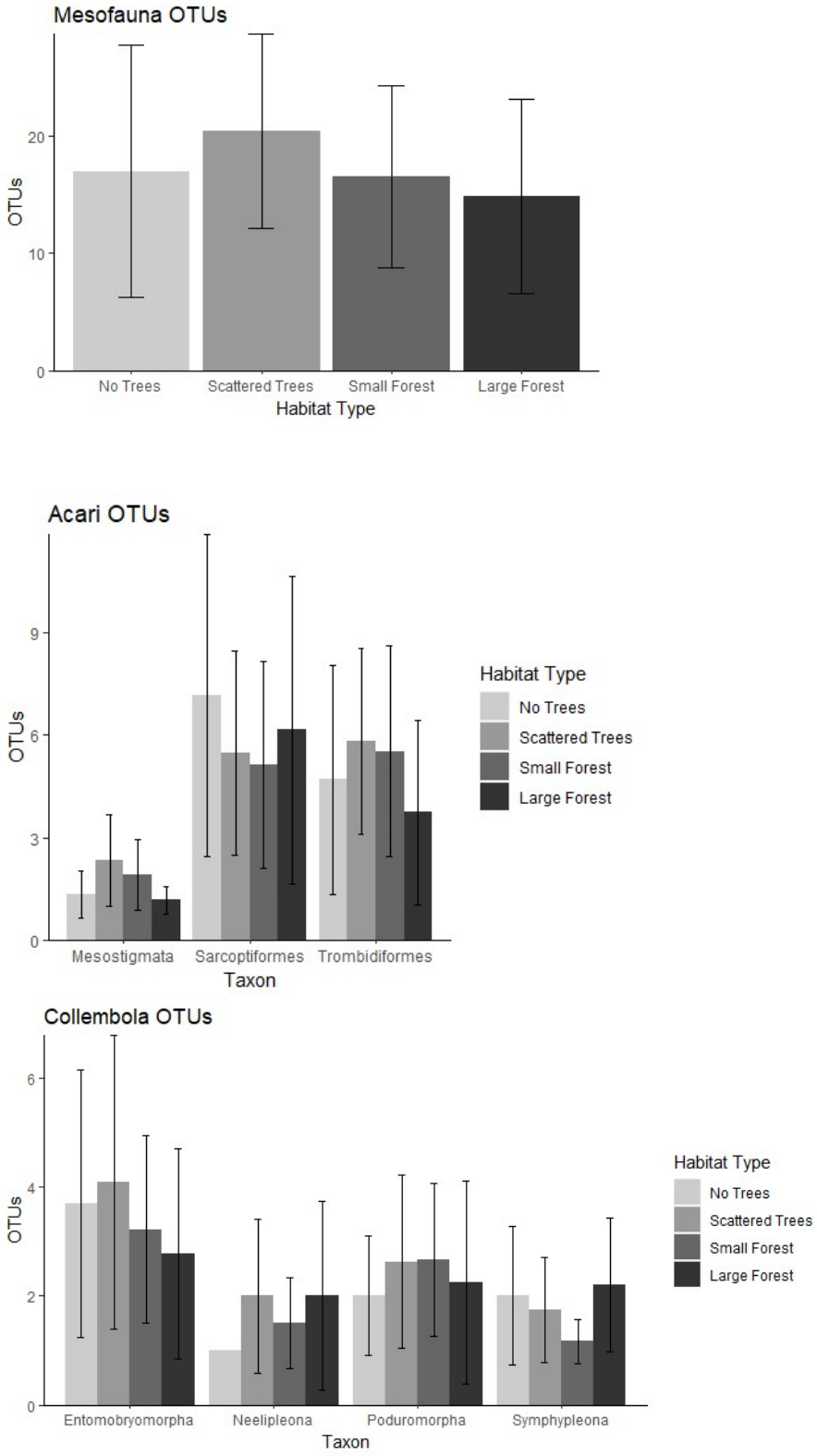


Figure 3.4. Mean (± SE) number of mesofauna OTUs for each habitat type; a. All mesofauna taxa combined; b. Acari orders; c. Collembola orders.

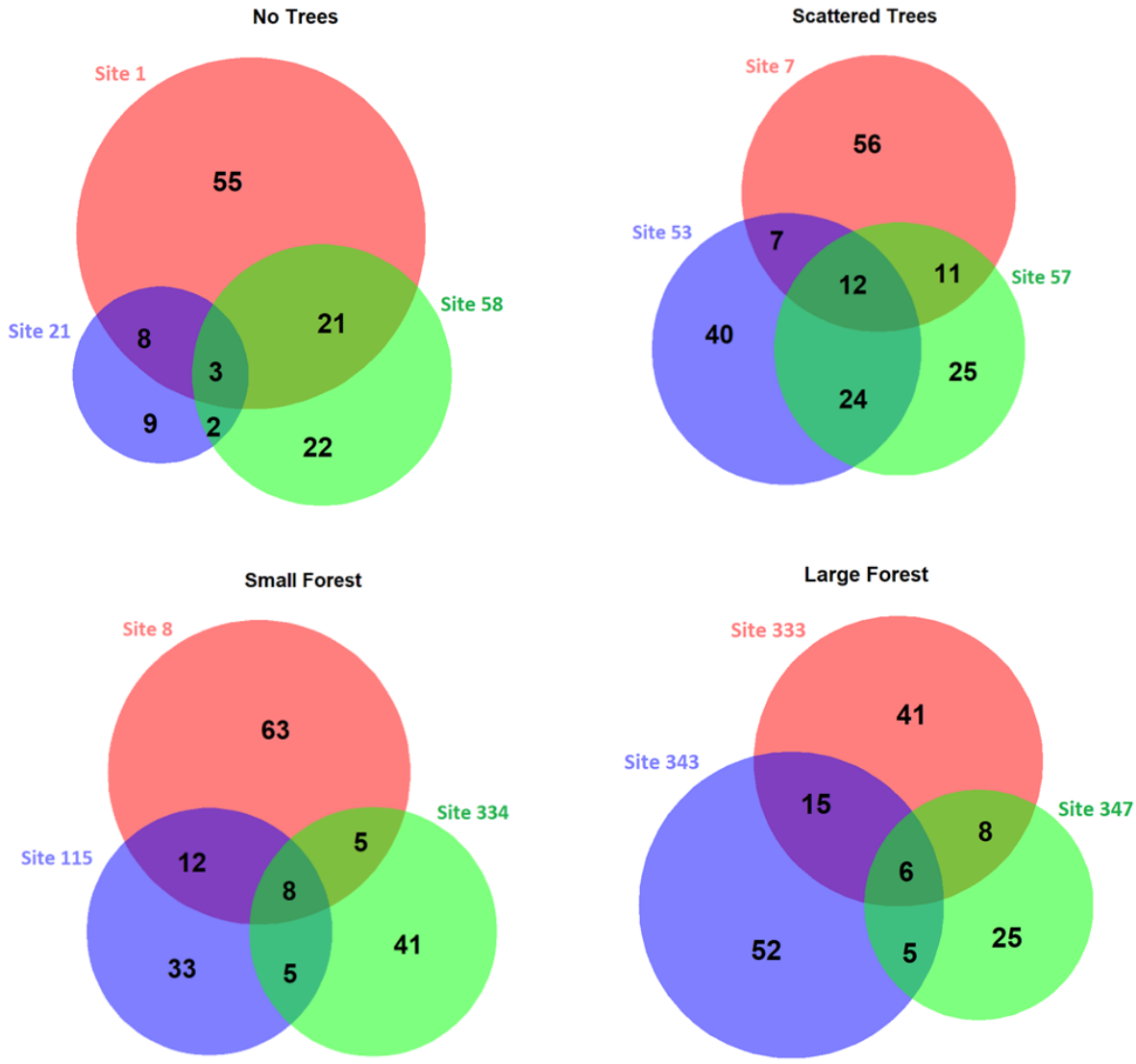


Figure 3.5. Overlap of mesofauna OTUs detected from the three sites within each habitat type.

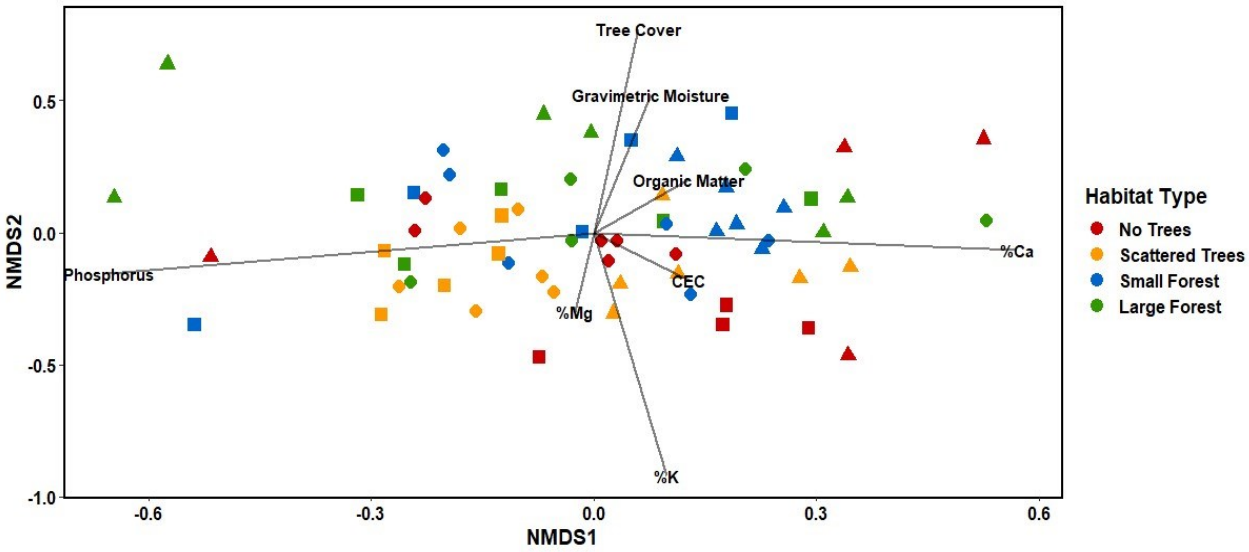


Figure 3.6. Mesofaunal community composition for each habitat type with soil parameters modelled using non-metric multidimensional scaling (NMDS) (stress value=0.19). Red = No trees: circles = Site 1, triangles = Site 21, squares = Site 58. Yellow = Scattered trees: circles = Site 53, triangles = Site 7, and squares = Site 57. Blue = Small forest: circles = Site 115, triangles = Site 8, and squares = Site 334. Green = Large forest: circles = Site 333, triangles = Site 343, and squares = Site 347.

