The influence of agricultural tillage practices on soil biodiversity

by

Jeff H.T. Strohm

A Thesis

Presented to

The University of Guelph

In partial fulfillment of requirements

for the degree of

Master of Science

in

Integrative Biology

Guelph, Ontario, Canada

© Jeff H.T. Strohm, April 2015
ABSTRACT

THE INFLUENCE OF AGRICULTURAL TILLAGE PRACTICES ON SOIL BIODIVERSITY

Jeff H.T. Strohm
The University of Guelph, 2015

Advisors:
Dr. Robert Hanner and Dr. Richard Heck

This was the first study to examine the influence of different tillage practices on soil biodiversity from all kingdoms of life. Four tillage treatments were compared with a no-till treatment maintained at the Elora Agricultural Research Station for 30 years. In terms of species or operational taxonomic unit (OTU) richness within taxonomic higher ranks, these treatments are quite similar. However, there were significant differences in the OTU richness of the Sordariomycetes (103±19 vs 86±10 OTUs; p= 3.12e-3) and the Dothideomycetes (29±6 vs 43±1 OTUs; p= 5.54e-5) fungal classes when comparing tilled and no-till soils. While all agricultural treatments were similar in terms of arthropod diversity, when compared with a nearby mature woodlot, compaction sensitive lineages of Collembola and mites were not present in the agricultural soils. These results suggest that even no-till soils are not comparable to the soil environment of an un-disturbed forest.
ACKNOWLEDGMENTS

Given that this was the Hanner lab's first foray into metagenomics, soil science, microbiology and next generation sequencing, I required the help of many people within and outside of our lab. I extend thanks on a broad scale to the other graduate students of the Hanner lab with special thanks given to my two undergraduate research assistants: Jennifer Gubala and Danielle Bourque. Jenny performed many of my microbial DNA soil extractions while Danielle spent months sorting and identifying soil arthropods. I would also like to thank Jeff Gross, Jing Zhang and Angela Hollis from the Advanced Analysis Center - Genomics Facility, University of Guelph, for their information and guidance regarding my Illumina MiSeq runs. In addition, I owe a great deal of thanks to Natalia Ivanova, Sean Prosser and Shadi Shokralla from the Centre for Biodiversity Genomics, University of Guelph, for their advice regarding sequence library prep. I must thank the creators of the open source software packages that I used, for their patient and rapid responses to my email: Robert Edgar from the UPARSE package, Daniel Huson from MEGAN, Maarja Opik from MaarjAM and Donovan Parks from STAMP. Finally, I extend a great deal of thanks to my advisory committee.

TABLE OF CONTENTS

ABSTRACT........................................................................................................................................... 2
ACKNOWLEDGMENTS................................................................................................................... iii
TABLE OF CONTENTS................................................................................................................... iii
TABLES................................................................................................................................................ v
FIGURES............................................................................................................................................. vi
ACRONYM LIST............................................................................................................................... xiii
CHAPTER I: INTRODUCTION AND REVIEWED LITERATURE .................................................... 1
1.1 TILLAGE IN AGRICULTURAL SYSTEMS................................................................. 2
1.2 FUNCTIONAL ECOLOGY OF AGROECOSYSTEMS.............................................. 4
1.3 SOIL STRUCTURE AND BIODIVERSITY............................................................. 9
1.4 QUANTIFYING THE MICRO/MACRO FAUNAL BIODIVERSITY OF SOILS WITH NEXT GENERATION DNA SEQUENCING ................................................................. 12
1.5 THE INFLUENCE OF TILLAGE PRACTICES ON SOIL BIODIVERSITY FROM A METAGENOMIC PERSPECTIVE ................................................................. 15
1.6 RATIONALE AND OBJECTIVES ........................................................................... 17
1.7 HYPOTHESES AND PREDICTIONS ................................................................... 17

CHAPTER II SOIL MICROBIAL COMMUNITIES UNDER DIFFERENT TILLAGE PRACTICES ........................................................................................................................... 18

2.1 STUDY SITES ......................................................................................................... 18
2.2 MICROBIAL SAMPLING METHODOLOGY ............................................................ 24
2.3 SOIL SUB-SAMPLING AND HOMOGENIZATION .................................................. 25
2.4 DNA EXTRACTION ............................................................................................... 26
2.5 MICROBIAL TAXON SPECIFIC PCR .................................................................... 26
2.6 PCR REACTION OPTIMIZATION .......................................................................... 27
2.7 MICROBIAL PCR REACTION SETUP .................................................................... 28
2.8 POOLED PCR PRODUCT PURIFICATION ............................................................. 30
2.9 INDEX PCR FOR MULTIPLEX SEQUENCING ...................................................... 30
2.10 INDEX PCR PURIFICATION ................................................................................. 33
2.11 DNA LIBRARY QUANTIFICATION AND NORMALIZATION ................................. 33
2.12 DNA SEQUENCE ANALYSIS ............................................................................... 34
2.13 BIOINFORMATIC PIPELINE ................................................................................. 34
2.14 SEQUENCE ANALYSIS OF INDIVIDUAL REPLICATES ...................................... 37
2.15 SOIL MICROBIAL DIVERSITY RESULTS- SEQUENCING SUCCESS .................. 40
2.17 SOIL MICROBIAL DIVERSITY RESULTS ........................................................... 41
2.18 MICROBIAL DIVERSITY FIGURES ...................................................................... 44
2.19 MICROBIAL DIVERSITY DISCUSSION ................................................................ 69
2.20 CONCLUSIONS AND RECOMMENDATIONS ..................................................... 78
CHAPTER III: SOIL ARTHROPOD DIVERSITY UNDER DIFFERENT TILLAGE PRACTICES ................................................................. 79
3.1 SOIL SAMPLING .................................................................................................................................................. 79
3.2 SOIL SAMPLE PROCESSING ................................................................................................................................. 80
3.3 PHYSICAL SORTING OF ARTHROPOD SPECIMENS ...................................................................................... 81
3.4 DNA EXTRACTION AND PURIFICATION ........................................................................................................... 82
3.5 PCR INHIBITOR REMOVAL ................................................................................................................................ 83
3.6 TAXON SPECIFIC PCR ...................................................................................................................................... 83
3.7 INDEX PCR ............................................................................................................................................................ 84
3.8 BIOINFORMATIC PIPELINE ................................................................................................................................ 85
3.9 ARTHROPOD RESULTS ..................................................................................................................................... 86
3.10 ARTHROPOD FIGURES ..................................................................................................................................... 88
3.11 ARTHROPOD DISCUSSION ................................................................................................................................. 98
3.12 CONCLUSIONS AND RECOMMENDATIONS ................................................................................................. 101
4. FINAL SUMMARY AND CONCLUSIONS ............................................................................................................ 102
REFERENCES ............................................................................................................................................................. 103

TABLES

Table 2.1. Taxon specific PCR primers chosen to amplify microbial DNA from bulk soil extracts

Table 2.2. Three PCR annealing temperatures chosen for each primer set based on annealing temperature gradient experiments.

Table 2.3. Number of sequence reads obtained for each of the four Loci at progressing stages of the UPARSE pipeline. Results for ITS extraction using ITSx are also shown. In both runs 3 and 4, two replicates from all 5 treatments were sequenced for a total of 10 samples per run.

Table 2.4. Diversity metrics computed in UPARSE for pooled OTU files from each treatment and gene. N represents the total number of reads that contributed to the OTUs.

Table 2.5. Phylogenetic diversity of 16S OTUs from pooled OTU files from each treatment

v
Table 2.6. Summary of all taxonomic identifications made for the 16S amplicons along with p-values from comparing OTU richness between treatment ANOVAs and Eta-squared effect sizes. Significant differences are boxed, see Figure 2.13 for post-hoc results.

Table 2.7. Summary of all taxonomic identifications made for the 18S amplicons along with p-values from comparing OTU richness between treatment ANOVAs and Eta-squared effect sizes. Significant differences across all treatments are boxed, post-hoc results yielded no significant differences.

Table 2.8. Summary of all taxonomic identifications made for the ITS2 amplicons along with p-values from comparing OTU richness between treatment ANOVAs and Eta-squared effect sizes. Significant differences across all treatments are boxed, see Figures 7,8,9 for post-hoc results.

Table 2.9. Summary of all Taxonomic identifications made for the 18S-AMF amplicons along with p-values from comparing OTU richness between treatment ANOVAs and Eta-squared effect sizes.

Table 3.1. Alpha-diversity metrics computed for COI OTUs in UPARSE. N dictates the number of quality filtered sequences that contributed to the OTUs.

Table 3.2. Summary of all Taxonomic identifications at the level of order or higher made for COI amplicons along with p-values from between treatment ANOVAs and Eta-squared effect sizes comparing OTU richness within taxonomic ranks.

FIGURES

Figure 2.1. Plot layout design at the Elora Minimum Tillage Experimental Plots. Plot numbers are denoted with three digits. Treatments sampled for this study are denoted by 1 (no-till), 3
(spring moldboard plow, cultivate and pack), 5 (fall moldboard plow, cultivate and pack), 9 (fall chisel plow, cultivate and pack) and 10 (spring tandem disc, twice, no cultivator). Other numbers denote treatments that were not sampled in this study. Microbial samples were collected from plots marked with a red square. Preceding numbers indicate the number of within-plot replicates taken. All treatments were intended to have 4 replicates each, but there was a sampling error and as a result treatment 9 has 5 replicates while treatment 10 has three. All others have 4. Arthropod samples were collected from plots marked with a blue square.

Figure 2.2. Stylized depiction of original sampling design. White dots represent samples and two dots per plot denote a within-replicate sampling event. The loci that were sequenced for each sample (dot) are listed at the bottom.

Figure 2.3. Plow types used at the Elora Experimental Agriculture Plots which were sampled in this study. A) Moldboard plow which was used in the spring and fall moldboard treatments. B) Chisel plow which was used in the fall chisel treatment. C) Tandem disc plow which was used in the spring disc treatment. Images on the right depict the influence of these plows on soil and crop residue. Images provided by Mr. Henk Wichers, University of Guelph, Department of Plant Agriculture.

Figure 2.4. Stylized depiction of taxon specific PCR setup for both microbial and arthropod samples. Each tube represents a single soil sample.