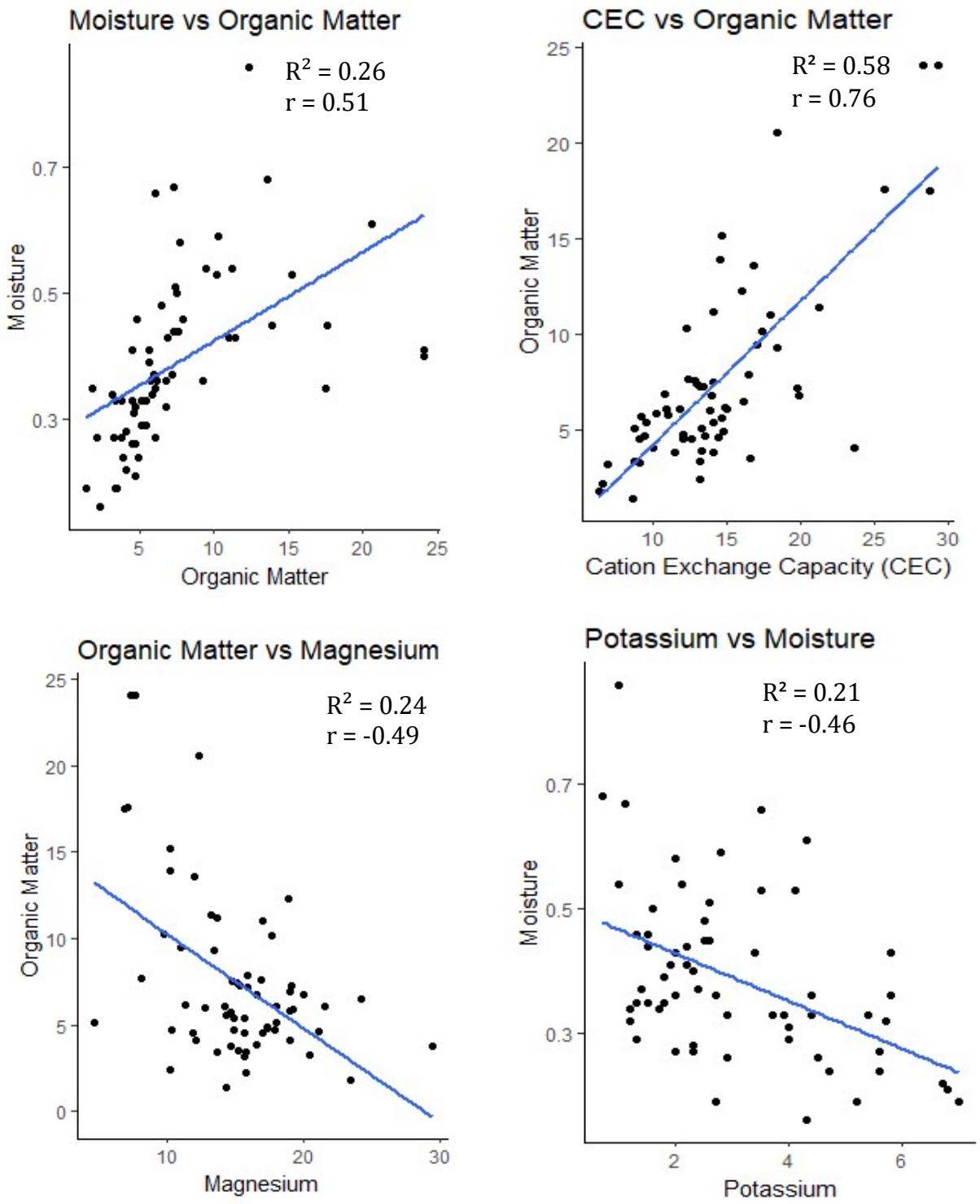


Figure 3.7. Correlation plots showcasing significant correlations between soil physicochemical parameters.

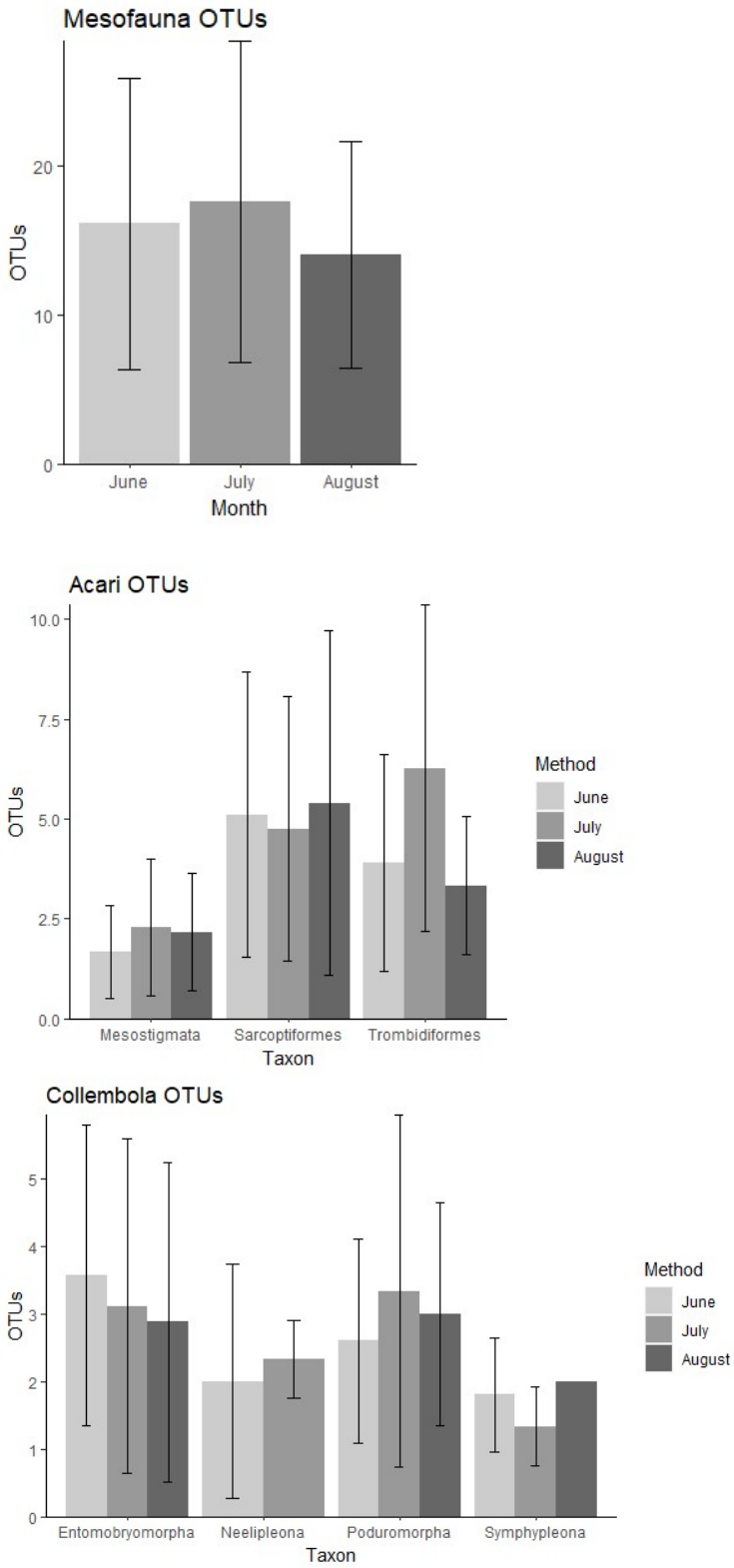


Figure 3.8. Mean (± SE) number of mesofauna OTUs for each month sampled; a. All mesofauna taxa combined; b. Acari orders; c. Collembola orders.

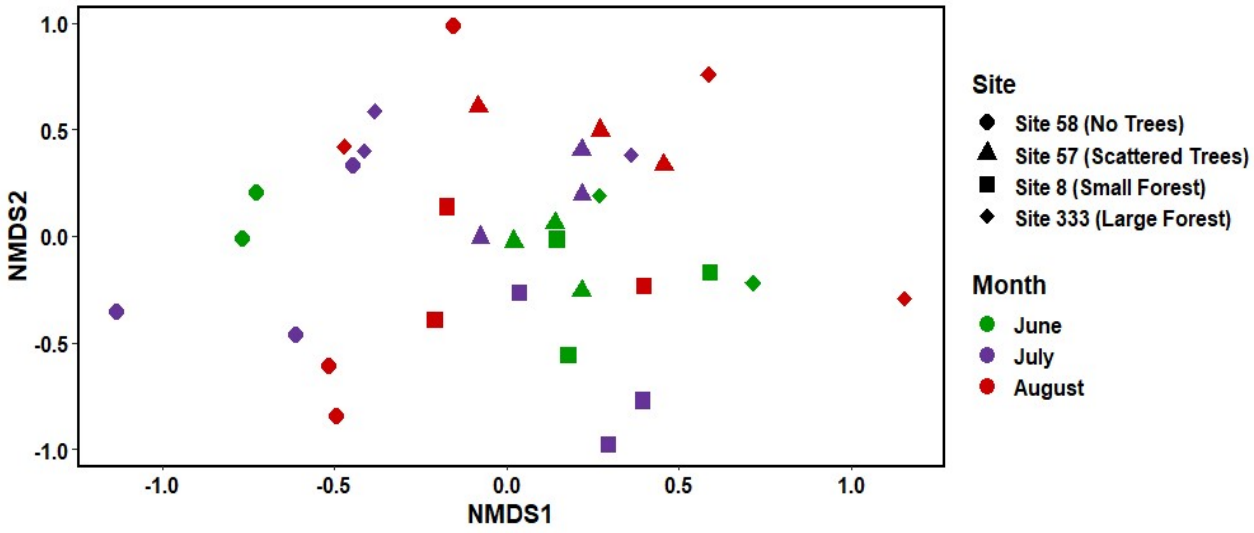


Figure 3.9. Community composition for the seasonal samples modelled using non-metric multidimensional scaling (NMDS) (stress value=0.14).

Chapter 4. General Conclusions

This thesis demonstrated the usefulness of DNA metabarcoding as a tool to analyze species richness and community composition of the soil mesofauna within agroecosystems. I did this by first performing DNA metabarcoding on two paired sample types (bulk soil samples versus bulk specimen samples) in order to compare the differences in soil mesofaunal community composition and species richness. I also compared the morphological identifications with the bulk specimen samples to assess whether both methods were able to capture similar mesofaunal species. I then used DNA metabarcoding to determine if there were differences in species richness and community composition for the soil mesofaunal community for agricultural sites with varying levels of natural capital across the growing season. Finally, I compared the soil physicochemical properties of each of these sites to determine if these differences had an impact on the soil mesofaunal community found at these sites.

In Chapter 2, I found that many more taxa were detected in the SOIL samples, which contained all the environmental DNA in the sample, compared to the BERLESE samples, which specifically targeted the soil fauna. With respect to assessing the soil mesofaunal community, however, the BERLESE samples were more efficient at capturing mesofaunal taxa than were the SOIL samples. When comparing community composition between the two samples, I found that there was much more overlap between the two molecular methods when the analysis was restricted to the soil mesofaunal community, with similar sites having similar communities, regardless of which method was used. I also compared morphological identifications at the family level with the bulk specimen sequencing results and found differences in mesofaunal taxa detected by the two methods. These differences in diversity and community composition between molecular and morphological methods could be attributed to the DNA extraction

method used for the BERLESE samples, which may have resulted in the loss of some smaller voucher specimens. As well, some species may not have been identifiable morphologically due to their life cycle stage (Dritsoulas & Duncan, 2020). The primers used in this study may have also failed to amplify certain taxa as a result of primer mismatch. Therefore, there are strengths and weaknesses in using molecular or morphological methods to identify the soil community. Studies that use a larger number of samples may favour metabarcoding due to time constraints, whereas studies may chose to use morphological identifications for a smaller number of samples, as this method would be beneficial for obtaining abundance data.

In Chapter 3, baseline taxonomic information on the soil mesofaunal community within agroecosystems was collected, providing a resource for future researchers to use when assessing changes in these communities (Table 3.3). This chapter aimed to explore mesofaunal species richness and community composition in different habitat patches within an agricultural matrix, ranging from small, untreed ditch margins (no trees and scattered trees) to varying patches of forest (small forest and large forest). The species richness for the soil mesofaunal community of the four different habitat types tested were found to not be significantly different from each other. The mesofaunal community composition did, however, vary between habitat types, which may due to differences in above-ground vegetation and other environmental site characteristics (Koehler & Born, 1989; Korboulewsky et al., 2016). Soil physicochemical properties, such as organic matter, moisture, and CEC may have also played a role in shaping mesofaunal communities. As well, this study demonstrated that small forested habitats contain similar mesofaunal communities as do larger forested habitats, indicating conservation of even small habitat patches can ultimately preserve soil biodiversity. Finally, this study did not find strong

seasonal effects on the mesofaunal community, suggesting that habitat or site has a stronger role in determining what taxa are present.

This thesis highlighted key advantages and disadvantages of using DNA metabarcoding on both bulk soil samples and bulk specimen samples in order to assess species richness and community composition. However, this study was limited to only exploring the soil mesofaunal community and does not examine these differences in the soil community as a whole. As well, the bulk soil method used in this study covered a wide range of non-animal taxa (e.g. bacteria and fungi), which may have taken up sequencing depth for target animal taxa. Continued experimentation with different primer sets to target mesofaunal groups using this molecular method is needed in order to ensure accuracy in interpretation of the results. The molecular results also contained an unexpectedly large portion of insect OTUs that were only identifiable to the order level, which could be a result of spurious sequences leading to an overestimation of diversity. Another pitfall I encountered with using DNA metabarcoding was the lack of taxonomic resolution for mesofaunal species due to incomplete reference libraries. There are also potential taxonomic issues involved with using COI as a species delineator, which may have resulted in overestimation of the true biological diversity present (Song et al., 2008). More taxonomic efforts are needed to contribute and verify mesofauna sequences contained within the Barcode of Life Database.

Along with taxonomic issues, methodological problems arose between the molecular methods tested. The bulk soil samples consisted of a time consuming and expensive DNA extraction process, so it is not feasible for a large number of samples. For the bulk specimen samples, specimens may have been lost during both the Berlese extraction and the DNA extraction steps, potentially resulting in less mesofaunal diversity being detected in those

samples. As well, the presence of PCR inhibitors, such as humic acid, contained within the soil samples from the four habitat types in Chapter 3, may have led to the PCR amplification step to work less efficiently, thereby resulting in DNA not being detected within these samples (Matheson et al., 2010). Despite these limitations, this thesis was able to use DNA metabarcoding to contribute to our knowledge of soil mesofauna species richness and community composition in a variety of habitats within an agricultural matrix. As well, this thesis bridges a knowledge gap that exists for identifying the soil mesofaunal species contained within agroecosystems in Eastern Ontario.

References

Akhila, A., & Entoori, K. (2022). Role of earthworms in soil fertility and its impact on agriculture: A review. *International Journal of Fauna and Biological Studies*, *9*(3), 55-63. https://doi.org/10.22271/23940522.2022.v9.i3a.907.

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*(1), 134-147. https://doi.org/10.1111/2041-210X.12849.

Anderson, M. J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA). *Wiley StatsRef: Statistics Reference Online*, 1-15. https://doi.org/10.1002/9781118445112.stat07841.

Andújar, C., Arribas, P., López, H., Arjona, Y., Pérez-Delgado, A., Oromí, P., Vogler, A. P., & Emerson, B. C. (2022). Community assembly and metaphylogeography of soil biodiversity: Insights from haplotype-level community DNA metabarcoding within an oceanic island. *Molecular Ecology*, *31*(15), 4078. https://doi.org/10.1111/mec.16560.

Arribas, P., Andújar, C., Hopkins, K., Shepherd, M., & Vogler, A. P. (2016). Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. *Methods in Ecology and Evolution*, *7*(9), 1071-1081. https://doi.org/10.1111/2041-210X.12557.

Arribas, P., Andujar, C., Salces-Castellano, A., Emerson, B. C., & Vogler, A. P. (2020). The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding. *Molecular Ecology*, *30*(1), 48-61. https://doi.org/10.1111/mec.15591.

Arroyo, J., Keith, A. M., Schmidt, O., & Bolger, T. (2013). Mite abundance and richness in an Irish survey of soil biodiversity with comments on some newly recorded species. *Irish Naturalists' Journal*, 19-27. https://www.jstor.org/stable/24394126.

Ashcroft, M. B., King, D. H., Raymond, B., Turnbull, J. D., Wasley, J., & Robinson, S. A. (2017). Moving beyond presence and absence when examining changes in species distributions. *Global Change Biology*, *23*(8), 2929-2940. https://doi.org/10.1111/gcb.13628.

Bacci, T., Trabucco, B., Marzialetti, S., Marusso, V., Lomiri, S., Vani, D., & Virno Lamberti, C. (2009). Taxonomic sufficiency in two case studies: where does it work better? *Marine Ecology*, *30*, 13-19. https://doi.org/10.1111/j.1439-0485.2009.00324.x.

Bardgett, R. D., & Cook, R. (1998). Functional aspects of soil animal diversity in agricultural grasslands. *Applied Soil Ecology*, *10*(3), 263-276. https://doi.org/10.1016/S0929-1393(98)00125-5.

Barrios, E. (2007). Soil biota, ecosystem services and land productivity. *Ecological Economics*, *64*(2), 269-285. https://doi.org/10.1016/j.ecolecon.2007.03.004.

Basset, Y., Hajibabaei, M., Wright, M. T., Castillo, A. M., Donoso, D. A., Segar, S. T., Souto-Vilarós, D., Soliman, D. Y., Roslin, T., Smith, M. A., Lamarre, G. P., De León, L. F., Decaëns, T., Palacios-Vargas, J. G., Castaño-Meneses, G., Scheffrahn, R. H., Rivera, M., Perez, F., Bobadilla, R., … Barrios, H. (2022). Comparison of traditional and DNA metabarcoding samples for monitoring tropical soil arthropods (Formicidae, Collembola and Isoptera). *Scientific Reports*, *12*(1), 10762. https://doi.org/10.1038/s41598-022-14915-2.

Baumann, J. (2021). Patterns of intraspecific morphological variability in soil mites reflect their dispersal ability. *Experimental and Applied Acarology*, *83*(2), 241-255. https://doi.org/10.1007/s10493-020-00587-y.

Beaulieu, F., Knee, W., Nowell, V., Schwarzfeld, M., Lindo, Z., Behan-Pelletier, V. M., Lumley, L., Young, M.R., Smith, I., Proctor, H.C., & Mironov, S.V. (2019). Acari of Canada. *ZooKeys*, (819), 77. https://doi.org/10.3897/zookeys.819.28307.

Behan-Pelletier, V. M. (1999). Oribatid mite biodiversity in agroecosystems: role for bioindication. *Agriculture, Ecosystems & Environment*, *74*(1-3), 411-423. https://doi.org/10.1016/S0167-8809(99)00046-8.

Behan-Pelletier, V. M. (2003). Acari and Collembola biodiversity in Canadian agricultural soils [Special Issue]. *Canadian Journal of Soil Science*, *83*, 279-288. https://doi.org/10.4141/S01-063.

Bellinger, P. F., Christiansen, K. A., & Janssens, F. (2024). Checklist of the Collembola of the World. http://www.collembola.org.

Bennewicz, J., & Barczak, T. (2020). Ground beetles (Carabidae) of field margin habitats. *Biologia*, *75*(10), 1631-1641. https://doi.org/10.2478/s11756-020-00424-y.

Bilde, T., Axelsen, J. A., & Toft, S. (2000). The value of Collembola from agricultural soils as food for a generalist predator. *Journal of Applied Ecology*, *37*(4), 672-683. https://doi.org/10.1046/j.1365-2664.2000.00527.x.

Blazka, P., & Fischer, Z. (2014). Moisture, water holding, drying and wetting in forest soils. *Open Journal of Soil Science*, *4*, 174-184. doi: 10.4236/ojss.2014.45021.

Boggs, L. M., Scheible, M. K., Machado, G., & Meiklejohn, K. A. (2019). Single fragment or bulk soil DNA metabarcoding: which is better for characterizing biological taxa found in surface soils for sample separation? *Genes*, *10*(6), 431. https://doi.org/10.3390/genes10060431.

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A. J., & Creer, S. (2022). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, *22*(4), 1231-1246. https://doi.org/10.1111/1755-0998.13512.

Brandon-Mong, G. J., Gan, H. M., Sing, K. W., Lee, P. S., Lim, P. E., & Wilson, J. J. (2015). DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. *Bulletin of Entomological Research*, *105*(6), 717-727. https://doi.org/10.1017/S0007485315000681.

Carew, M. E., & Hoffmann, A. A. (2015). Delineating closely related species with DNA barcodes for routine biological monitoring. *Freshwater Biology*, *60*(8), 1545-1560. https://doi.org/10.1111/fwb.12587.

Carpio, A. J., Castro, J., & Tortosa, F. S. (2019). Arthropod biodiversity in olive groves under two soil management systems: presence versus absence of herbaceous cover crop. *Agricultural and Forest Entomology*, *21*(1), 58-68. https://doi.org/10.1111/afe.12303.

Coissac, E., Riaz, T., & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular Ecology*, *21*(8), 1834-1847. https://doi.org/10.1111/j.1365-294X.2012.05550.x.

Colwell, R. K. (2009). EstimateS: Statistical estimation of species richness and shared species from samples. Version 8.2. User's Guide and application published at: http://purl.oclc.org/estimates.

Coulibaly, S. F., Winck, B. R., Akpa-Vinceslas, M., Mignot, L., Legras, M., Forey, E., & Chauvat, M. (2019). Functional assemblages of Collembola determine soil microbial communities and associated functions. *Frontiers in Environmental Science*, *7*, 52. https://doi.org/10.3389/fenvs.2019.00052.

Crotty, F. V., Demirer, U. A., Norris, S. L., Liu, W., & Murray, P. J. (2023). Evaluating the Impact of Long-Term Land Use Change and Age since Disturbance on Soil Faunal Diversity. *Forests*, *14*(9), 1882. https://doi.org/10.3390/f14091882.

Czigány, S., Sarkadi, N., Lóczy, D., Cséplő, A., Balogh, R., Fábián, S. Á., Ciglič, R., Ferk, M., Pirisi, G., Imre, M. and Nagy, G., … Pirkhoffer, E. (2023). Impact of agricultural land use types on soil moisture retention of loamy soils. *Sustainability*, *15*(6), 4925. https://doi.org/10.3390/su15064925.

de Graaff, M. A., Hornslein, N., Throop, H. L., Kardol, P., & van Diepen, L. T. (2019). Effects of agricultural intensification on soil biodiversity and implications for ecosystem functioning: a meta-analysis. *Advances in Agronomy*, *155*, 1-44. https://doi.org/10.1016/bs.agron.2019.01.001.

Decaëns, T., Jiménez, J. J., Gioia, C., Measey, G. J., & Lavelle, P. (2006). The values of soil animals for conservation biology. *European Journal of Soil Biology*, *42*, S23-S38. https://doi.org/10.1016/j.ejsobi.2006.07.001.

Decocq, G., Andrieu, E., Brunet, J., Chabrerie, O., De Frenne, P., De Smedt, P., Deconchat, M., Diekmann, M., Ehrmann, S., Giffard, B., Mifsud, E. G., Hansen, K., Hermy, M., Kolb, A., Lenoir, J., Liira, J., Moldan, F., Prokofieva, I., Rosenqvist, L., ... & Wulf, M. (2016). Ecosystem services from small forest patches in agricultural landscapes. *Current Forestry Reports*, *2*, 30-44. https://doi.org/10.1007/s40725-016-0028-x.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., De Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872-5895. https://doi.org/10.1111/mec.14350.

Donhauser, J., Briones, M. J., Mikola, J., Jones, D. L., Eder, R., Filser, J., Frossard, A., Krogh, P. H., Sousa, J. P., Cortet, J., Desie, E., Domene, X., Djuric, S., Hackenberger, D., Jimenez, J. J., Iamandei, M., Rissmann, C., Schmidt, O., Shanskiy, M., ... & Frey, B. (2023). Extracting DNA from soil or directly from isolated nematodes indicate dissimilar community structure for Europe-wide forest soils. *Soil Biology and Biochemistry*, *185*, 109154. https://doi.org/10.1016/j.soilbio.2023.109154.

Dontsova, K. M., & Norton, L. D. (2002). Clay Dispersion, Infiltration, and Erosion as Influenced by Exchangeable Ca and Mg. *Soil Science*, *167*(3), 184-193. https://doi.org/10.1097/00010694-200203000-00003.

Dopheide, A., Tooman, L. K., Grosser, S., Agabiti, B., Stevens, M. I., Nelson, N., Buckley, T. R., Drummond, A. J., & Newcomb, R. D. (2019). Estimating the biodiversity of terrestrial invertebrates on a forested island using DNA barcodes and metabarcoding data. *Ecological Applications*, *29*(4), e01877. https://doi.org/10.1002/eap.1877.

Dritsoulas, A., & Duncan, L. W. (2020). Optimizing for taxonomic coverage: a comparison of methods to recover mesofauna from soil. *Journal of Nematology*, *52*, e2020-104. https://doi.org/10.21307/jofnem-2020-104.

Druciarek, T., Lewandowski, M., & Tzanetakis, I. (2019). A new, sensitive and efficient method for taxonomic placement in the Eriophyoidea and virus detection in individual eriophyoids. *Experimental and Applied Acarology*, *78*, 247–261. https://doi.org/10.1007/s10493-019-00382-4.

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*(10), 996–998. https://doi.org/10.1038/nmeth.2604.

Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, *31*(21), 3476-3482. https://doi.org/10.1093/bioinformatics/btv401.

Eitzinger, B., Micic, A., Körner, M., Traugott, M., & Scheu, S. (2013). Unveiling soil food web links: New PCR assays for detection of prey DNA in the gut of soil arthropod predators. *Soil Biology and Biochemistry*, *57*, 943-945. https://doi.org/10.1016/j.soilbio.2012.09.001.

Elbrecht, V. (2018). JAMP: just another metabarcoding pipeline. R Package Version 0.67.

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—sequence relationships with an innovative metabarcoding protocol. *PLoS ONE*, *10*(7), e0130324. https://doi.org/10.1371/journal.pone.0130324.