Figure 2.5. Index PCR tests on 16S amplicons. The three lanes to the left tested the influence of the number of cycles. The lanes to the right tested different annealing temperatures in a 30 cycle program. On the far right, un-indexed amplicons were loaded into the 2% gel as a control. The right-most and left-most lanes are 100bp DNA ladder.
Figure 2.6. Stylized depiction of index PCR layout in a standard 96 well PCR plate (only half was used in this study). Each red arrow represents a unique forward primer while each green arrow represents a unique reverse primer.

Figure 2.7. Nucleotide sequences for each unique index primer used in this study. The 'N5' series are the forward primers, while the 'N7' series are reverse primers.

Figure 2.8. Taxonomic results for 18S amplicons from four identical DNA templates sequenced on two separate MiSeq runs at different loading concentrations (R1 = 12pM, R2 = 6pM). Significant differences in OTU richness within a taxonomic rank are marked with an asterisk via the default settings for a two sample comparison in STAMP (two sided G-test w/Yates' + Fisher). All taxa with OTU richness too low to appear on this figure were discarded to make efficient use of space. There were no significant differences in these taxa. A) Spring moldboard B) Fall moldboard C) Spring moldboard D) Spring moldboard

Figure 2.9. Taxonomic results for 18S amplicons from eight within-plot replicates (blue or orange bars) sampled from four tillage treatment types. Significant differences in OTU richness within a taxonomic rank are marked with an asterisk via the default settings for a two sample comparison in STAMP (two sided G-test w/Yates' + Fisher). All taxa with OTU richness too low to appear on this figure were discarded to make efficient use of space. There were no significant differences in these taxa. A) No-till B) Spring disc C) Spring moldboard D) Fall chisel (fall moldboard omitted due to space limitations)

Figure 2.10. Taxonomic results for 16S amplicons from eight within plot replicates (blue or orange bars) sampled from four tillage treatment types. Significant differences in OTU richness within a taxonomic rank are marked with an asterisk via the default settings for a two sample comparison in STAMP (two sided G-test w/Yates' + Fisher). All taxa with OTU richness too low
to appear on this figure were discarded to make efficient use of space. There were no significant differences in these taxa. A) No-till B) Spring disc C) Spring moldboard D) Fall chisel (fall moldboard omitted due to space limitations)

Figure 2.11. Taxonomic results for ITS2 amplicons from eight within plot replicates (blue or orange bars) sampled from four tillage treatment types. Significant differences are marked with an asterisk via the default settings for a two sample comparison in STAMP (two sided G-test w/Yates' + Fisher). All taxa with OTU richness too low to appear on this figure were discarded to make efficient use of space. There were no significant differences in these taxa. A) No-till B) Spring disc C) Spring moldboard D) Fall chisel (fall moldboard omitted due to space limitations)

Figure 2.12. Principal component analysis results comparing taxonomic identifications and their OTU richness. Computed in STAMP using default settings. Comparisons were made using taxonomically identified OTUs from all treatment replicates from five tillage regimes. In the legend, the numbers in parentheses correspond to the number of replicates that were analyzed.

Figure 2.13. 16S taxonomic groups that yielded significant differences in OTU richness of OTUs (y-axis) between tillage treatments. ANOVAs followed by Games-Howell post-hoc tests were used to identify significant differences (p < 0.05) in the program STAMP. The p-values listed were generated from ANOVAs. OTUs were clustered at 99% similarity from 400bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures. The letters above boxes indicate the other box that they are significantly different from.

Figure 2.14. ITS2 taxonomic groups that yielded significant differences in OTU richness of OTUs (y-axis) between tillage treatments but no broad scale patterns were apparent. ANOVAs followed by Games-Howell post-hoc tests were used to identify significant differences (p < 0.05)
in the program STAMP. The p-values listed were generated from ANOVAs. OTUs were clustered at 99% similarity from 144bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures. The letters above boxes indicate the other box that they are significantly different from.

Figure 2.15. ITS2 taxonomic groups that yielded significant differences in OTU richness of OTUs (y-axis) between tillage treatments and demonstrated interesting patterns. ANOVAs followed by Games-Howell post-hoc tests were used to identify significant differences (p < 0.05) in the program STAMP. The p-values listed were generated from ANOVAs. OTUs were clustered at 99% similarity from 144bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures. The letters above boxes indicate the other box that they are significantly different from.

Figure 2.16. Detailed Games-Howell post-hoc test results for the Dothideomycetes and Sordariomycetes data generated in STAMP. All parameters are identical to those described in Figure 2.15.

Figure 2.17. Genera within Sordariomycetes that were detected in the different tillage treatments from ITS2 reads. This figure was generated from pooled replicates in MEGAN 5. The legend is sorted from most diverse to least diverse.

Figure 2.18. Genera within Dothideomycetes that were detected in the different tillage treatments from ITS2 reads. This figure was generated from pooled replicates in MEGAN 5. The legend is sorted from most diverse to least diverse.

Figure 2.19. Heat-map plot detailing the number of 16S sequences (OTUs) that were assigned to certain taxonomic groups using the SILVA 119 database and the SINA classifier. Treatment
replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.

Figure 2.20. Heat-map plot detailing the number of 18S sequences (OTUs) that were assigned to certain taxonomic groups using the SILVA 119 database and the SINA classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.

Figure 2.21. Heat-map plot detailing the number of ITS2 sequences (OTUs) that were assigned to certain taxonomic groups using the UNITE database and the BLASTn classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.

Figure 2.22. Heat-map plot detailing the number of 18S-AMF sequences (OTUs) that were assigned to certain taxonomic groups using a custom database and the BLASTn classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.

Figure 2.23. Rarefaction curves generated in UPARSE from the pooled sample libraries for each tillage treatment. The x-axis indicates the size (%) of the sub-sampled dataset while the y-axis indicates OTUs.

Figure 3.0 Stylized depiction of the Berlese funnel set-up that was used to coax soil arthropods out of bulk soil samples and into ethanol for storage and processing.

Figure 3.1. Principal component analysis results from STAMP acquired using default settings. Comparisons were made using taxonomically identified CO1 OTUs from all treatment replicates.
from five tillage regimes and a mature woodlot. In the legend, the numbers in parentheses correspond to the number of replicates that were analyzed. Plot A includes all taxonomic identifications that were made for all OTUs. Plot B only includes OTUs that were identified as either Collembola or Mites (Arachnida) both of which comprised the majority of these soil communities.

Figure 3.2. Relationship between bulk soil sample dry mass and the number of CO1 OTUs obtained from those samples. Samples marked in red were collected from the mature woodlot while samples marked in blue were collected from the tillage plots.

Figure 3.3. CO1 OTU richness of Collembola OTUs (y-axis) between tillage treatments and a mature woodlot. An ANOVA did not identify any significant differences (p > 0.05). OTUs were clustered at 99% similarity from 400bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures.

Figure 3.4. Raw specimen abundances for Collembola sorted from all soil samples and visually identified.

Figure 3.5. CO1 OTU richness of Mite (Arachnida) OTUs (y-axis) between tillage treatments and a mature woodlot. An ANOVA did not identify any significant differences (p > 0.05). OTUs were clustered at 99% similarity from 400bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures.

Figure 3.6. Raw specimen abundances for Mites (Arachnida) sorted from all soil samples and visually identified.
Figure 3.7. Heat-map plot detailing the number of CO1 sequences (OTUs) that were assigned to Collembola using the NCBI database (excluding environmental samples) and the BLASTn classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.

Figure 3.8. Heat-map plot detailing the number of CO1 sequences (OTUs) that were assigned to mites using the NCBI database (excluding environmental samples) and the BLASTn classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.

Figure 3.9. Rarefaction curves generated in UPARSE from the pooled sample libraries from three replicates for each tillage treatment and the mature woodlot.

ACRONYM LIST

OTU  Operational taxonomic unit
PCR  Polymerase chain reaction
N  Nitrogen
C  Carbon
S  Sulfur
SOM  Soil organic matter
UV  Ultraviolet radiation
AMF  Arbuscular mycorrhizal fungi
DNA  Deoxyribonucleic acid
NGS  Next generation DNA sequencing
UAN  Urea and ammonium nitrate
16S  16S rRNA gene
18S  18S rRNA gene
ITS2
Cytochrome c oxidase subunit 1 gene
Taq
Thermus aquaticus DNA polymerase
KAPA HiFi
High fidelity DNA polymerase manufactured by KAPA Biosystems
ddH2O
Double distilled molecular biology grade water
GTR
Generalized time reversible model
STAMP
Statistical analysis of microbial profiles
LCA
Least common ancestor
NCBI
The National Center for Biotechnology Information
EC2
Elastic cloud computing offered by Amazon Web Services
ANOVA
Analysis of variance
qPCR
Quantitative polymerase chain reaction
DGGE
PCR–denaturing gradient gel electrophoresis
PLFA
Phospholipid fatty acid
BOLD
Barcode of Life Datasystems
CHAPTER I: INTRODUCTION AND REVIEWED LITERATURE

Soils harbour the most biodiverse microbial ecosystems on Earth (Delmont et al. 2011), yet our knowledge of this diversity and how it is influenced by environmental factors is very preliminary. This is in part due to the fact that the vast majority of soil microorganisms do not respond to "traditional" culturing techniques (Delmont et al. 2011). In the past this roadblock prevented us from studying the functional and phylogenetic diversity of these important ecosystems. Very recently, advances in next generation DNA sequencing have begun to illuminate this "black box" beneath our feet (Caporaso et al. 2010). We can now generate millions of DNA sequences from complex environmental samples such as soils and use them to examine the connections between soil biodiversity and the environment.

In the case of our research program, we are targeting agricultural soils. With 40% of the Earth's surface currently occupied by agricultural activities, we must find innovative management practices in order to reduce demands for more land (Foley et al. 2005). One such practice is known as conservation tillage, and while there are many varieties, the overarching philosophy is reducing or eliminating tillage and therefore soil disturbance, while maintaining a cover of plant material from the previous season (Coughenour and Chamala 2000). The ultimate goal of this practice is to reduce erosion and increase soil organic matter. Our study is the first to examine how different tillage practices, including conservation tillage, influence the entire soil community. With our next generation sequencing approach, we are targeting bacteria, archaea, fungi and protozoa. We are also isolating soil arthropods and sequencing their diversity with the same techniques. We have chosen phylogenetically informative markers to sequence, as they can inform us of the biodiversity under different tillage practices. We can then infer the ecological
roles of these organisms by using taxonomic identifications to link with studies that have examined their physiology and function.