Elias, M., Hill, R. I., Willmott, K. R., Dasmahapatra, K. K., Brower, A. V., Mallet, J., & Jiggins, C. D. (2007). Limited performance of DNA barcoding in a diverse community of tropical butterflies. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1627), 2881-2889. https://doi.org/10.1098/rspb.2007.1035.

FAO, ITPS, GSBI, CBD, & EC. (2020). *State of knowledge of soil biodiversity - Status, challenges and potentialities*, *Report 2020*. Rome, FAO. https://doi.org/10.4060/cb1928en.

FAO. (2022). *Framework for Action on Biodiversity for Food and Agriculture*. FAO Commission on Genetic Resources for Food and Agriculture. Rome. https://doi.org/10.4060/cb8338en.

Ferguson, S. H. (2001). Changes in trophic abundance of soil arthropods along a grass-shrub-forest gradient. *Canadian Journal of Zoology*, *79*(3), 457-464. https://doi.org/10.1139/z01-002.

Ferguson, S. H., & Joly, D. O. (2002). Dynamics of springtail and mite populations: the role of density dependence, predation, and weather. *Ecological Entomology*, *27*(5), 565-573. https://doi.org/10.1046/j.1365-2311.2002.00441.x.

Frézal, L., & Leblois, R. (2008). Four years of DNA barcoding: current advances and prospects. *Infection, Genetics and Evolution*, *8*(5), 727-736. https://doi.org/10.1016/j.meegid.2008.05.005.

Gao, M., He, P., Zhang, X., Liu, D., & Wu, D. (2014). Relative roles of spatial factors, environmental filtering and biotic interactions in fine-scale structuring of a soil mite

community. *Soil Biology and Biochemistry*, *79*, 68-77. https://doi.org/10.1016/j.soilbio.2014.09.003.

Gardi, C., Jeffery, S., & Saltelli, A. (2013). An estimate of potential threats levels to soil biodiversity in EU. *Global Change Biology*, *19*(5), 1538-1548. https://doi.org/10.1111/gcb.12159.

Geisen, S., Wall, D. H., & van der Putten, W. H. (2019). Challenges and opportunities for soil biodiversity in the anthropocene. *Current Biology*, *29*(19), R1036-R1044. https://doi.org/10.1016/j.cub.2019.08.007.

George, P. B., Keith, A. M., Creer, S., Barrett, G. L., Lebron, I., Emmett, B. A., Robinson, D. A., & Jones, D. L. (2017). Evaluation of mesofauna communities as soil quality indicators in a national-level monitoring programme. *Soil Biology and Biochemistry*, *115*, 537-546. https://doi.org/10.1016/j.soilbio.2017.09.022.

Gergócs, V., & Hufnagel, L. (2009). Application of oribatid mites as indicators. *Applied Ecology and Environmental Research*, *7*(1), 79-98.

Girvan, M. S., Campbell, C. D., Killham, K., Prosser, J. I., Glover, L. A. (2005). Bacterial diversity promotes community stability and functional resilience after perturbation. *Environmental Microbiology*, *7*(3), 301-313. https://doi.org/10.1111/j.1462-2920.2005.00695.x.

Guillot, G., & Rousset, F. (2013). Dismantling the Mantel tests. *Methods in Ecology and Evolution*, *4*(4), 336-344. https://doi.org/10.1111/2041-210x.12018.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270*, 313–321. https://doi.org/10.1098/rspb.2002.2218.

Hebert, P. D., Ratnasingham, S., & De Waard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*, S96-S99. https://doi.org/10.1098/rsbl.2003.0025.

Hebert, P. D., Ratnasingham, S., Zakharov, E. V., Telfer, A. C., Levesque-Beaudin, V., Milton, Pedersen, S., Jannetta, P. M., & deWaard, J. R. (2016). Counting animal species with DNA barcodes: Canadian insects. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1702), 20150333. https://doi.org/10.1098/rstb.2015.0333.

Heděnec, P., Jiménez, J. J., Moradi, J., Domene, X., Hackenberger, D., Barot, S., Frossard, A., Oktaba, L., Filser, J., Kindlmann, P., & Frouz, J. (2022). Global distribution of soil fauna functional groups and their estimated litter consumption across biomes. *Scientific Reports*, *12*(1), 17362. https://doi.org/10.1038/s41598-022-21563-z.

Heidemann, K., Hennies, A., Schakowske, J., Blumenberg, L., Ruess, L., Scheu, S., & Maraun, M. (2014). Free-living nematodes as prey for higher trophic levels of forest soil food webs. *Oikos*, *123*(10), 1199-1211. https://doi.org/10.1111/j.1600-0706.2013.00872.x.

Hendrickx, F., Maelfait, J. P., Van Wingerden, W., Schweiger, O., Speelmans, M., Aviron, S., Augenstein, I., Billeter, R., Bailey, D., Bukacek, R. and Burel, F., Diekötter, T., Dirksen, J., Herzog, F., Liira, J., Roubalova, M., Vandomme, V., & Bugter, R. (2007). How landscape structure, land-use intensity and habitat diversity affect components of total

arthropod diversity in agricultural landscapes. *Journal of Applied Ecology*, *44*(2), 340-351. https://doi.org/10.1111/j.1365-2664.2006.01270.x.

Herman, P. M., & Heip, C. (1988). On the use of meiofauna in ecological monitoring: who needs taxonomy? *Marine Pollution Bulletin*, *19*(12), 665-668. https://doi.org/10.1016/0025-326X(88)90386-4.

Hermans, S. M., Lear, G., Buckley, T. R., & Buckley, H. L. (2022). Environmental DNA sampling detects between-habitat variation in soil arthropod communities, but is a poor indicator of fine-scale spatial and seasonal variation. *Ecological Indicators*, *140*, 109040. https://doi.org/10.1016/j.ecolind.2022.109040.

Heydari, M., Eslaminejad, P., Kakhki, F. V., Mirab-balou, M., Omidipour, R., Prévosto, B., Kooch, Y., & Lucas-Borja, M. E. (2020). Soil quality and mesofauna diversity relationship are modulated by woody species and seasonality in semiarid oak forest. *Forest Ecology and Management*, *473*, 118332. https://doi.org/10.1016/j.foreco.2020.118332.

Hoage, J. F. J. (2018). *Metabarcoding soil microarthropods for soil quality assessment: Importance of integrated taxonomy, phylogenetic marker selection and sampling design* [Doctoral dissertation, Laurentian University of Sudbury]. https://zone.biblio.laurentian.ca/handle/10219/3188.

Hopkin, S. P. (1997). *Biology of the Springtails:(Insecta: Collembola)*. Oxford University Press.

Hulsen T. (2021). BioVenn: Create Area-Proportional Venn Diagrams from Biological Lists. R Package Version 1.1.3. https://cran.r-project.org/package=BioVenn.

Iatrou, S. A., Kavallieratos, N. G., Palyvos, N. E., Buchelos, C. T., & Tomanović, S. (2010). Acaricidal effect of different diatomaceous earth formulations against Tyrophagus putrescentiae (Astigmata: Acaridae) on stored wheat. *Journal of Economic Entomology*, *103*(1), 190-196. https://doi.org/10.1603/EC08213.

Joly, F. X., Coq, S., Coulis, M., Nahmani, J., & Hättenschwiler, S. (2018). Litter conversion into detritivore faeces reshuffles the quality control over C and N dynamics during decomposition. *Functional Ecology*, *32*(11), 2605-2614. https://doi.org/10.1111/1365-2435.13178.

Kaila, L., & Ståhls, G. (2006). DNA barcodes: Evaluating the potential of COI to diffentiate closely related species of Elachista (Lepidoptera: Gelechioidea: Elachistidae) from Australia. *Zootaxa*, *1170*(1), 1-26. https://doi.org/10.11646/zootaxa.1170.1.1.

Kamczyc, J., Dyderski, M. K., Horodecki, P., & Jagodziński, A. M. (2022). Temperature and precipitation affect seasonal changes in mite communities (Acari: Mesostigmata) in decomposing litter of broadleaved and coniferous temperate tree species. *Annals of Forest Science*, *79*(1), 12. https://doi.org/10.1186/s13595-022-01129-9.

Kardol, P., Reynolds, W. N., Norby, R. J., & Classen, A. T. (2011). Climate change effects on soil microarthropod abundance and community structure. *Applied Soil Ecology*, *47*(1), 37-44. https://doi.org/10.1016/j.apsoil.2010.11.001.

Keck, F., Couton, M., & Altermatt, F. (2023). Navigating the seven challenges of taxonomic reference databases in metabarcoding analyses. *Molecular Ecology Resources*, *23*(4), 742-755. https://doi.org/10.1111/1755-0998.13746.

Kennedy, S. R., Prost S., Overcast, I., Rominger, A. J., Gillespie, R. G., & Krehenwinkel, H. (2020). High-throughput sequencing for community analysis: the promise of DNA barcoding to uncover diversity, relatedness, abundances and interactions in spider communities. *Development Genes and Evolution*, *230*, 185-201. https://doi.org/10.1007/s00427-020-00652-x.

Kheradmand, K., Kamali, K., Fathipour, Y., & Goltapeh, E. M. (2007). Development, life table and thermal requirement of Tyrophagus putrescentiae (Astigmata: Acaridae) on mushrooms. *Journal of Stored Products Research*, *43*(3), 276-281. https://doi.org/10.1016/j.jspr.2006.06.007.

Kindt, R., & Coe, R. (2005). *Tree Diversity Analysis. A Manual and Software for Common Statistical Methods for Ecological and Biodiversity Studies*. World Agroforestry Centre.

Kirichenko-Babko, M. B., Danko, Y. M., Danylkiv, J. M., & Majerek, D. (2021). Comparison of the use of species abundance and presence-absence data for diversity assessment. *Journal of Physics: Conference Series*, *1736*(1), 012044. doi: 10.1088/1742-6596/1736/1/012044.

Kirse, A., Bourlat, S. J., Langen, K., & Fonseca, V. G. (2021). Metabarcoding Malaise traps and soil eDNA reveals seasonal and local arthropod diversity shifts. *Scientific Reports*, *11*(1), 10498. https://doi.org/10.1038/s41598-021-89950-6.

Koehler, H. H. (1999). Predatory mites (Gamasina, Mesostigmata). *Agriculture, Ecosystems & Environment*, *74*(1-3), 395-410. https://doi.org/10.1016/S0167-8809(99)00045-6.

Koehler, H., & Born, H. (1989). The influence of vegetation structure on the development of soil mesofauna. *Agriculture, Ecosystems & Environment*, *27*(1-4),253-269. https://doi.org/10.1016/0167-8809(89)90090-X.

Kokořová, P., & Starý, J. (2017). Communities of oribatid mites (Acari: Oribatida) of naturally regenerating and salvage-logged montane spruce forests of Šumava Mountains. *Biologia*, *72*(4), 445-451. https://doi.org/10.1515/biolog-2017-0050.

Korboulewsky, N., Perez, G., & Chauvat, M. (2016). How tree diversity affects soil fauna diversity: a review. *Soil Biology and Biochemistry*, *94*, 94-106. https://doi.org/10.1016/j.soilbio.2015.11.024.

Koudji, E. G., Despland, E., Caron, A. S., & Handa, I. T. (2023). Soil Springtail Communities Are Resilient to Forest Tent Caterpillar Defoliation in Quebec Mixed Hardwood Forests. *Forests*, *14*(7), 1302. https://doi.org/10.3390/f14071302.

Krantz, G. W., & Walter, D. E. (2009). *A Manual of Acarology* (3rd ed.). Texas Tech U.

Krehenwinkel, H., Graze, M., Rödder, D., Tanaka, K., Baba, Y. G., Muster, C., & Uhl, G. (2016). A phylogeographical survey of a highly dispersive spider reveals eastern Asia as a major glacial refugium for Palaearctic fauna. *Journal of Biogeography*, *43*(8), 1583-1594. https://doi.org/10.1111/jbi.12742.

Kuznetsova, N. A. (2006). Long-term dynamics of Collembola in two contrasting ecosystems. *Pedobiologia*, *50*(2), 157-164. https://doi.org/10.1016/j.pedobi.2005.12.004.

Lakshmi, G., Okafor, B. N., & Visconti, D. (2020). Soil microarthropods and nutrient cycling. *Environment, Climate, Plant and Vegetation Growth*, 453-472. https://doi.org/10.1007/978-3-030-49732-3_18.

Lavelle, P., Decaëns, T., Aubert, M., Barot, S., Blouin, M., Bureau, F., Margerie, P., Mora, P. & Rossi, J.P. (2006). Soil invertebrates and ecosystem services. *European Journal of Soil Biology*, *42*, S3-S15. https://doi.org/10.1016/j.ejsobi.2006.10.002.

Lebrun, P., & van Straalen, N. M. (1995). Oribatid mites: prospects for their use in ecotoxicology. *Experimental & Applied Acarology*, *19*, 361-379. https://doi.org/10.1007/BF00145154.

Lee, K. E. (1994). The biodiversity of soil organisms. *Applied Soil Ecology*, *1*(4), 251-254. https://doi.org/10.1016/0929-1393(94)90001-9.

Lehmitz, R., Russell, D., Hohberg, K., Christian, A., & Xylander, W. E. (2012). Active dispersal of oribatid mites into young soils. *Applied Soil Ecology*, *55*, 10-19. https://doi.org/10.1016/j.apsoil.2011.12.003.

Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., & Burridge, C. P. (2020). A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology*, *45*(3), 373-385. https://doi.org/10.1111/een.12831.

Lukic, D., Eberle, J., Thormann, J., Holzschuh, C., & Ahrens, D. (2021). Excluding spatial sampling bias does not eliminate oversplitting in DNA-based species delimitation analyses. *Ecology and Evolution*, *11*(15), 10327-10337. https://doi.org/10.1002/ece3.7836.

Lumley, L., Beaulieu, F., Behan-Pelletier, V., Knee, W., Lindquist, E. E., Mark, M., Proctor, H., & Walter, D. (2013). The status of systematic knowledge of the Acari of Canada: tickin'away with some mitey progress. *Newsletter of the Biological Survey of Canada*, *32*(2), 23-37.

Manninen, N., Soinne, H., Lemola, R., Hoikkala, L., & Turtola, E. (2018). Effects of agricultural land use on dissolved organic carbon and nitrogen in surface runoff and subsurface

drainage. *Science of the Total Environment*, *618*, 1519-1528. https://doi.org/10.1016/j.scitotenv.2017.09.319.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet*, *17*(1), 10–12. https://doi.org/10.14806/ej.17.1.200.

Martinez, A. P. (2020). pairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.4.

Matheson, C. D., Gurney, C., Esau, N., & Lehto, R. (2010). Assessing PCR inhibition from humic substances. *The Open Enzyme Inhibition Journal*, *3*(1), 38-45. https://doi.org/10.2174/1874940201003010038.

Maraun, M., & Scheu, S. (2000). The structure of oribatid mite communities (Acari, Oribatida): patterns, mechanisms and implications for future research. *Ecography*, *23*(3), 374-382. https://doi.org/10.1111/j.1600-0587.2000.tb00294.x.

McMurtry, J. A., Sourassou, N. F., & Demite, P. R. (2015). The Phytoseiidae (Acari: Mesostigmata) as biological control agents. *Prospects for Biological Control of Plant Feeding Mites and Other Harmful Organisms*, 133-149. https://doi.org/10.1007/978-3-319-15042-0_5.

Meehan, M. L., Song, Z., Lumley, L. M., Cobb, T. P., & Proctor, H. (2019). Soil mites as bioindicators of disturbance in the boreal forest in northern Alberta, Canada: Testing taxonomic sufficiency at multiple taxonomic levels. *Ecological Indicators*, *102*, 349-365. https://doi.org/10.1016/j.ecolind.2019.02.043.

Meier, R., Shiyang, K., Vaidya, G., & Ng, P. K. (2006). DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic Biology*, *55*(5), 715-728. https://doi.org/10.1080/10635150600969864.

Montgomery, I., Caruso, T., & Reid, N. (2020). Hedgerows as ecosystems: service delivery, management, and restoration. *Annual Review of Ecology, Evolution, and Systematics*, *51*, 81-102. https://doi.org/10.1146/annurev-ecolsys-012120-100346.

Morris, A. J., Bailey, C. M., Dillon, I. A., Gruar, D. J., & Westbury, D. B. (2010). Improving floristically enhanced field margins for wildlife. *Aspects of Applied Biology*, (100), 353-357.

Needelman, B. A., Kleinman, P. J., Strock, J. S., & Allen, A. L. (2007). Drainage Ditches: Improved management of agricultural drainage ditches for water quality protection: An overview. *Journal of Soil and Water Conservation*, *62*(4), 171-178.

Ondov, B. D., Bergman, N. H., & Phillippy, A. M. (2011). Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics*, *12*(1), 1-10. https://doi.org/10.1186/1471-2105-12-385.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., … Wagner, H. (2015). vegan: Community ecology package. R Package Version 2.3.

Oliverio, A. M., Gan, H., Wickings, K., & Fierer, N. (2018). A DNA metabarcoding approach to characterize soil arthropod communities. *Soil Biology and Biochemistry*, *125*, 37-43. https://doi.org/10.1016/j.soilbio.2018.06.026.

Petrova, V., Salmane, I., & Çudare, Z. (2004). The predatory mite (Acari, Parasitiformes: Mesostigmata (Gamasina); Acariformes: Prostigmata) community in strawberry agrocenosis. *Acta Universitatis Latviensis, Biology*, *676*, 87-95.

Porco, D., Decaëns, T., Deharveng, L., James, S. W., Skarżyński, D., Erséus, C., Butt, K. R., Richard, B., & Hebert, P. D. (2013). Biological invasions in soil: DNA barcoding as a monitoring tool in a multiple taxa survey targeting European earthworms and springtails in North America. *Biological Invasions*, *15*, 899-910. https://doi.org/10.1007/s10530-012-0338-2.

Porter, T. M., Morris, D. M., Basiliko, N., Hajibabaei, M., Doucet, D., Bowman, S., Emilson, E. J., Emilson, C. E., Chartrand, D., Wainio-Keizer, K., Séguin, A., Séguin, A., & Venier, L. (2019). Variations in terrestrial arthropod DNA metabarcoding methods recovers robust beta diversity but variable richness and site indicators. *Scientific Reports*, *9*(1), 18218. https://doi.org/10.1038/s41598-019-54532-0.

Postma-Blaauw, M. B., De Goede, R. G., Bloem, J., Faber, J. H., & Brussaard, L. (2012). Agricultural intensification and de-intensification differentially affect taxonomic diversity of predatory mites, earthworms, enchytraeids, nematodes and bacteria. *Applied Soil Ecology*, *57*, 39-49. https://doi.org/10.1016/j.apsoil.2012.02.011.

Potapov, A. M., Beaulieu, F., Birkhofer, K., Bluhm, S. L., Degtyarev, M. I., Devetter, M., Goncharov, A. A., Gongalsky, K. B., Klarner, B., Korobushkin, D. I., Liebke, D. F., Maraun, M., Mc Donnell, R. J., Pollierer, M. M., Schaefer, I., Shrubovych, J., Semeyuk, I. I., Sendra, A., Tuma, J., ... & Scheu, S. (2022). Feeding habits and multifunctional classification of soil-associated consumers from protists to vertebrates. *Biological Reviews*, *97*(3), 1057-1117. https://doi.org/10.1111/brv.12832.

Potapov, A., Bellini, B. C., Chown, S. L., Deharveng, L., Janssens, F., Kováč, Ľ., Kuznetsova, N., Ponge, J. F., Potapov, M., Querner, P., Russell, D., Sun, X., Zhang, F., & Berg, M. P. (2020). Towards a global synthesis of Collembola knowledge: challenges and potential solutions. *Soil Organisms*, *92*(3), 161-188. https://hal.science/hal-03040767.

Potapov, A. M., Sun, X., Briones, M. J., Brown, G. G., Cameron, E. K., Cortet, J., Eisenhauer, N., Fujii, S., Geisen, S., Gongalsky, K. B., Guerra, C., Handa, I. T., Janion-Scheepers, C., Lindo, Z., Mathieu, J., Murvanidze, M., Nielsen, U. N., Scheu, S., Schmidt, O., ... & Wall, D. H. (2022). Global monitoring of soil animal communities using a common methodology. *bioRxiv*, 2022-01. https://doi.org/10.1101/2022.01.11.475793.

R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rantalainen, M. L., Fritze, H., Haimi, J., Pennanen, T., & Setälä, H. (2005). Colonisation of newly established habitats by soil decomposer organisms: the effect of habitat corridors in relation to colonisation distance and habitat size. *Applied Soil Ecology*, *28*(1), 67-77. https://doi.org/10.1016/j.apsoil.2004.06.004.

Ratnasingham S., & Hebert P. D. N. (2007). BOLD: The Barcode of Life Data system (http://www.barcodinglife.org). *Molecular Ecology Notes*, *7*(3), 355-364. https://doi.org/10.1111/j.1471-8286.2007.01678.x.

Read, D. S., Sheppard, S. K., Bruford, M. W., Glen, D. M., & Symondson, W. O. C. (2006). Molecular detection of predation by soil micro-arthropods on nematodes. *Molecular Ecology*, *15*(7), 1963-1972. https://doi.org/10.1111/j.1365-294X.2006.02901.x.

Recuero, E., Etzler, F. E., & Caterino, M. S. (2023). Most soil and litter arthropods are unidentifiable based on current DNA barcode reference libraries. *Current Zoology*, zoad051. https://doi.org/10.1093/cz/zoad051.

Roesch, L. F., Fulthorpe, R. R., Riva, A., Casella, G., Hadwin, A. K., Kent, A. D., Daroub, S. H., Camargo, F. A., Farmerie, W. G., & Triplett, E. W. (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *The ISME Journal*, *1*, 283-290. https://doi.org/10.1038/ismej.2007.53.

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. doi: 10.7717/peerj.2584.

Rueda-Ramírez, D., Carta, L., Mowery, J., Bauchan, G., Ochoa, R., Young, M., Santos, J. C., & Palevsky, E. (2022). In memory of Gary Bauchan: Integrated taxonomy of soil mites in farming systems. *Systematic and Applied Acarology*, *27*(2), 181-208. https://doi.org/10.11158/saa.27.2.3.

Ruf, A., & Beck, L. (2005). The use of predatory soil mites in ecological soil classification and assessment concepts, with perspectives for oribatid mites. *Ecotoxicology and Environmental Safety*, *62*(2), 290-299. https://doi.org/10.1016/j.ecoenv.2005.03.029.

Sabu, T. K., Shiju, R., Vinod, K., & Nithya, S. (2011). A comparison of the pitfall trap, Winkler extractor and Berlese funnel for sampling ground-dwelling arthropods in tropical Montane cloud forests. *Journal of Insect Science*, *11*, 28. https://doi.org/10.1673/031.011.0128.

Saitoh, S., Aoyama, H., Fujii, S., Sunagawa, H., Nagahama, H., Akutsu, M., Shinzato, N., Kaneko, N. & Nakamori, T. (2016). A quantitative protocol for DNA metabarcoding of

springtails (Collembola). *Genome*, *59*(9), 705-723. https://doi.org/10.1139/gen-2015-0228@gen-iblf.issue01.

Schenk, J., Kleinbölting, N., & Traunspurger, W. (2020). Comparison of morphological, DNA barcoding, and metabarcoding characterizations of freshwater nematode communities. *Ecology and Evolution*, *10*(6), 2885-2899. https://doi.org/10.1002/ece3.6104.

Schirmel, J., Thiele, J., Entling, M. H., & Buchholz, S. (2016). Trait composition and functional diversity of spiders and carabids in linear landscape elements. *Agriculture, Ecosystems & Environment*, *235*, 318-328. https://doi.org/10.1016/j.agee.2016.10.028.

Schneider, K., & Maraun, M. (2009). Top-down control of soil microarthropods–Evidence from a laboratory experiment. *Soil Biology and Biochemistry*, *41*(1), 170-175. https://doi.org/10.1016/j.soilbio.2008.10.013.

Semenov, M. V. (2021). Metabarcoding and metagenomics in soil ecology research: achievements, challenges, and prospects. *Biology Bulletin Reviews*, *11*, 40-53. https://doi.org/10.1134/S2079086421010084.

Shaw, R. F., Johnson, P. J., Macdonald, D. W., & Feber, R. E. (2015). Enhancing the biodiversity of ditches in intensively managed UK farmland. *PLoS ONE*, *10*(10), e0138306. https://doi.org/10.1371/journal.pone.0138306.