This research will be novel in several ways. As previously stated it is the first to reconstruct the entire soil community and it is hoped that we will be able to detect interesting patterns that span all the kingdoms of life. We will also be generating an order of magnitude more sequences per sample than previous research in this field (Carbonetto et al. 2014; Eilers et al 2012; Opik et al. 2009). This will increase our ability to detect rare taxa. Finally we will be comparing no-till soils to four different tillage treatments which doubles the number of treatments examined when compared to previous work.

Great emphasis must be placed on the fact that these techniques are all extremely new and are subject to biases in the laboratory and analytical phases. Many of the software packages that were used here did not exist at the outset of this project. As such, previous research has warned of the biases that can be introduced in the analytical phase by the manner in which sequence data are processed (Schloss et al. 2011). Previous research also cautions that "universal" PCR primers, annealing temperatures and polymerase choices can all influence study results (Deagle et al. 2013). Therefore, for this "first look" in this system, every attempt will be made to rationalize study design choices and support these choices with published research. However, due to the novelty of this study and of this field, these results will most certainly require future investigation.

1.1 TILLAGE IN AGRICULTURAL SYSTEMS
Tillage in some form has been employed since the dawn of human agriculture. It serves to disrupt and overturn existing plant communities while loosening the soil and distributing organic matter throughout the upper soil horizon. This prepares the soil for planting. However, depending on the practices employed, tillage can degrade the soil environment to a point where it may no longer support cropping systems (Lal 1993). Important characteristics such as soil structure, resistance to wind and water erosion, cycles of water, nutrients and organic matter may be disrupted in a manner that is very difficult to rectify (Lal 1993). In North America, prior to the 1930's, removal of leftover plant material or residue (crop residue) from the previous crop and extensive plowing was the only method of preparing the soil for further cultivation (Coughenour and Chamala 2000). This subjected the soil to fierce wind and water erosion but farmers had no other means of planting within crop residue and controlling weeds. The production of herbicides in the 1940s allowed farmers to begin to experiment with reduced or no-till cropping systems (Coughenour and Chamala 2000). Not only did a reduction in tillage reduce soil degradation, it also saved farmers time and fossil fuel expenditure. This ushered in a new era of conservation cropping systems that employ management practices that leave a varying amount of crop residue between harvests and planting. The latest National Crop Residue Management survey of tillage practices in the U.S. found that 41.5% of planted crop acres employ conservation tillage (30% residue); a further 21.5% of these planted crop acres employ reduced till (15-30% residue) practices (CTIC 2008; www.ctic.purdue.edu/CRM/). This revolution in management techniques has resulted in cropping systems that are more resilient to degradation, resource efficient and productive (Holland 2004). However, due to the relative youth of these practices, much is unknown about their long-term influence on soil biodiversity and soil structure.
Large-scale cultivation of monoculture crops will quickly lead to a serious reduction in soil organic matter and other nutrients (Holland and Coleman 1987). Modern conventional systems remedy this by increasing the application of synthetic fertilizers. Given that these systems usually employ vigorous and deep tillage techniques, they possess poor soil structure and water retention is reduced (Franzluebbers 2002). The resulting runoff carries a large quantity of synthetic fertilizers into local water bodies where they lead to the decline in ecosystem health associated with eutrophication (Gregory et al. 2004).

1.2 FUNCTIONAL ECOLOGY OF AGROECOSYSTEMS

Soils are extremely biodiverse environments and may contain many orders of magnitude more species than any other terrestrial ecosystem (Thiele-Bruhn et al. 2012). Soil organisms have a fundamental role in nutrient cycles as they decompose organic matter, fix nitrogen and sequester carbon. The vigor to which these processes occur is largely dictated by the functional diversity of microbial taxa (Heemsbergen et al 2004).

Nitrogen is a fundamentally important nutrient for all ecosystems on Earth due to its presence in DNA, RNA and amino acids. Nitrogen can enter agricultural systems from anthropogenic inputs or it can be fixed from atmospheric N\textsubscript{2} into ammonium (NH\textsubscript{4}+) by N- fixing bacterial taxa that are associated with legumes such as Rhizobia (Miransari 2013). Additionally, N may be fixed by free-living colonies of bacteria and cyanobacteria (Berman-Frank et al. 2003). Reduced or no-tillage regimes influence N-fixation as crop residues are decomposed by certain bacteria and fungi on the soil surface, and they provide ammonia for the nitrifying bacteria. In fact, allowing residues to remain on the surface may lead to 100-fold increases in the abundance and 10-fold increases in the activity of N-fixing bacteria (Roper 1983). As organic matter is
decomposed by a variety of bacteria and fungi, the organic nitrogen is converted to \( \text{NH}_4^+ \) in a process called mineralization. Tillage regimes that leave organic matter to decompose on the surface as opposed to plowing it into the soil have higher quantities of mineralized nitrogen and greater N retention (Holland and Coleman 1987).

Nitrification is the next crucial step in the N cycle; it involves the conversion of ammonium to nitrite (\( \text{NO}_2^- \)) by bacteria in the genera *Nitrosomonas* and *Nitrosospira* and then nitrate (\( \text{NO}_3^- \)) via the bacterial genus *Nitrobacter* (Chu et al. 2007). Different species of these nitrifying bacteria have been associated with varying land use types, fertilizer applications and in turn, rates of nitrification (Carney et al 2004; Chu et al. 2007). Nitrification and mineralization are very important in agricultural and natural systems as they lead to the decomposition of organic matter back into fundamental N compounds for plant metabolism (Sooksa-nguan et al. 2009). Additionally, \( \text{NO}_3^- \) liberated via nitrification is highly water soluble and represents the principal vector for N leaching in terrestrial systems.

The final step in the soil nitrogen cycle is denitrification. This involves the conversion of \( \text{NO}_3^- \) to inert \( \text{N}_2 \) or nitric oxide (a powerful greenhouse gas) by bacteria in anaerobic conditions which use \( \text{NO}_3^- \) as a terminal electron acceptor for respiration (Wang et al. 2005). Bacterial species that have been linked with this process are *Pseudomonas aeruginosa*, *P. fluorescens*, *P. stutzeri*, *Paracoccus denitrificans* and *Thiobacillus denitrificans* (Carlson and Ingraham 1983; Thomas et al. 1994; Wang et al. 2005). Denitrification is of obvious interest as it represents a loss of plant-useable N from the system and it is dependent on soluble carbon, water for \( \text{NO}_3^- \) movement, temperature and porosity among other things (Luo et al. 1999).

Phosphorus is another essential plant nutrient with availability for plant absorption that is highly dependent on microbial activities. There is no P equivalent for N-fixation so the P within
a system remains constant unless it is added from an external source. This, coupled with its low solubility, means it is usually a growth-limiting nutrient in agricultural systems (Vandermeer 2010). Across a wide pH range, plant-available $\text{H}_2\text{PO}_4^-$ becomes inorganically bound with Al, Fe, Ca or Kaolinite clays and exists in a crystallized non-soluble form that is inaccessible to plants (McDowell and Sharpley 2003). However, many different bacterial and fungal taxa are capable of liberating it from this state; a recent review by Han et al. (2007) lists 22 taxa of fungi and actinomycetes in addition to 21 bacterial taxa that are capable of solubilizing inorganically bound P in soil by excreting various acids. Maintaining robust populations of these microflora is therefore of paramount interest in agricultural systems, where additions of P may be rendered insoluble before they can be utilized by plants.

Microbial taxa are also heavily involved in the decomposition of soil organic matter (SOM) to facilitate the liberation of carbon for plant metabolism. The soil carbon cycle is deceptively simple; it consists of CO$_2$ fixation by photosynthetic organisms or those that oxidize reduced inorganic compounds; this organic carbon is then decomposed by heterotrophic organisms, resulting in a release of CO$_2$ and completion of the cycle (Prosser 2007). The decomposition process and carbon cycle is complicated by a number of factors including the availability of carbon compounds, the mode of cellular degradation (i.e. internal versus external enzymes), links to other nutrient cycles like the N cycle, and the fact that some organic compounds are degraded by single taxa or large groups (Prosser 2007). Plant material comprises the bulk of decomposing matter in agricultural systems and approximately half of that matter is cellulose, making bacterial and fungal species with the ability to degrade it very important. Bacterial genera that have been implicated in cellulose digestion include: *Clostridium, Sorangium, Micromonospora*, *Cellulomonas, Thermonospora, Polyangium, Acetovibrio, Cytophaga, Cellovibrio, Bacteroides*.
Archangium and Sporocytophaga (Adl 2003). Lignin is another abundant plant polysaccharide but unlike cellulose it is almost entirely degraded by various fungi such as Trametes versicolor, Phlebia tremellosa, Phellinus pini and Scytinostroma galactinum working together in syntrophy to employ a diverse repertoire of enzymes (Blanchette 1995). From this review, it should already be apparent that the quantity of SOM in agricultural systems is highly dependent on tillage practices. There is a broad body of literature demonstrating that reduced/no tillage systems possess more SOM than conventional systems (Six et al 1999; Roger-Estrade et al. 2010; Kong and Six 2010). This trend has been linked with an increased stability of soil aggregates in reduced/no-tillage systems, which in turn fosters a protective environment for SOM retention and reduced decomposition rate (Bronick and Lal 2005; Ashagrie et al 2007). In addition, leaving crop residues on the soil surface as opposed to incorporation via tillage has been found to support 144% greater surface fungal biomass and a greater retention of C in the system (Holland and Coleman 1987).

Sulfur is yet another important nutrient for plant growth with an abundance that is largely dictated by microbial activities. S is required to make the amino acids methionine and cystine along with various other vitamins and organic compounds. The S cycle resembles the P cycle in that there is an alternation between bioavailable and unavailable crystalline forms driven by bacterial and fungal catalysis. Sulfates represent the main source of S for plants, but first they must be mineralized from organic compounds by aerobic bacteria and fungi (Davet 2004). Other inaccessible mineral forms of S can be oxidized to sulphates by very specialized bacteria such as Thiobacillus, Beggiatoa and some archaea (Davet 2004). Some oxidized mineral forms of sulfur must be reduced in order to become bioavailable, and this reaction only seems to be catalyzed by
a few bacteria such as those in the genera *Desulfovibrio* and *Desulfotomaculum* (Parshina et al 2010).