Sire, L., Yáñez, P. S., Bézier, A., Courtial, B., Mbedi, S., Sparmann, S., Larrieu, L., Rougerie, R., Bouget, C., Monaghan, M. T., Herniou, E.A., & Lopez-Vaamonde, C. (2023). Persisting roadblocks in arthropod monitoring using non-destructive metabarcoding from collection media of passive traps. *PeerJ*, 11, e16022. doi: 10.7717/peerj.16022.

Smith, J., Chapman, A., & Eggleton, P. (2006). Baseline biodiversity surveys of the soil macrofauna of London's green spaces. *Urban Ecosystems*, *9*, 337-349. https://doi.org/10.1007/s11252-006-0001-8.

Smith, A. P., Marín-Spiotta, E., & Balser, T. (2015). Successional and seasonal variations in soil and litter microbial community structure and function during tropical postagricultural forest regeneration: a multiyear study. *Global Change Biology*, *21*(9), 3532-3547. https://doi.org/10.1111/gcb.12947.

Somers, K. M., & Jackson, D. A. (2022). Putting the Mantel test back together again. *Ecology*, *103*(10), e3780. https://doi.org/10.1002/ecy.3780.

Song, H., Buhay, J. E., Whiting, M. F., & Crandall, K. A. (2008). Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences*, *105*(36), 13486-13491. https://doi.org/10.1073/pnas.0803076105.

Srivastava Lab. (2012). *Dichotomous Key for Microarthropod Identification-Mites And Other Microarthropods*. University of British Columbia. https://www.zoology.ubc.ca/~srivast/mites/key.html.

Stroot, L., Brinkert, A., Hölzel, N., Rüsing, A., & Bucharova, A. (2022). Establishment of wildflower strips in a wide range of environments: a lesson from a landscape-scale project. *Restoration Ecology*, *30*(4), e13542. https://doi.org/10.1111/rec.13542.

Szigeti, N., Berki, I., Vityi, A., & Winkler, D. (2022). Soil mesofauna and herbaceous vegetation patterns in an agroforestry landscape. *Agroforestry Systems*, *96*(4), 773-786. https://doi.org/10.1007/s10457-022-00739-6.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045-2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x.

Tibbett, M., Fraser, T. D., & Duddigan, S. (2020). Identifying potential threats to soil biodiversity. *PeerJ*, *8*, e9271. https://doi.org/10.7717/peerj.9271.

Tomlin, A. D., & Miller, J. J. (1987). Composition of the soil fauna in forested and grassy plots at Delhi, Ontario. *Canadian Journal of Zoology*, *65*(12), 3048-3055. https://doi.org/10.1139/z87-462.

Turnbull, M. S., & Stebaeva, S. (2019). Collembola of Canada. *ZooKeys*, (819), 187. doi: 10.3897/zookeys.819.23653.

Udawatta, R. P., Rankoth, L. M., & Jose, S. (2019). Agroforestry and biodiversity. *Sustainability*, *11*(10), 2879. https://doi.org/10.3390/su11102879.

Valdez, J. W., Callaghan, C. T., Junker, J., Purvis, A., Hill, S. L., & Pereira, H. M. (2023). The undetectability of global biodiversity trends using local species richness. *Ecography*, *2023*(3), e06604. https://doi.org/10.1111/ecog.06604.

Vanbergen, A. J., Watt, A. D., Mitchell, R., Truscott, A. M., Palmer, S. C., Ivits, E., Eggleton, P., Jones, T. H., & Sousa, J. P. (2007). Scale-specific correlations between habitat heterogeneity and soil fauna diversity along a landscape structure gradient. *Oecologia*, *153*, 713-725. https://doi.org/10.1007/s00442-007-0766-3.

van Bruggen, A. M. & Semenov, A. (2000). In search of biological indicators for soil health and disease suppression. *Applied Soil Ecology*, *15*(1), 13-24. https://doi.org/10.1016/S0929-1393(00)00068-8.

van Capelle, C., Schrader, S., & Brunotte, J. (2012). Tillage-induced changes in the functional diversity of soil biota–A review with a focus on German data. *European Journal of Soil Biology*, *50*, 165-181. https://doi.org/10.1016/j.ejsobi.2012.02.005.

van der Heyde, M., Bunce, M., & Nevill, P. (2022). Key factors to consider in the use of environmental DNA metabarcoding to monitor terrestrial ecological restoration. *Science of the Total Environment*, *848*, 157617. https://doi.org/10.1016/j.scitotenv.2022.157617.

van Straalen, N. M. (1998). Evaluation of bioindicator systems derived from soil arthropod communities. *Applied Soil Ecology*, *9*(1-3), 429-437. https://doi.org/10.1016/S0929-1393(98)00101-2.

Varela, E., Verheyen, K., Valdés, A., Soliño, M., Jacobsen, J. B., De Smedt, P., Ehrmann, S., Gärtner, S., Górriz, E., & Decocq, G. (2018). Promoting biodiversity values of small forest patches in agricultural landscapes: Ecological drivers and social demand. *Science of the Total Environment*, *619*, 1319-1329. https://doi.org/10.1016/j.scitotenv.2017.11.190.

Virgilio, M., Backeljau, T., Nevado, B., & De Meyer, M. (2010). Comparative performances of DNA barcoding across insect orders. *BMC Bioinformatics*, *11*(1), 1-10. https://doi.org/10.1186/1471-2105-11-206.

Walter, D. E., Latonas, S., Byers, K., & Lumley, L. (2014). Almanac of Alberta Oribatida Part I, Version 2.4. Edmonton, Alberta, Canada. doi: 10.13140/RG.2.2.15348.07044.

Watts, C., Dopheide, A., Holdaway, R., Davis, C., Wood, J., Thornburrow, D., & Dickie, I. A. (2019). DNA metabarcoding as a tool for invertebrate community monitoring: a case

study comparison with conventional techniques. *Austral Entomology*, *58*(3), 675-686. https://doi.org/10.1111/aen.12384.

Wilson, J. J., Sing, K. W., Floyd, R. M., & Hebert, P. D. (2017). DNA barcodes and insect biodiversity. *Insect Biodiversity: Science and Society*, 575-592. https://doi.org/10.1002/9781118945568.ch17.

Wissuwa, J., Salamon, J. A., & Frank, T. (2013). Oribatida (Acari) in grassy arable fallows are more affected by soil properties than habitat age and plant species. *European Journal of Soil Biology*, *59*, 8-14. https://doi.org/10.1016/j.ejsobi.2013.08.002.

Wu, P., & Wang, C. (2019). Differences in spatiotemporal dynamics between soil macrofauna and mesofauna communities in forest ecosystems: the significance for soil fauna diversity monitoring. *Geoderma*, *337*, 266-272. https://doi.org/10.1016/j.geoderma.2018.09.031.

Yang, C., Wang, X., Miller, J. A., de Blécourt, M., Ji, Y., Yang, C., Harrison, R. D., & Douglas, W. Y. (2014). Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. *Ecological Indicators*, *46*, 379-389. https://doi.org/10.1016/j.ecolind.2014.06.028.

Young, M. R., & Hebert, P. D. (2022). Unearthing soil arthropod diversity through DNA metabarcoding. *PeerJ*, *10*, e12845. https://doi.org/10.7717/peerj.12845.

Zhang, B., Chen, T. W., Mateos, E., Scheu, S., & Schaefer, I. (2019). DNA-based approaches uncover cryptic diversity in the European Lepidocyrtus lanuginosus species group (Collembola: Entomobryidae). *Invertebrate Systematics*, *33*(4), 661-670. https://doi.org/10.1071/IS18068.

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., et al. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, *28*(8), 1857-1862. https://doi.org/10.1111/mec.15060.

Appendices

Appendix 1. Species list of all animals identified in this thesis through metabarcoding.

- **Annelida**
  - **Clitellata**
    - Crassiclitellata
      - **Lumbricidae**
        - *Aporrectodea caliginosa*
        - *Aporrectodea* spp.
        - *Bimastos rubidus*
        - *Dendrobaena octaedra*
        - *Dendrodrilus rubidus*
        - *Lumbricus rubellus*
        - *Lumbricus terrestris*
        - *Octolasion tyrtaeum*
    - Enchytraeida
      - **Enchytraeidae**
        - *Buchholzia appendiculata*
        - *Enchytraeus dichaetus*
        - *Enchytraeus lacteus*
        - *Fridericia bulboides*
        - *Fridericia christeri*
        - *Fridericia dura*
        - *Fridericia eiseni*
        - *Fridericia galba*
        - *Fridericia paroniana*
        - *Fridericia ratzeli*
        - *Fridericia* spp.
        - *Lumbricillus latithecatus*
        - *Stercutus niveus*
        - Enchytraeidae spp.
- **Arthropoda**
  - **Arachnida**
    - Araneae
      - **Araneidae**
        - *Hypsosinga pygmaea*
      - **Linyphiidae**
        - *Agyneta fabra*
        - *Islandiana flaveola*
        - *Mermessus trilobatus*
        - *Pocadicnemis americana*
      - **Lycosidae**
        - *Trochosa ruricola*
      - **Theridiidae**

Robertus spinifer

Mesostigmata

Ameroseiidae

Ameroseius corbiculus

Ameroseius sp.

Ascidae

Arctoseius cetratus

Asca spp.

Gamasellodes sp.

Leioseius minusculus

Blattisociidae

Cheiroseius sp.

Digamasellidae

Dendrolaelaps sp.

Digamasellidae spp.

Laelapidae

Cosmolaelaps sp.

Gaeolaelaps aculeifer

Ololaelaps spp.

Laelapidae spp.

Macrochelidae

Geholaspis sp.

Pachylaelapidae

Pachyseius sp.

Pachylaelapidae sp.

Parasitidae

Lysigamasus misellus

Lysigamasus vagabundus

Lysigamasus spp.

Parasitidae spp.

Parholaspididae

Krantzholaspis spp.

Phytoseiidae

Proprioseiopsis okanagensis

Typhlodromips sp.

Rhodacaridae

Multidentorhodacarus sp.

Rhodacarellus silesiacus

Rhodacaridae spp.

Veigaiidae

Cyrthydrolaelaps spp.

Veigaia nemorensis

Veigaia sp.

Veigaiidae spp.

Pseudoscorpiones

Neobisiidae
Microbisium pygmaeum
Sarcoptiformes
Acaridae
Rhizoglyphus robini
Achipteriidae
Achipteria spp.
Astegistidae
Astegistidae spp.
Brachychthoniidae
Brachychthoniidae spp.
Ceratozetidae
Ceratozetes gracilis
Ceratozetes sp.
Ceratozetidae spp.
Chamobatidae
Chamobates cuspidatus
Chamobates sp.
Eniochthoniidae
Eniochthonius crosbyi
Eniochthoniidae spp.
Epilohmanniidae
Epilohmannia pallida
Epilohmannia sp.
Eremobelbidae
Eremobelba sp.
Euphthiracaridae
Rhysotritia ardua
Euzetidae
Euzetes globulus
Galumnidae
Acrogalumna longipluma
Acrogalumna sp.
Galumna lanceata
Galumna spp.
Pergalumna spp.
Galumnidae spp.
Gustaviidae
Gustavia microcephala
Gustavia spp.
Haplozetidae
Haplozetidae spp.
Hypochthoniidae
Hypochthonius luteus
Hypochthonius spp.

Mochlozetidae
Podoribates pratensis
Nanhermanniidae
Nanhermanniidae sp.
Nanorchestidae
Nanorchestes spp.
Speleorchestes spp.
Nanorchestidae spp.
Nothridae
Nothrus anauniensis
Nothrus spp.
Nothridae sp.
Oppiidae
Oppia nitens
Oppiella nova
Oppiella spp.
Oppiidae spp.
Oribatulidae
Oribatula tibialis
Oribatula spp.
Parakalummidae
Parakalummidae spp.
Phenopelopidae
Eupelops occultus
Phthiracaridae
Phthiracaridae sp.
Punctoribatidae
Punctoribates palustris
Punctoribatidae sp.
Scheloribatidae
Scheloribates badius
Scheloribates clavilanceolatus
Scheloribates spp.
Scheloribatidae spp.
Suctobelbidae
Suctobelbidae spp.
Tectocepheidae
Tectocepheus sarekensis
Tectocepheus spp.
Tectocepheidae spp.
Tegoribatidae
Tegoribatidae sp.
Trombidiformes
Ereynetidae
Ereynetidae spp.

Eriophyidae
Aculodes mckenziei
Erythraeidae
Erythraeidae sp.
Eupodidae
Eupodes spp.
Eupodidae spp.
Penthalodidae
Penthalodidae sp.
Pygmephoridae
Pygmephoridae spp.
Rhagidiidae
Rhagidia sp.
Rhagidiidae spp.
Scutacaridae
Scutacaridae spp.
Tarsonemidae
Tarsonemidae spp.
Trombidiidae
Trombidiidae spp.
Tydeidae
Microtydeus spp.
Tydeidae spp.

Chilopoda
Geophilomorpha
Geophilidae
Geophilus proximus
Geophilus spp.
Geophilidae spp.
Schendylidae
Schendyla nemorensis
Lithobiomorpha
Henicopidae
Lamyctes africanus
Lamyctes emarginatus
Lithobiidae
Lithobius forficatus
Lithobius microps
Lithobius sp.
Lithobiidae spp.

Collembola
Entomobryomorpha
Entomobryidae
Entomobrya multifasciata
Lepidocyrtus cyaneus

Lepidocyrtus paradoxus
Lepidocyrtus spp.
Orchesella villosa
Orchesella sp.
Pseudosinella alba
Pseudosinella spp.
Entomobryidae spp.

Isotomidae
Folsomia candida
Folsomia fimetaria
Folsomia spp.
Isotoma viridis
Isotomodes productus
Isotomodes sp.
Parisotoma notabilis
Parisotoma spp.
Isotomidae spp.

Tomoceridae
Pogonognathellus sp.
Tomoceridae spp.

Neelipleona

Neelidae
Megalothorax spp.
Neelidae spp.

Poduromorpha

Hypogastruridae
Hypogastruridae spp.

Neanuridae
Friesea claviseta
Neanura muscorum
Neanura sp.
Neanuridae spp.

Onychiuridae
Onychiurus yodai
Protaphorura spp.
Supraphorura furcifera
Supraphorura sp.
Thalassaphorura houtanensis
Thalassaphorura udhagaiensis
Thalassaphorura sp.
Onychiuridae spp.

Tullbergiidae
Mesaphorura spp.
Tullbergia macrochaeta
Tullbergiidae spp.

- Symphypleona
  - **Arrhopalitidae**
    - Arrhopalitidae sp.
  - **Bourletiellidae**
    - *Bourletiella* spp.
    - Bourletiellidae sp.
  - **Dicyrtomidae**
    - *Dicyrtomina* sp.
  - **Katiannidae**
    - *Sminthurinus elegans*
    - *Sminthurinus* spp.
    - Katiannidae spp.
  - **Sminthuridae**
    - Sminthuridae sp.
  - **Sminthurididae**
    - *Sminthurides* sp.
    - *Sphaeridia* sp.
    - Sminthurididae spp.
- **Diplopoda**
  - Julida
    - **Blaniulidae**
      - *Choneiulus palmatus*
    - **Julidae**
      - *Brachyiulus pusillus*
      - *Cylindroiulus caeruleocinctus*
      - *Cylindroiulus latestriatus*
      - *Cylindroiulus* sp.
      - *Julus scandinavius*
      - *Ophyiulus pilosus*
      - *Ophyiulus* spp.
      - Julidae spp.
  - Polydesmida
    - **Polydesmidae**
      - Polydesmidae sp.
- **Diplura**
  - Rhabdura
    - **Campodeidae**
      - *Campodea fragilis*
      - *Campodea plusiochaeta*
      - *Campodea* spp.
      - Campodeidae spp.
- **Insecta**
  - Blattodea
    - **Ectobiidae**
      - *Ectobius pallidus*

Ectobiidae spp.

Coleoptera

Cantharidae

Polemius laticornis

Carabidae

Agonum palustre

Agonum spp.

Anisodactylus harrisii

Anisodactylus sanctaecrucis

Bembidion mimus

Bradycellus neglectus

Dyschirius globulosus

Harpalus herbivagus

Carabidae sp.

Cryptophagidae

Atomaria ephippiata

Curculionidae

Exomias pellucidus

Polydrusus formosus

Polydrusus impressifrons

Sciaphilus asperatus

Elateridae

Agriotes mancus

Agriotes spp.

Elateridae spp.

Hydrophilidae

Anacaena globulus

Ptiliidae

Ptinella sp.

Ptinidae

Ptinus villiger

Scarabaeidae

Melinopterus prodromus

Staphylinidae

Amischa analis

Anotylus sp.

Atheta formicaensis

Bibloplectus minutissimus

Bledius gallicus

Drusilla canaliculata

Meotica pallens

Mocyta luteola

Philhygra clemens

Sunius melanocephalus

Tachinus limbatus

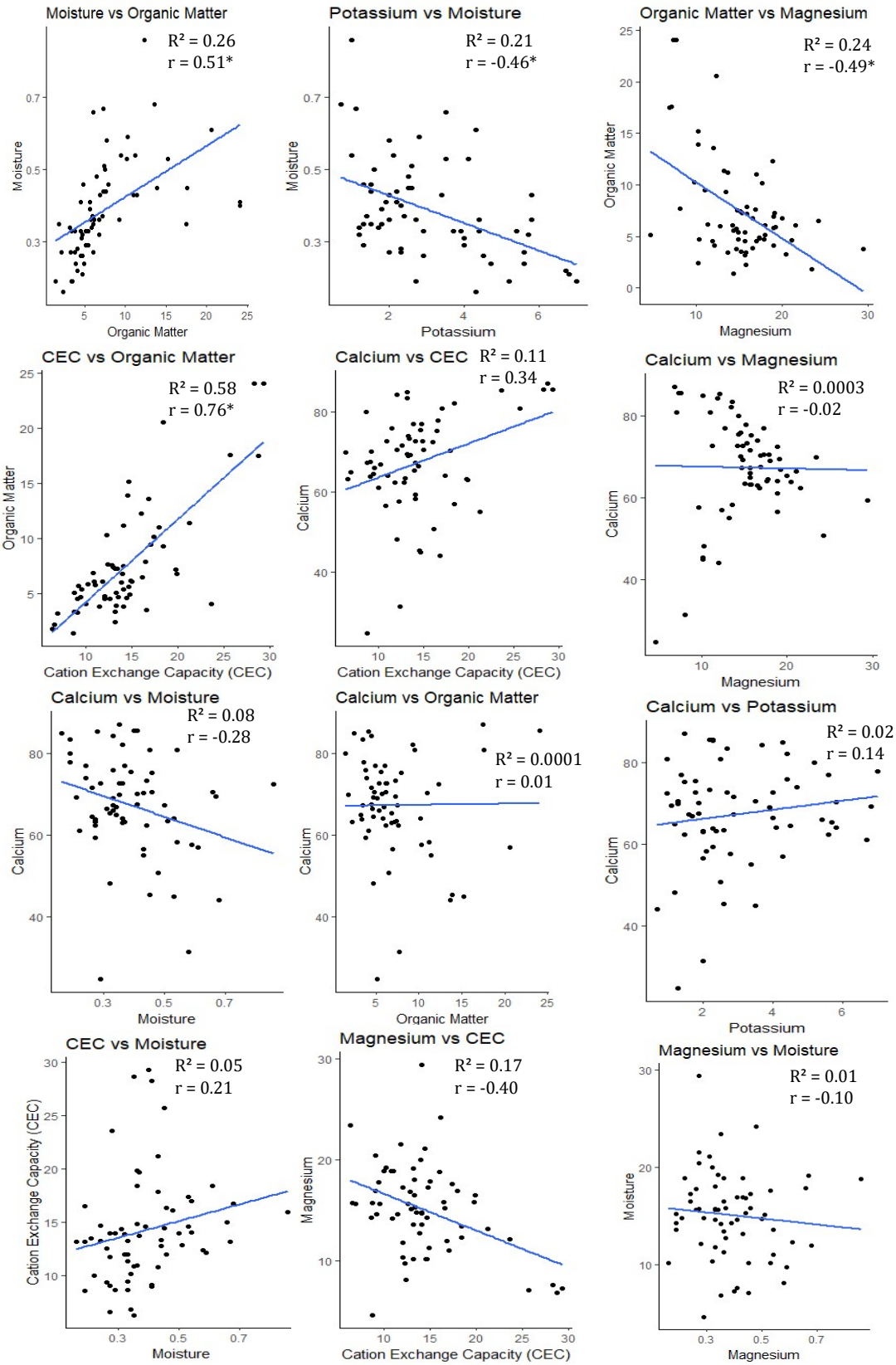
Diptera
- **Cecidomyiidae**
  - *Aphidoletes aphidimyza*
  - Cecidomyiidae sp.
- **Chironomidae**
  - *Cricotopus diversus*
- **Dolichopodidae**
  - *Chrysotus* sp.
- **Hybotidae**
  - *Platypalpus major*
- **Limoniidae**
  - *Rhipidia* sp.
- **Phoridae**
  - *Diplonevra nitidula*
  - *Megaselia* sp.
- **Sciaridae**
  - *Bradysia urticae*
  - *Corynoptera cursor*
  - *Corynoptera perpusilla*
  - *Corynoptera umbrata*
  - *Cratyna* sp.

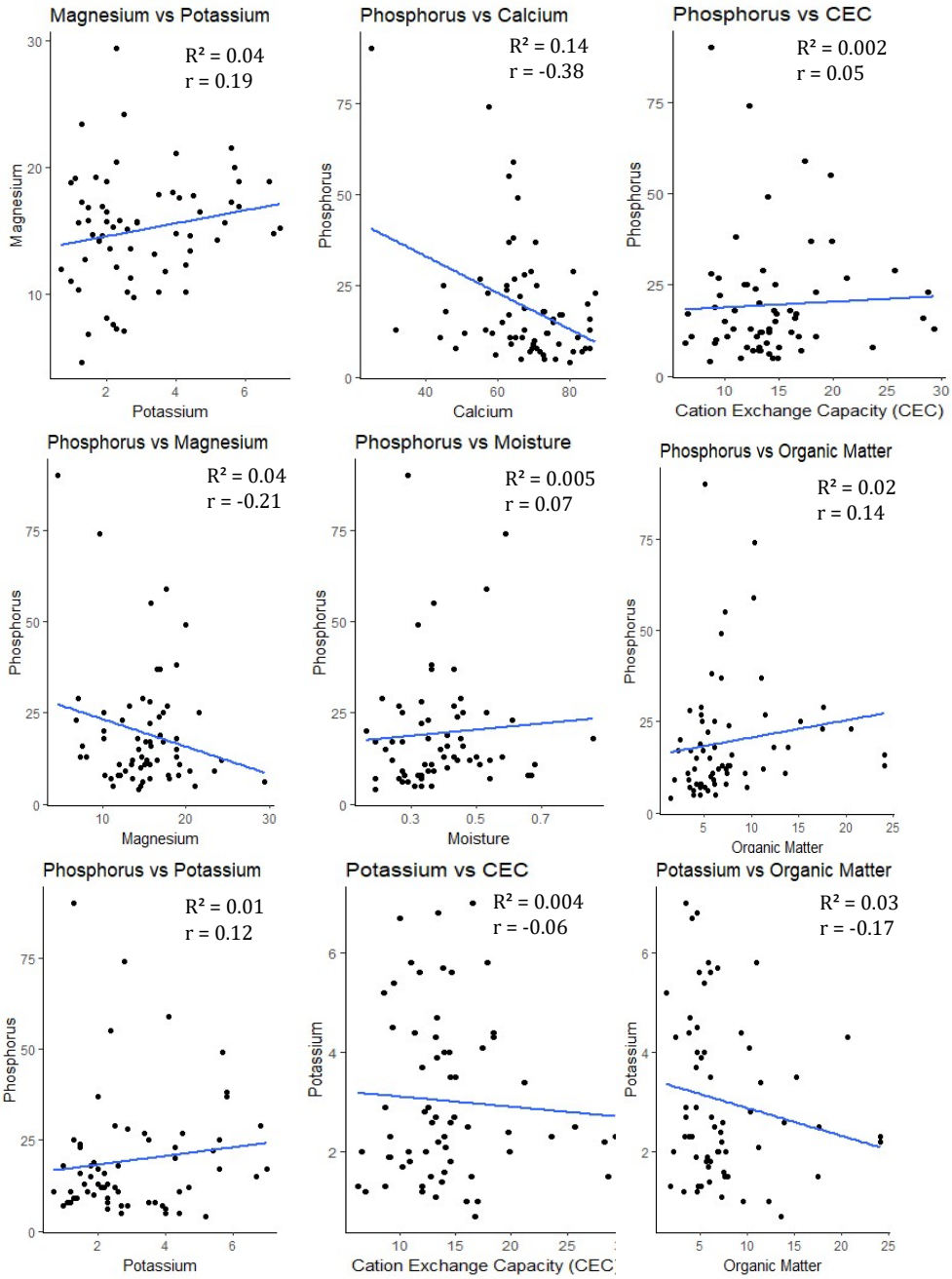
Hemiptera
- **Aphididae**
  - *Forda marginata*
  - *Pemphigus* sp.
  - *Prociphilus* spp.
- **Dictyopharidae**
  - Dictyopharidae sp.
- **Pseudococcidae**
  - Pseudococcidae sp.
- **Rhyparochromidae**
  - *Drymus unus*
  - *Scolopostethus thomsoni*