In the soil food web, the bacteria and fungi reviewed so far are classified as primary saprotrophs due to their ability to release enzymes into the environment for nutrient solubilization (Adl 2003). However, in order for this microbe-mediated decomposition to take place efficiently, coarse organic matter must be shredded by multicellular creatures such as insects, mites, annelids, diplopods and isopods. Detritivores such as nematode and oligochaete worms and various protozoa then process this fragmented material to an even finer level in a process collectively known as secondary decomposition (Adl 2003). Therefore, both mechanical and chemical processes are integral to nutrient cycling.

Secondary saprotrophs are a guild of creatures that feed on primary saprotrophs and the organic matter that has been processed by them. Various myxobacteria, nematodes, protozoa, collembolans and tardigrades feed selectively on bacterial taxa of their liking and can therefore control nutrient flow in agricultural systems in a top-down manner (Darbyshire 1994; Adl 2003). This grazing can liberate nutrients that were otherwise inaccessible for plant metabolism, and bacterial-grazing nematodes and amoebas have been measured to contribute 83% of mineralized N by soil fauna in grasslands (Verhoef and Brussaard 1990). Soil protozoa also excrete up to 60% of the N and P that they consume as excess, and these nutrients are readily available for plant uptake (Darbyshire 1994). In addition, many act as keystone species through their selective grazing of bacterial taxa. Protozoa seem to favour larger and dividing bacteria, single bacteria versus aggregates, and also select prey based upon chemical contents, colony size and whether the cells are adhered to a surface or in suspension (Caron 1987; Sibbald and Albright 1988). This selective grazing fosters biodiversity among bacterial taxa by enabling competing taxa to coexist.
The functional diversity among bacteria that emerges from this is crucial for nutrient cycling and decomposition (Griffiths 1994). Fungivores are also important for food web dynamics as they control the liberation of nutrients from woody material, and most are amoeboid and other protozoa, collembolans and nematodes (Adl 2003).

1.3 SOIL STRUCTURE AND BIODIVERSITY

Soil structure refers to the shape, size, continuity and arrangement of pores and solids (Bronick and Lal 2005). There are two perspectives from which to view soil structure: from the solid phase, or from the void phase (Chenu and Cosentino, 2011). The former describes interactions between soil particles of differing sizes and shapes that result in aggregate formation, while the latter views the soil environment as an interconnected network of pores that allow for the movement or accumulation of various substances such as fluids, organic and inorganic material and the penetration of roots. Macropores are often biological in origin, as various creatures tunnel their way through subterranean habitats. Macropores allow for rapid drainage, as water may bypass smaller pores that constitute the bulk of the soil void space. In addition, macropores influence the movement of various anthropogenic chemicals and the degree of leaching of organic and inorganic compounds (Flury et al. 1994; Gerrard, 2000).

Aggregation is another soil structural feature, and it is defined as “a naturally occurring cluster or group of particles in which the forces holding the particles together are much greater than the forces between the aggregates” (Martin et al. 1955). Aggregation depends highly upon the type and sizes of particles, the distribution and concentration of various organic compounds, and the climatic conditions (Gerrard, 2000). Microaggregates (<250 µm) form around organic molecules bound to clays and polyvalent cations such as Ca$^{2+}$ and Fe$^{3+}$ to form particles which
adhere together and create macroaggregates (>250 µm) (Tisdall 1982). Macroaggregates may also form around particulate organic matter, as primary saprotrophs release digestive compounds which increase macroaggregate stability and leads to the formation of microaggregates inside (Bronick and Lal 2005). The soil structure that emerges in turn influences the growth of plant roots and other organisms via water and oxygen conductivity and retention. Roots and fungal hyphae physically stabilize aggregate clusters and release compounds that have adhesive properties; the same is true for the polysaccharide exudates of bacterial colonies (Tisdall 1994). In terms of importance for plant growth and an influence on soil structure, the arbuscular mycorrhizal fungi (AMF) are of extreme interest. These fungi exist in a symbiotic relationship with 80-90% of all plant species, and their colonies penetrate plant roots and the soil simultaneously (Wilson et al. 2009). AMF aid plants in extracting nutrients such as P, N and S from the soil in exchange for sugars, and the rate of N transfer is directly linked to C supplied by the plant (Fellbaum et al. 2012). AMF also modify soil structure by promoting aggregate formation and thus organic matter and water stability with their fungal hyphae, cohesive exudates and a protein (or group) called glomalin (Miller and Jastrow 1990; Rilling 2004). Soil algae and cyanobacteria also have positive effects on soil structure by increasing aggregate stability and resistance to erosion (Eldridge and Leys 2003). They also have the ability to fix N and are among the first organisms to colonize new soils (Starks et al. 1981). The diversity of cyanobacteria has also been demonstrated to vary with different land uses and has thus been suggested as an indicator group for ecosystem degradation (Zancan et al. 2006).

Tillage can cause rapid nutrient turnover, reduced water and oxygen availability, compaction, aggregate disruption, increased desiccation and exposure to UV (Franzluebbers 2002; Pagliai et al. 2004). Despite the fact that tillage is used in part to increase soil porosity for
root penetration, increases in porosity can be quite short-term as tilled soil is less structurally stable due to the reduction in organic matter and biological activity associated with it (Gregory et al. 2007; Liiri et al. 2012). Additionally, the reduction in soil biodiversity associated with conventionally tilled systems leads to a reduction in biopores created by various invertebrates such as earthworms, ants and beetle grubs as they construct tunnels and chambers (Pagliai et al. 2004). These biopores are of importance to plant growth because they have a high degree of connectivity and allow for easy root penetration and water movement; damage to soil structure can be recognized by a reduction of this pore type (Pagliai et al. 2004). Unlike these large macrofauna, microfauna such as nematodes, springtails, tardigrades and mites have a negligible influence on soil structure as they use existing pore networks for dispersal (Collins and Qualset 1999). Bacterial colonies are also dependent on the pore networks of the soil, although, due to their incredibly small size, their reliance on pores for water and gases such as oxygen far surpass their need for habitable space. Since no-till soils are typically wetter, cooler and denser, they have been found to support higher microbial biomass and diversity than tilled systems in the upper 30cm (Young and Ritz 2000; Welbaum et al. 2004). As a result, conventionally-tilled soils have a reduced rate of nutrient cycling and decomposition (Roger-Estrade et al. 2010). The mechanical disturbance associated with tillage can obviously harm soil organisms as it rips apart fungal colonies and destroys the tunnels of worms and ants. The impact of this mechanical stress on bacteria is less clear; if the aggregates that house the bacteria remain intact and in a soil horizon with similar gases and nutrients, they may remain unharmed (Young and Ritz 2000). In conclusion, soil organisms support the existence of beneficial soil structure, water availability, nutrient cycling; therefore, soil organisms are of great interest for the restoration of exploited systems and the maintenance of resource-efficient cropping systems.
1.4 QUANTIFYING THE MICRO/MACRO FAUNAL BIODIVERSITY OF SOILS WITH NEXT GENERATION DNA SEQUENCING

In the early days of soil microbiology and biodiversity, culture plates were inoculated with soil microbes, incubated and colonies were then counted and morphologically or biochemically examined to identify the taxa that were present (Galvez et al. 1998). This led to a phenomenon known as the “great plate count anomaly” as it is estimated that 1% of wild bacterial species will grow in the nutrients, temperatures, physical properties and oxygen concentration of agar plates (Stanley and Konopka 1985). This 1% only represents about 8 bacterial genera which are easy to grow in situ and can thus be called “weeds” of the microbial world (Hugenholtz 2002). This serious issue was not remedied until the mid-1980s when Norman Pace proposed that bulk DNA could be extracted from a soil sample and a gene of interest with enough interspecific diversity and little intraspecific diversity (such as the 16S rRNA gene) could be amplified with PCR primers and then cloned into artificial plasmids (Pace et al. 1985). Despite this huge advance in methodology, biodiversity surveys using the bacterial and archaean 16S gene may still be confounded by horizontal gene transfer between these organisms, despite the argument that it rarely occurs for rRNA due to their core role in metabolic processes (Yap et al. 1999; Zhaxybayeva et al. 2006). These advances lead to a new field of study collectively known as metagenomics, and these methods were employed until a revolution in DNA sequencing technology. Traditional Sanger DNA sequencing works well on PCR product amplified from one specimen, but fails when multiple species are present in the sample. Given the hyperdiversity of soils, Sanger sequencing is a hopeless choice for obtaining sequence data for individual species. Next-generation DNA sequencing (NGS) provides a massively
parallel alternative to Sanger sequencing; there are several NGS varieties, but the Illumina MiSeq platform will be the version discussed here. It is the machine used for this research due to its ability to generate large quantities of high quality, 500bp sequences which are of sufficient length to examine biodiversity from an amplicon sequencing perspective (Saliplante et al. 2014) (for a detailed review of other next generation platforms see Shokralla et al. 2012).

In order to begin a metagenomics project on the MiSeq, bulk DNA from soil samples is extracted using a commercially available kit that can remove various PCR inhibiting substances that exist in soil such as humic acids. A gene region of interest is then amplified with universal PCR primers for a broad taxonomic group such as bacteria or fungi. However, these primers differ from the standard ones used throughout molecular biology as they contain a sequence tail of DNA that is complimentary to universal Illumina adapter primers. Once the gene region of interest is amplified using these modified PCR primers, the products are checked on an agarose gel to ensure they are of the desired length and quality. Then a second PCR is performed with primers that contain the other half of the Illumina adapter sequences in addition to an index region. The Illumina adapter sequences are identical for every PCR amplicon and are complimentary to oligonucleotides (oligos) bound to the sequencing flow cell on the inside of the MiSeq machine. This technology is similar to a microarray but all oligos are identical and bind the adapter regions in the tails of the amplicon library. The index region that is added to the amplicons during index PCR possesses a unique DNA sequence for every primer, and allows the user to tag all amplicons from a given soil sample with this same sequence. Once digital sequencing reads are obtained from the MiSeq, bioinformatic software is used to sort the sequences based on the soil sample that they were extracted from. This allows researchers to sequence DNA from multiple soil samples in one MiSeq run, resulting in significant monetary
and time savings. Once the final set of indexed amplicons is obtained, they are pooled together and loaded into the MiSeq. The Illumina adapter sequences on the amplicons hybridize with their complements on the flow cell and a polymerase is added that generates millions of amplicon copies or clusters. In the sequencing phase, standard sequencing primers are washed over the flow cell, which hybridize with the millions of new DNA clusters. Then, DNA polymerases are added in addition to fluorescently labeled nucleotides that possess a chemically inactivated 3’-OH (Shokralla et al. 2012). Therefore, extension of the amplicons within clusters takes place one nucleotide at a time. Following each base addition, they are excited, fluoresce and this fluorescence is detected in a base-specific manner. Digital sequence files are constructed from the consecutive images of fluorescent nucleotides along the entire DNA molecule. The MiSeq platform is currently capable of generating 250bp paired end reads which amount to 500bp of sequencing capability per amplicon. Amplicons are sequenced 250bp in from the 5’ or 3’ end, and the overlapping region is used to merge reads from separate forward and reverse files, into contiguous reads with downstream software such as UPARSE (Edgar 2013). The MiSeq is capable of generating over 15 million of these sequence reads in a span of a few days. This is why next generation sequencing has changed the manner in which metagenomic research is conducted. It allows for the reconstruction of entire environmental communities without ever observing the organisms themselves (pending marker gene amplification with universal primers) and without having to clone their DNA into plasmid vectors; which is a complicated procedure and may lead to biases and a reduction in species detection. Therefore, metagenomics has ushered in a new paradigm in the study of biodiversity as the majority of taxa will only be known from a DNA sequence that is never matched to an observed organism. In order to catalogue this unseen diversity, which often has a murky taxonomic framework at the species level, the
molecular operational taxonomic unit (MOTU or often OTU in the microbiology literature) is used to characterize biodiversity by clustering all sequences together above a common sequence similarity threshold and treating these as species in downstream analyses (Blaxter and Floyd 2003; Edgar 2013).