Hymenoptera
- **Ceraphronidae**
  - Ceraphronidae sp.
- **Formicidae**
  - *Formica cinerea*
  - *Formica fusca*
  - *Lasius alienus*
  - *Lasius nearcticus*
  - *Lasius neoniger*
  - *Lasius umbratus*
  - *Lasius* spp.
  - *Myrmecina americana*

Myrmica fracticornis
Myrmica spp.
Stenamma brevicorne
Stenamma sp.
Stigmatomma pallipes
Tapinoma sessile
Tetramorium caespitum
Formicidae sp.
Platygastridae
Leptacis sp.
Psocodea
Liposcelididae
Liposcelis bostrychophila
Psyllipsocidae
Dorypteryx domestica
Thysanoptera
Phlaeothripidae
Haplothrips sp.
Thripidae
Taeniothrips inconsequens
Thrips spp.
Malacostraca
Isopoda
Trichoniscidae
Hyloniscus riparius
Trichoniscus pusillus
Protura
Eosentomata
Eosentomidae
Eosentomidae spp.
Mollusca
Gastropoda
Stylommatophora
Agriolimacidae
Deroceras laeve
Arionidae
Arion distinctus
Valloniidae
Vallonia costata
Nematoda
Chromadorea
Rhabditida
Diplogastridae
Pristionchus entomophagus
Rhabditidae

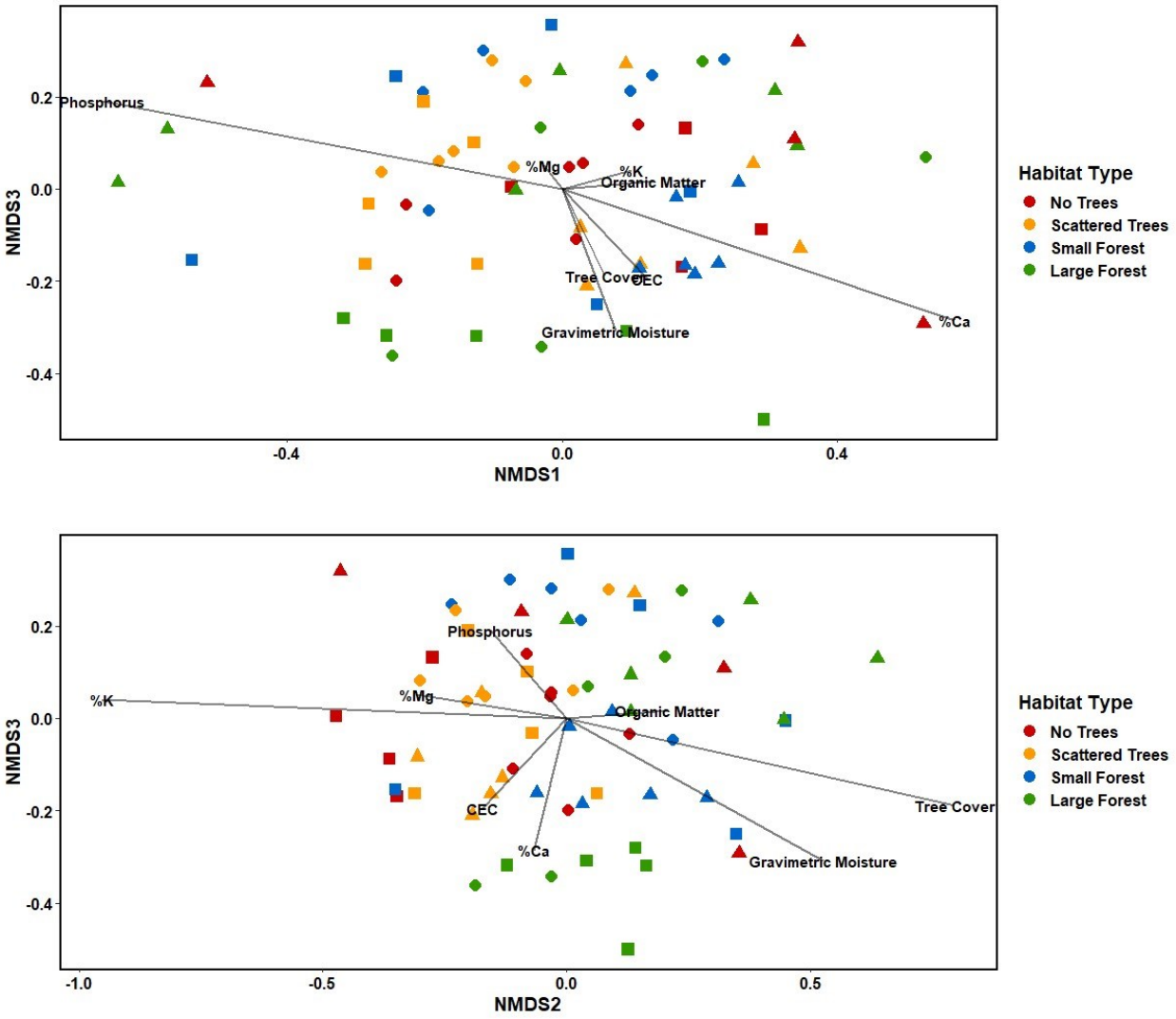
Rhabditidae spp.
Tylenchida
Aphelenchidae
Aphelenchus avenae
Parasitylenchidae
Parasitylenchidae sp.
Enoplea
Dorylaimida
Longidoridae
Xiphinema sp.
Platyhelminthes
Rhabditophora
Tricladida
Geoplanidae
Rhynchodemus sp.
Rotifera
Bdelloidea
Adinetida
Adinetidae
Adinetidae spp.
Philodinida
Habrotrochidae
Habrotrochidae spp.


Moisture vs Organic Matter
R² = 0.26
r = 0.51*
Moisture
Organic Matter
Potassium vs Moisture
R² = 0.21
r = -0.46*
Moisture
Potassium
Organic Matter vs Magnesium
R² = 0.24
r = -0.49*
Organic Matter
Magnesium
CEC vs Organic Matter
R² = 0.58
r = 0.76*
Organic Matter
Cation Exchange Capacity (CEC)
Calcium vs CEC
R² = 0.11
r = 0.34
Calcium
Cation Exchange Capacity (CEC)
Calcium vs Magnesium
R² = 0.0003
r = -0.02
Calcium
Magnesium
Calcium vs Moisture
R² = 0.08
r = -0.28
Calcium
Moisture
Calcium vs Organic Matter
R² = 0.0001
r = 0.01
Calcium
Organic Matter
Calcium vs Potassium
R² = 0.02
r = 0.14
Calcium
Potassium
CEC vs Moisture
R² = 0.05
r = 0.21
Cation Exchange Capacity (CEC)
Moisture
Magnesium vs CEC
R² = 0.17
r = -0.40
Magnesium
Cation Exchange Capacity (CEC)
Magnesium vs Moisture
R² = 0.01
r = -0.10
Moisture
Magnesium

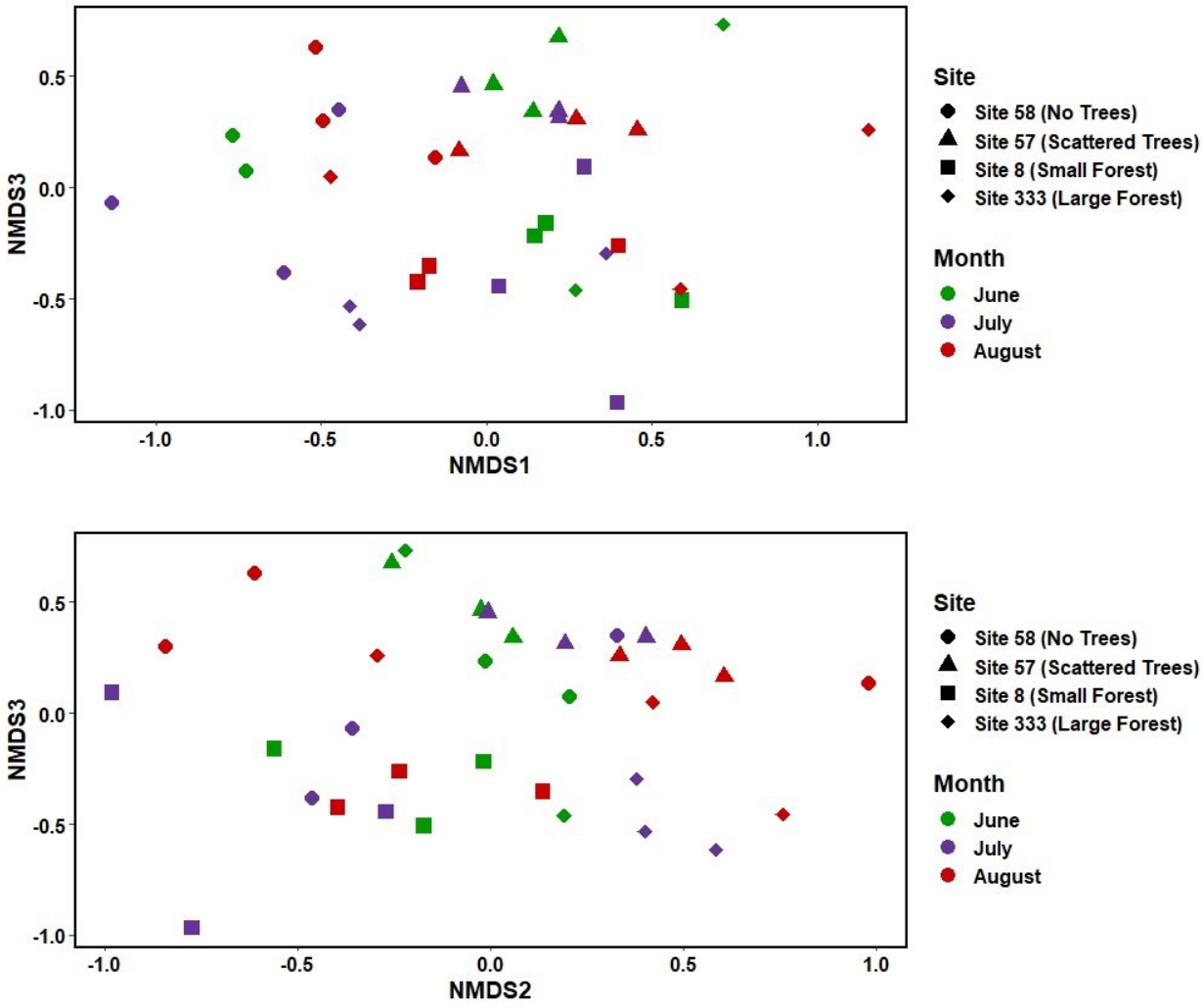


*Moderate to high correlations

Appendix 2. Scatterplot correlation graphs comparing soil physicochemical parameters with their associated correlation coefficients.



Appendix 3. Mesofaunal community composition for each habitat type with soil parameters modelled using non-metric multidimensional scaling (NMDS), showing NMDS 1 vs 3, and NMDS 2 vs 3 (stress value=0.19). Red = No trees: circles = Site 1, triangles = Site 21, squares = Site 58. Yellow = Scattered trees: circles = Site 53, triangles = Site 7, and squares = Site 57. Blue = Small forest: circles = Site 115, triangles = Site 8, and squares = Site 334. Green = Large forest: circles = Site 333, triangles = Site 343, and squares = Site 347.



Appendix 4. Community composition for the seasonal samples modelled using non-metric multidimensional scaling (NMDS), showing NMDS 1 vs 3, and NMDS 2 vs 3 (stress value=0.14).