1.5 THE INFLUENCE OF TILLAGE PRACTICES ON SOIL BIODIVERSITY FROM A METAGENOMIC PERSPECTIVE

Thus far, many of the publications generated in this field are based on study sites in South America. One study used a shotgun sequencing approach to compare the microbial diversities of conventionally tilled soils versus no-till in Southern Brazil (Souza et al. 2013). They found that tilled soils contained a more abundant and diverse community of bacteria responsible for organic matter decomposition and nutrient cycling of nitrogen and carbon. They also found that tilled soils had a more abundant and diverse community of Eukaryotes, with the majority being Fungi. They do add the caveat that 46% of all reads were not taxonomically identified, and that no-till soils contained a higher proportion of these. It should be noted that while they collected 4 treatment replicates to a depth of 10cm, these were homogenized prior to DNA extraction so a comparison of between replicate variability is impossible. In a follow-up study using the same research site, the metabolic profiles of tilled versus no-till soils (also to a depth of 10cm) were studied using shotgun sequencing (Souza et al. 2015). Their results suggest that tilled soils also contain a more diverse repertoire of microbial genes involved in carbohydrate metabolism. This study had a much lower level of gene sequence diversity compared with their previous study that examined taxonomic diversity. They suggest that this is evidence for a high level of functional redundancy. In both studies they re-enforce the notion
that, despite tilled soils containing a more diverse community, tilled soils have been demonstrated to lose organic matter and nutrients with time (Franchini et al. 2007; Hungria et al. 2009). They suggest that the more diverse taxonomic and metabolic profiles of tilled soils in their study have been selected via environmental filtering to process a wider range of nutrient sources due to their gradual depletion from erosion. Conversely, another study from Brazil that used a 16S Illumina sequencing approach, found that no-till soils contained significantly more rich communities at the taxonomic levels of order, family, genus and species in several lineages (De Quadros et al. 2012).

A recent study from Argentinean long-term tillage experimental plots yielded contradictory results from the above studies (Carbonetto et al. 2014). No-till soils were found to have a significantly higher diversity of nitrogen cycling *Nitrospirae* OTUs when compared with tilled soils. No-till soils also supported a higher diversity of the anaerobic, syntrophic *Syntrophobacterales*. This study also examined metabolic profiles between treatments but could not find any significant differences. While this study did analyze 6 replicates per treatment separately instead of pooling them, they "air dried" their samples for an unspecified amount of time prior to being frozen. This may confound comparisons to the aforementioned studies from Brazil. In addition, sampling depth was not specified.

Instead of using a DNA sequencing approach, Lupwayi et al. (2012) examined the functional capabilities of microbial communities of the upper 7.5cm in tilled versus no-till Canadian prairie soils using BioLog Eco® plates which measure metabolism of carbon compounds. Their results indicate that no-till soils contain both a greater microbial biomass and a higher level of functional diversity, although many soil microbes may not respond to the Biolog culture environment.
Another recent study from Mexico seems to support the findings of Souza et al. (2013) (Navarro-Noya et al. 2013). The well studied notion that no-till systems contain higher organic carbon concentrations was supported. Additionally, they found that tilled soils harboured a greater average number of bacterial OTUs compared to no-till in the upper 20cm that were analyzed. However, from an OTU or phylogenetic diversity perspective there were no significant differences related to treatment type. The notion that tilled soils have a more taxonomically and metabolically diverse microbiome was further supported by a recent study in Laos (Lienhard et al. 2013). These conclusions were based on respiration measurements following wheat residue addition and 454 pyrosequencing of 16S and 18S amplicons.

### 1.6 RATIONALE AND OBJECTIVES

The influence of tillage practices on soil biodiversity is an extensively studied area of research. However, the gap in our understanding lies in the diversity of the entire soil community. To our knowledge there is yet to be a study that examines the diversity of bacteria, fungi, algae, nematodes, protozoa, arthropods and other multicellular fauna in the context of different agricultural management regimes. Therefore our proposed research aims to reconstruct the entire soil community of these agroecosystems as interesting patterns could be missed by only focusing on certain taxonomic groups. In addition, we will be examining more than one form of tillage treatment which will also yield novel results.

### 1.7 HYPOTHESES AND PREDICTIONS
H1. Agricultural fields under no-till have an increased retention of organic matter, pore
connectivity and decreased physical disturbance compared to tilled fields. These environmental
characteristics are more congenial for a more diverse community due differential requirements
among taxa.

P1. No-till fields harbor a more diverse community of bacteria, protozoa, arthropods and fungi
and high pore connectivity.

H2. Microbial diversity is fostered by the grazing of secondary saprotrophs.

P2. Fields with higher microbial diversity will also contain a high diversity of nematodes and
protozoa that graze on bacteria (causation in this case will require manipulation in future
experiments).

H3. Tillage is traumatic for microorganisms due to shearing forces and other physical
disturbances and there is differential sensitivity to disturbances among taxa.

P3. No-till soils will harbor a greater microbial diversity.

H4. Tillage disturbances are harmful for soil arthropods due to shearing forces and other
physical disturbances and there is differential sensitivity to disturbances among taxa.

P4. No-till soils will contain a greater richness of arthropods when compared to tilled soils.

CHAPTER II SOIL MICROBIAL COMMUNITIES UNDER DIFFERENT
TILLAGE PRACTICES

2.1 STUDY SITES
The experimental agricultural plots operated by the University of Guelph in Elora, Ontario provide a unique environment to study the influence of tillage on biodiversity. For the past 30 years they have maintained certain plots on strict regiments which either lack tillage entirely or have been subjected some common tillage practices. The treatments that will be examined in this study are: no-till, spring moldboard plow with cultivation and pack, fall moldboard plow with cultivation and pack, fall chisel plow with cultivation and pack or spring tandem disc (twice) without a cultivator (Figure 2.1). The direction of tillage is reversed between even and odd years. The plots measure 6.1m in width by 12.2m in length.
Figure 2.1. Plot layout design at the Elora Minimum Tillage Experimental Plots. Plot numbers are denoted with three digits. Treatments sampled for this study are denoted by 1 (no-till), 3 (spring moldboard plow, cultivate and pack), 5 (fall moldboard plow, cultivate and pack), 9 (fall
chisel plow, cultivate and pack) and 10 (spring tandem disc, twice, no cultivator). Other numbers denote treatments that were not sampled in this study. Microbial samples were collected from plots marked with a red square. Preceding numbers indicate the number of within-plot replicates taken. All treatments were intended to have 4 replicates each, but there was a sampling error and as a result treatment 9 has 5 replicates while treatment 10 has three. All others have 4. Arthropod samples were collected from plots marked with a blue square.

<table>
<thead>
<tr>
<th>Tillage Type</th>
<th>Microbes</th>
<th>Arthropods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year 1</td>
<td>Year 2</td>
</tr>
<tr>
<td></td>
<td>Collected in a single sampling event</td>
<td>Collected in a single sampling event</td>
</tr>
<tr>
<td>Fall Chisel</td>
<td>![Images]</td>
<td>![Images]</td>
</tr>
<tr>
<td>Fall Moldboard</td>
<td>![Images]</td>
<td>![Images]</td>
</tr>
<tr>
<td>No Till</td>
<td>![Images]</td>
<td>![Images]</td>
</tr>
<tr>
<td>Spring Tandem Disc</td>
<td>![Images]</td>
<td>![Images]</td>
</tr>
<tr>
<td>Spring Moldboard</td>
<td>![Images]</td>
<td>![Images]</td>
</tr>
<tr>
<td>Woodlot</td>
<td>![Images]</td>
<td>![Images]</td>
</tr>
</tbody>
</table>

**Loci**
- **16S** - Bacteria/ Archaea
- **18S** - Eukaryotes (Protozoa)
- **ITS2** - Fungi
- **18S** - Mycorrhizal Fungi
- **CO1** - Arthropods
Figure 2.2. Stylized depiction of original sampling design. White dots represent samples and two dots per plot denote a within-replicate sampling event. The loci that were sequenced for each sample (dot) are listed at the bottom.

The plots are planted with a corn-soy annual rotation. For years in which corn is planted, 78,000 seeds/ha of Pioneer 38B14 corn is fertilized with 43L/ha of 6-24-6 (N-P-K) liquid fertilizer, 157kg/ha of 5-20-20 solid fertilizer and a side dressing of 420L/ha of UAN (urea and ammonium nitrate). The herbicide Callisto is applied at 0.3L/ha along with 3.5L/ha of Primextra2Magnum. Corn stalks are collected as silage after harvest. For the years in which soybeans are planted, fertilizer (0-0-50 potassium sulfate) is applied at a concentration of 200 lbs per acre prior to planting. A Kinze planter is then used to sow S05-B3 soybeans in 15” rows. The herbicide Lorox is applied at 2.4L/ha along with the herbicide Dual 11 Magnum at 1.75 L/ha.

Samples were collected from five tillage treatments. Mr. Henck Wichers maintains these plots and provided the following details regarding their tillage treatments (Scaiff 2011). In the no-till plots, seeds are planted directly into the crop residue from the previous season and no tillage implements of any kind are used. There are two moldboard plow treatments with the first being spring moldboard. The concave shape of the moldboard plow results in an inversion of the soil down to a depth of 18cm and little crop residue is left at the surface (Figure 2.3A). The second moldboard treatment is performed in the fall and employs secondary tillage in the form of a soil cultivator and packer. This is done in the spring, prior to planting as moldboard plowing is very intense and leads to a loose and uneven seedbed. The third tillage treatment studied here is the use of a chisel plow in the fall (Figure 2.3B). It also disturbs the soil down to a depth of 18cm, but does so by dragging shanks through the entire plow layer, but more crop residue is left at the surface compared to moldboard plowing. The final tillage treatment is
tandem disc plowing in the spring, which inverts the soil in a similar manner to the moldboard plow, but to a much shallower depth of 6cm (Figure 2.3C).
Figure 2.3. Plow types used at the Elora Experimental Agriculture Plots which were sampled in this study. A) Moldboard plow which was used in the spring and fall moldboard treatments. B) Chisel plow which was used in the fall chisel treatment. C) Tandem disc plow which was used in the spring disc treatment. Images on the right depict the influence of these plows on soil and crop residue. Images provided by Mr. Henk Wichers, University of Guelph, Department of Plant Agriculture.

2.2 MICROBIAL SAMPLING METHODOLOGY

All tillage treatment samples were collected on September 26, 2012 from the minimum tillage experimental plots at the Elora Agricultural Research Station, Ontario, Canada (43°38’38.0”N 80°24’20.5”W). The soil type in these plots is of the Luvisolic order under the Canadian soil classification system. Samples were collected from three replicate plots for each treatment. Within the third replicate plot, two samples were collected to examine within plot variability (Figure 2.2). For the between plot replicates, samples were collected 2m up the third planting row from the left plot border. For the within-replicate plots, an additional sample was collected 2m up a row from the first sample. All samples were taken 10cm to right of the center line between planting rows to avoid the small trench made by the fertilizer applicator.

Two sterile 15mL Falcon tubes were used for collecting each sample. The first tube was hammered into the soil until fully submerged. It was then removed by digging down next to it with a trowel. A small wall of soil was left between the trowel and the sample tube hole to avoid contamination between sites. The first tube was removed and capped. The second tube was then inserted into the hole made by the first tube, and the process was repeated. This dual tube method
was used in case a blockage prevented an even sample being collected across the plow layer, as microbial communities exhibit stratification by depth (Eilers et al. 2012). This method allowed the upper 15cm of the plough layer to be sampled. All samples were isolated in plastic bags to avoid cross contamination and placed in a cooler kept in the shade. Samples were frozen at -80°C until DNA extraction.

2.3 SOIL SUB-SAMPLING AND HOMOGENIZATION

Sub-sampling was performed in a fume hood as it allows for aerosols to be drawn out of the workspace and keeps solid contaminants contained for easy cleanup. In addition, it provides isolation from the cabinets used for PCR reaction set-up. The fume hood was first sterilized with 70% ethanol. Eliminase was then applied to all surfaces and rinsed twice with ddH₂O. Soil sample tubes were thawed at 4°C. Scoopulas were cleaned with detergent, followed by 70% ethanol, then Eliminase and finally autoclaved. One large weigh boat and one sterile reagent reservoir were UV sterilized for every sample for 15 minutes.

Each sample was sub-sampled separately (including within-plot replicates). Nitrile gloves and a mask were always worn. All sub-sampling was performed on a 3 ply Kimwipe to catch large contaminating spills. Samples tubes A and B were opened and a scoopula was used to scrape the soil into the reagent reservoir. Once all soil from both tubes was in the reservoir, the same scoopula was used to thoroughly mix the sample by breaking up all of the soil aggregates. Manual homogenization was chosen as opposed to mechanical, with the intention of reducing cross contamination between samples, as scoopulas are far easier to sterilize for each sample, compared with mechanized implements.
9 grams of homogenized soil was then placed in a PowerBead tube from the MO BIO PowerSoil extraction kit and frozen for DNA extraction, which took place within 24hrs. Nitrile gloves were changed, all disposable equipment was sealed in a bag and the work surface was sterilized before proceeding to the next sample.

2.4 DNA EXTRACTION

The MO BIO PowerMax Soil extraction kit was chosen as it can remove PCR inhibitors such as phenols, humic acids and a variety of other organic molecules from the DNA extract. In addition, this kit allows for extraction from up to 10g of soil (many other kits only allow processing of 0.25g). DNA extraction was performed in the fume hood using the same sterilization protocols as were used for sub-sampling. The manufacturer's extraction protocols were strictly adhered to. DNA extract was stored at -80°C until PCR.

2.5 MICROBIAL TAXON SPECIFIC PCR

The following primer sets were selected from the literature:

Table 2.1. Taxon specific PCR primers chosen to amplify microbial DNA from bulk soil extracts

<table>
<thead>
<tr>
<th>Target</th>
<th>Sequence</th>
<th>Length (bp)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prokaryote 16S V3-V4 rDNA</td>
<td>F 5’- ACTCCTACGGGACGAGCAG R 5’- TACNVGGGTATCTAATCC</td>
<td>410</td>
<td>Claesson et al. (2010)</td>
</tr>
<tr>
<td>Arbuscular Mycoryzal Fungi 18S SSU rDNA NS31/AM1 region</td>
<td>F 5’- TTGGAGGGCAAGTCTGGTGCC R 5’- GTTTCCCGTAAGGCGCCGAA</td>
<td>550</td>
<td>Opik et al. (2009)</td>
</tr>
<tr>
<td>Universal Fungi ITS2</td>
<td>F 5’- TCCTCCGCTTTATTGATATGC R 5’- GGAAGTAAAAAGTCGAACAGG</td>
<td>Variable from 490-640</td>
<td>Schoch et al. (2012)</td>
</tr>
</tbody>
</table>
All primers were ordered from Integrated DNA Technologies with Illumina adapter tails F 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG and R 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG. Primers were re-suspended in 10mM Tris-HCL (pH 7.6).

2.6 PCR REACTION OPTIMIZATION

The high fidelity polymerase, KAPA HiFi HotStart in ReadyMix format was chosen because, according to the manufacturer, it incorporates 100X fewer erroneous bases when compared with wildtype Taq. An independent assessment by Quail et al. (2012) determined it to be the ideal polymerase for NGS library preparation because, it has fewer G+C, A+T and primer biases compared to its competitors.

Annealing temperature gradients were performed for each primer set and were used to select "low, medium and high" annealing temperatures to be used for all samples as recommended by Schmidt et al. (2013). The 16S primers used to target prokaryotes were chosen based on a comparative study by Claesson et al. (2010). They compared primer sets that spanned all six regions of the 16S gene and determined that the set targeting the V3/V4 region had the best simulated taxonomic coverage but due to the shorter sequence lengths available at that time actual results were taxonomically biased. Subsequent work on effective study designs has reinforced the notion that the commonly used V3/V4 region is ideal for metagenomic studies (Mizrahi-Man et al. 2013). ITS2 was selected as it has been chosen by the fungal barcoding
community as a standard marker (Schoch et al. 2012). The 18S-AMF specific primers were chosen to compliment ITS2 AMF sequences and because they provided an alternative to other study designs that required several AMF primers and different cycling conditions (Opik et al. 2009). The 18S universal primers used in this study were chosen simply because Baldwin et al. (2013) was one of the first studies to examine soil eukaryotes with metagenomic techniques.

Table 2.2. Three PCR annealing temperatures chosen for each primer set based on annealing temperature gradient experiments.

<table>
<thead>
<tr>
<th>Primer set</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prokaryote 16S V3-V4 rDNA</td>
<td>63°C</td>
<td>66°C</td>
<td>69°C</td>
</tr>
<tr>
<td>Arbuscular Mycorhizal Fungi 18S SSU rDNA</td>
<td>66°C</td>
<td>69°C</td>
<td>72°C</td>
</tr>
<tr>
<td>Universal Fungi ITS2</td>
<td>58°C</td>
<td>61°C</td>
<td>64°C</td>
</tr>
<tr>
<td>Eukaryotic 18S SSU rDNA</td>
<td>63°C</td>
<td>66°C</td>
<td>69°C</td>
</tr>
</tbody>
</table>

Additional PCR experiments were conducted to determine that the minimum number of cycles needed to get a "bright" agarose gel band was 30 cycles for all primer sets. Further experiments determined that in 12.5μL reactions, 4μL of template DNA yielded brighter agarose gel bands compared to the 2μL which is commonly used. Template DNA for all primer tests was from sample EM1-307 (no-till).

2.7 MICROBIAL PCR REACTION SETUP

Decontamination protocols were as follows:

1. Pipettes were disassembled and decontaminated with Eliminase. All parts were rinsed twice with ddH2O before a 15 minute UV sterilization, air drying and re-assembly.
2. The AirClean PCR hood to be used for reaction setup was decontaminated with Eliminase. All surfaces were rinsed twice with ddH₂O before a 15 minute UV sterilization with nothing in the cabinet that could shield contaminants.

3. All reaction tubes were UV sterilized for 30 minutes. Before PCR set up began, a Kimberly-Clark Sterile Cleanroom full body suit was donned to reduce the risk of personal microbiome DNA contamination. For each sample, PCR reactions for the four primer sets were set up independently (Table 2.1). For each of these, a total reaction volume (including template) of 37.5 μL was mixed and then three 12.5 μL aliquots were moved into separate wells. KAPA’s reaction setup protocols were adhered to with 6.25 μL of ReadyMix, 4 μL of template, 0.375 μL of forward or reverse primer (0.3 μM) and 1.5 μL ddH₂O being used. Eppendorf Mastercycler Pro S thermal cyclers were used to run the following protocol: 95°C for 5 minutes followed by 30 cycles of denaturation at 98°C for 20 seconds, primer annealing for 15 seconds and extension at 72°C for 35 seconds. A final extension at 72°C for 5 minutes was used. These reactions were run at the three different annealing temperatures that were previously selected (Table 2.2, Figure 2.4). For each primer set, a negative control was created using the same ddH₂O that was used for the PCR master mix. All reactions were vortex mixed and briefly centrifuged before being placed in the thermal cycler.

Reaction success was confirmed with an Invitrogen E-gel and reactions run at the three different annealing temperatures were then pooled. Amplicons were stored at 4°C to be purified the following day.
2.8 POOLED PCR PRODUCT PURIFICATION

Aline PCRClean DX was used to purify all amplicons and a size selection protocol was adopted from Ampure beads in order to help reduce primer carryover. For 16S, ITS2 and AMF-18S, a ratio of 0.55 (beads):1 (sample) was used. For 18S, a ratio of 0.9:1 was used due to its shorter length. Amplicons were eluted in 10mM Tris-HCL (pH 7.6). An Invitrogen E-gel was used to confirm purification success. Amplicons were stored at 4°C for index PCR the following day.

2.9 INDEX PCR FOR MULTIPLEX SEQUENCING
Prior to proceeding with the full scale workflow, small scale experiments were done to examine the influence of the number of cycles and annealing temperature on index PCR success (Figure 2.5)

![Figure 2.5. Index PCR tests on 16S amplicons. The three lanes to the left tested the influence of the number of cycles. The lanes to the right tested different annealing temperatures in a 30 cycle program. On the far right, un-indexed amplicons were loaded into the 2% gel as a control. The right-most and left-most lanes are 100bp DNA ladder.](image)

For the full scale workflow, a master mix tube was prepared for each index primer. Half the total volume of master mix was added to each reaction well from either master mix tube containing the forward or reverse index primer (Figure 2.6). The total reaction volume was 50μL. The following volumes were used for each reaction: 25μL KAPA HiFi HotStart ReadyMix, 1.5μL of forward or reverse index primer, 4μL template amplicons, 18μL ddH2O. Index PCR reactions were also set up for all negative controls from taxon specific PCR. The following cycling conditions were used: 95°C for 5min, followed by 10 cycles of 98°C for 20s, 64°C for
15s, 72°C for 35s. A final extension of 72 °C for 5min was used. Amplicons were stored at 4°C to be purified the following day. This index PCR layout allows amplicons in each well of a PCR plate to be tagged with a unique combination of forward and reverse index primers (Figure 2.7).

Figure 2.6. Stylized depiction of index PCR layout in a standard 96 well PCR plate (only half was used in this study). Each red arrow represents a unique forward primer while each green arrow represents a unique reverse primer.
2.10 INDEX PCR PURIFICATION

PCRClean DX was used, but this time at the full concentration of 1.8X (beads): 1(sample). This was done in order to minimize amplicon loss. Amplicons were stored at 4°C to be quantified and normalized the following day.

2.11 DNA LIBRARY QUANTIFICATION AND NORMALIZATION

An Invitrogen Q-Bit Fluorometer was used to measure all amplicon concentrations as recommended by Simbolo et al. (2013). In order to determine whether to use the high sensitivity
or the broad range kit, a few samples with the brightest agarose gel were bands Nano-Dropped to see if their concentration was in excess of 20ng/µL. 5 µL of amplicons were used for measurement. All samples were normalized to 4nmol with the addition of 10mM Tris-HCL (pH 7.6). Amplicons were then pooled and submitted to the Advanced Analysis Center Genomics Facility at the University of Guelph for Illumina MiSeq sequencing using the 250bp paired end read chemistry. Amplicons were loaded into the MiSeq as a 6pM denatured library which was 15% PhiX viral genome, recommended by Illumina to increase the sequence diversity of a library for improved base call quality. In low diversity libraries, florescence from identical bases makes it difficult for the optical detector to discern the correct signal for each amplicon.

2.12 DNA SEQUENCE ANALYSIS

Two MiSeq runs were used to sequence the amplicons for the microbial portion of this study. In the first run, amplicons from all four primer sets were sequenced for 10 between-plot replicates (two replicates per tillage type from different plots) plus the four negative controls. In the second run, amplicons from all four primer sets were sequenced for 10 within-plot replicates (two replicates per tillage type from the same plot) plus the four negative controls. Two MiSeq runs prior to these two were loaded with a library that was too concentrated and as a result, most reads failed to pass either the MiSeq internal quality filter, or the external filter command in UPARSE (Edgar 2013). These two "failed" runs were not used in the final analysis.

2.13 BIOINFORMATIC PIPELINE
Primers were trimmed from the raw reads using Cutadapt (Martin 2011). UPARSE was chosen to quality filter and cluster all reads as it is one of the few packages that was benchmarked against a microbial mixture with known taxa as part of its original release (Edgar 2013). For 16S and 18S, forward and reverse reads were merged using -fastq_mergepairs in UPARSE. 16S reads were then trimmed to a common length of 400bp with reads of a shorter length being discarded. Length trimming for all genes was chosen by examining output from the fastq_stats command. This was done to find the optimal balance between trimming poor quality sequence ends and retaining as much sequence length as possible. For 18S, reads were trimmed to a common length of 175bp. All reads with an expected error rate higher than one base pair per read were discarded (Table 2.3). For ITS2 and AMF-18S the amplicons for these genes were larger than 500bp which is the maximum sequence length that the MiSeq can produce. As a result, the reverse reads were discarded as merging the forward and reverse reads was not possible. In cases such as this reverse reads are chosen to be discarded as they typically have lower quality scores (UPARSE manual; http://drive5.com/usearch/manual/fastq_choose_filter.html). The forward reads for AMF-18S were quality filtered in the same manner as previously mentioned except that they were trimmed to 227bp. Forward reads for ITS2 were quality filtered in the same manner. All treatment replicates were pooled as recommended by UPARSE (analysis of individual replicates was done later see 2.14). Each replicate also received a negative control for the gene region in question from each MiSeq run for a total of 8 negative controls per treatment type. Full-length duplicate sequences were removed using UPARSE, with the '-size_out' command providing the original number of sequences for each representative. These unique sequences were sorted by their
abundance and singletons and doubletons were discarded (UPARSE manual; http://drive5.com/usearch/manual/singletons.html).

At this point, the ITS2 reads were processed with an additional step. ITSx (Bengtsson-Palme et al. 2013) was used to extract the ITS2 region as recommended by the fungal pipeline proposed by Bálint et al (2013). This was done as the LSU region constituted the 3' end of the reads. This region is highly conserved and can influence clustering and taxonomic ID results. The extraction of ITS2 resulted in variation in total sequence length so all ITS2 reads were trimmed again to a common length of 144bp.

All reads were then clustered in UPARSE at 99% sequence similarity as opposed to the "standard" 97%. This has recently been demonstrated to reduced the sequence diversity lost by clustering at lower thresholds (Powell and Sikes 2014; Tikhonov 2014; Yamamoto et al. 2014). Basic ecological diversity metrics were calculated for pooled OTU files for each treatment and gene using the UPARSE command '-fasta_diversity' with 100 iterations. These metrics were total OTU richness, Shannon entropy and Renyi entropy. Shannon entropy provides a measure of the complexity of a sample by calculating the average minimum number of bits needed to characterize the OTUs present based on their frequency. Renyi entropy is similar to Shannon entropy but based upon the quadratic distribution of OTUs. Faith's phylogenetic diversity was computed for all 16S OTUs for each treatment from their respective branch lengths (Faith 1992). This was done by first aligning all OTUs of a given treatment with the RDPipline 16S Aligner (Cole et al. 2014). A maximum likelihood phylogeny was then computed using FastTree 2.1 with the default settings of the generalized time reversible (GTR) model and CAT approximation selected (Price et al. 2010). Across all five phylogenies, six OTUs were deleted as outliers as their branch lengths were longer than the depth of their phylogenies. Phylogenies were used in
conjunction with an OTU table generated by UPARSE to compute the phylogenetic diversities in mothur (Schloss et al. 2009).

For 16S and 18S, the SILVA 119 release was used as a reference for taxonomic identifications (Quast et al. 2013). The identifications themselves were made using the SINA package which was downloaded and run on an Amazon Web Services EC2 instance. 16S and 18S sequences were first aligned and then identified against the SILVA 119 release using default SINA settings. For ITS2 taxonomic identifications, the UNITE release date 2014/09/10 database (Kõljalg et al. 2013) was downloaded and queried with a local installation of BLASTn using default search parameters. BLASTn was chosen as it is optimized for sequences that are "somewhat similar" to those in the library- this is ideal for unexplored soil diversity. For AMF-18S the Maarjam database (Opik et al. 2010) was downloaded in addition to a custom NCBI database. The search keywords: "Glomeromycota" and "arbuscular mycorrhiza 18S" yielded 87,817 and 3534 records respectively. A custom BLASTn database was constructed with all of these sequences. This method was chosen over the SILVA package as NCBI contains many more representatives from AMF taxa. This subset of the entire NCBI database was chosen as the AMF-18S primers also amplify 18S from many other taxa. By only downloading AMF sequences, sorting taxonomic identifications was much simpler in addition to the query time that was saved. Taxonomic identifications from all four primer sets were loaded into the MEGAN5 (5.6.3) package and the "most correct" identifications were chosen using the least common ancestor (LCA) algorithm with default settings (Huson et al. 2011).

2.14 SEQUENCE ANALYSIS OF INDIVIDUAL REPLICATES
The following pipeline was used for both the microbial results and the arthropod (CO1) results from chapter III.

In order to assess the statistical significance of patterns of diversity between treatments, all replicates had to be individually processed prior to being analyzed with Statistical Analysis of Metagenomic Profiles (STAMP) (Parks et al. 2014). The same UPARSE pipeline and settings were used to process these individual replicates for all 4 genes. For ITS2, the ITSx program was used once again to extract the ITS2 region from the surrounding conserved domains prior to sequence dereplication and clustering. This was a highly time consuming step, so 10 Amazon Web Services EC2 instances were purchased to reduce the total compute time to 2 days.

In order to avoid the time and resource intensive process associated with querying all replicates against the entire SILVA, UNITE or Genbank databases for taxonomic identifications, the following workflow was developed. In general, for the pooled samples from each treatment type, taxonomic identifications were linked with their respective OTUs in a single FASTA file. Since these OTUs were generated from pooled and dereplicated sequences from all replicate samples, then all OTUs from every replicate are present. Therefore, the FASTA file of all OTUs and taxonomic identifications could be used as a database to assign taxonomic identifications to OTUs from the individual replicates.

Specifically, this workflow comprised the following steps:

1. The OTU and SILVA log files for each treatment and gene were loaded into MEGAN 5 and the OTU names (Illumina format) and their associated taxonomic identifications were exported in .DSV format. For 16S, 18S and ITS2, class was the lowest taxonomic rank that was selected. For 18S-AMF and CO1, genus was selected. However, all taxonomic identifications that were at
a level higher than these were still included along with all unclassified reads. Therefore no OTUs
would be excluded from downstream statistical analyses. The 18S primers amplified DNA from
non-target plants, fungi and bacteria, so all identifications within these taxa were collapsed to
Viridiplantae, Fungi and Bacteria to simplify downstream analyses.

2. The MEGAN 5 .DSV output was opened in Microsoft Excel and sorted alphabetically by read
names.

3. http://darwin.biochem.okstate.edu/fasta2tab/ to was used to convert the separate OTU FASTA
file to tab delimited format. This OTU tab file was then opened in a separate Microsoft Excel
window and was sorted by read name. This created a means of linking the OTU sequences in one
file with the taxonomic identifications in another file via the sorted Illumina format read names
that they shared.

4. The two files were merged and all redundant information was deleted. A FASTA file was then
created that contained the OTU sequences with their taxonomic identifications as headers. These
files will now be referred to as libraries.

5. For each gene, libraries from all 5 treatments were concatenated together to make a master
library. The master libraries were dereplicated using the UPARSE -derep_fulllength command.

6. For each gene, OTUs from every replicate sample were labeled using the sed command from
the UPARSE pipeline: sed "-es/^>\(.*\)/\1;barcodelabel=SampleID;" < input.fasta >
output.fasta and then concatenated together into a single "query file" for each gene using the cat
command.

7. The search_exact command was used in UPARSE to match the sequences in the "query files"
for each gene against their respective master libraries. The output was generated in .UC format
and thus provided taxonomic identifications for all queries.
8. The sed command: sed "-es/size=[0-9]*://" < input.uc > output.uc was used to remove the "size=" information from each query OTU or the OTU table building script would erroneously include the number of reads that OTUs were generated from.

9. The UPARSE python script uc2otutab.py was used to generate an OTU table for each gene from all sample replicates.

10. This OTU table was then loaded into STAMP for statistical analysis.

11. OTUs within taxonomic ranks from all treatments were compared using the multiple groups option in STAMP with ANOVA selected. The Games-Howell post-hoc test was used as recommended by the STAMP creators to elucidate which treatments differed significantly.

2.15 SOIL MICROBIAL DIVERSITY RESULTS- SEQUENCING SUCCESS

Unfortunately, the first two MiSeq runs were loaded with amplicons at a concentration that yielded a high cluster density. This resulted in a very high frequency of poor quality base calls and many samples yielded too few high quality reads to be useful. Data from these first two runs are not part of the final analysis in this study. However, some useful data were still obtained from these runs. Due to the shorter length of the 18S amplicons, ~20,000 -60,000 reads per sample were obtained for both runs. Much of this variability is likely due to the fact that a lower cluster density was used in the second run. Regardless, in four cases, an identical sample of bulk DNA template was sequenced in the first and second run. This provided a mean of examining the variability introduced by laboratory and sequencing protocols. Despite different cluster densities, the majority of 18S taxonomic groups do not have significantly different OTU richness between the two MiSeq runs (Figure 2.8). Ultimately, it is difficult to know how the variability between identical templates was generated. This is most likely due to the different cluster densities but
could also reflect variations in PCR. In the future this form of control will used in a more reliable manner. Only 18S was used as this form of control due to all other amplicon types from run 2 having a high level of bad base calls, making comparisons with run 3 impossible. One reason that 18S sequencing was largely successful in this environment may be due to its shorter length (175bp average as opposed to >400bp). With the adjustments made to amplicon library preparation and MiSeq loading conditions (6pM library instead of 12pM), attempts to re-sequence all samples in runs 3 and 4 yielded a high read depth (Table 2.3).

2.17 SOIL MICROBIAL DIVERSITY RESULTS

When examining within-plot sample replicates for 16S, 18S and ITS2, the majority of OTU richness at a given taxonomic level is not significantly different (Figures 2.9-2.11). There is, however, some variability with 7.5±3.1% of 16S, 6.9±1.0% of 18S and 0.64±0.6% (standard error) of ITS2 taxonomic groups being significantly different in terms of OTU richness between within-plot replicates (these means include the 5th treatment that was not included in Figures 2.9-2.11 to make an efficient use of space). Results for 18S-AMF are not shown here as the vast majority of reads were unassigned, as these primers amplified 18S from many Eukaryotes not within Glomeromycota (Figure 2.22). For the AMF-18S reads that were assigned taxonomic identifications, 10±6.1% (standard error) taxonomic identifications were different in terms of OTU richness between replicates. Although, this mean is not highly informative in this case as there were only four different taxonomic identifications available for comparison, and three sets of within-plot replicates had zero significant differences, while two sets had one.

The ecological diversity metrics computed for all treatments and genes demonstrate that their alpha diversities are similar (Table 2.4). However, the results from the phylogenetic
diversity analysis of 16S OTUs suggests that these treatments may still have their own "fingerprints" of diversity (Table 2.5).

It is interesting to note that the AMF OTUs detected here are an order of magnitude more rich than some previous work on AMF suggests (Opik et al. 2009). This is likely due to the 99% clustering threshold being used in this study compared with the standard 97%, and some of this increase in diversity is likely within-species variability.

When examining the principal component analysis results, only the fungal community (ITS2) exhibits clustering of treatment replicates, with no-till being the most distinct (Figure 2.12). This is also the case when examining the nearest neighbour clustering results obtained from the taxonomy heat-map plots (Figures 2.19- 2.22). ITS2 was the only amplicon to have three replicates cluster together (Figure 2.21). More detailed statistical testing by means of ANOVAs and Games-Howell post-hoc testing also supports the notion that these 5 tillage treatments only differ significantly in the OTU richness of their fungal communities (Figures 2.15, 2.16). With use of ANOVAs, the proportion of the taxonomic identifications that differ in terms of OTU richness between treatments was calculated to be 26% (Table 2.8). Of these significant differences between treatments, no discernible pattern was apparent for the Eurotiomycetes, Chytridiomycota, Cystobasidiomycetes and the Microbotryomycetes (Figure 2.14). However, the post-hoc results for the Sordariomycetes and the Dothideomycetes were very interesting (Figure 2.16). Overall, the Sordariomycetes were the most diverse fungal group (Figure 2.21). However, the no-till plots had a significantly lower Sordariomycetes OTU richness compared to three of the four tillage treatments (Eta-squared effect size = 0.63) (Figures 2.15, 2.16). Despite spring moldboard not having a significantly higher OTU richness than no-till, its mean is comparable to the other tillage treatments (Figure 2.15). When examining the
Sordariomycetes taxonomic identifications at the level of genus, many of the OTUs that make the fall moldboard and fall chisel treatments the most rich are from known pathogens such as *Fusarium*, *Cylindrocarpon* and *Metarhizium* (Figure 2.17).

Statistical tests indicate that the OTU richness of Dothideomycetes follows the opposite trend (Figure 2.15). No-till plots were significantly more rich than three of the four tillage treatments with the only large effect size seen in this study of 0.79. The tillage treatment that was not significantly different, was spring tandem disc and this can be attributed to its higher variability between replicates (Figure 2.16). When examining the Dothideomycetes taxonomic identifications at the level of genus, an extremely interesting pattern is apparent (Figure 2.18). The increased OTU richness of the no-till plots can be attributed to the obligate coprophilovore, *Sporormiella* of which 22 OTUs were detected. In all four replicates, from all four of the tillage treatments, only one *Sporormiella* OTU was found (Figure 2.18). On a broad scale, when examining the heat-map plots (Figures 2.19-2.22), the OTUs of fungal community are more evenly spread across multiple taxonomic groups leading to a lower proportion of rare taxa when compared with 16S, 18S or 18S-AMF. However, ITS2 also had the highest frequency of unassigned reads with ~25% being unassigned when compared with 16S and 18S (not including the AMF specific primers).

ANOVA detected significant differences in OTU richness in 12.5% of all bacterial taxonomic identifications between treatment types (Table 2.6). Spring tandem disc had a significantly higher OTU richness in 4 of the 6 significantly different taxonomic groups (Figure 2.13). However these p-values are not supported with large effect sizes or by general trends apparent in Figure 2.13 and can most likely be attributed to random biological variation (future work from a temporal perspective could confirm or refute this). Of all four amplicon types, 16S
was the only one in which all OTUs were assigned taxonomic identifications. However, many were only assigned to Bacteria (Figure 2.19). The dominant taxa of the bacterial communities in these soils are the Acidobacteria and the Actinobacteria (Figure 2.19).

ANOVA detected significant differences in OTU richness between treatments for 3% of universal 18S taxonomic identifications (Table 2.7). ANOVA did not detect any significant differences in OTU richness for 18S-AMF taxonomic identifications (Table 2.9). However, for both the universal 18S and the AMF-18S amplicons, post-hoc tests did not detect any significant differences and therefore no box plots are provided for these groups. Universal 18S amplicons were most often assigned the identification of Eukaryota (Figure 2.20), which suggests that either the majority of the Eukaryotic community present in these soils does not have representatives in the SILVA 119 database; or that there are taxonomic inconsistencies which results in the LCA algorithm in MEGAN assigning this very coarse identification.

Rarefaction curves indicate that there is still some unexplored diversity in these soils (Figure 2.23). It is worth reiterating that the majority of 18S-AMF amplicons were non-target sequences from other Eukaryotes (Figure 2.22) and this likely explains the linear trend obtained for these primers.

2.18 MICROBIAL DIVERSITY FIGURES

Table 2.3. Number of sequence reads obtained for each of the four Loci at progressing stages of the UPARSE pipeline. Results for ITS extraction using ITSx are also shown. In both runs 3 and 4, two replicates from all 5 treatments were sequenced for a total of 10 samples per run.
Table 2.4. Diversity metrics computed in UPARSE for pooled OTU files from each treatment and gene. N represents the total number of reads that contributed to the OTUs.

<table>
<thead>
<tr>
<th>Loci</th>
<th>Run 1</th>
<th>Run 2</th>
<th>Run 3</th>
<th>Run 4</th>
<th>Run 3</th>
<th>Run 4</th>
<th>Run 3</th>
<th>Run 4</th>
<th>Run 3</th>
<th>Run 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>18S</td>
<td>123</td>
<td>123</td>
<td>123</td>
<td>123</td>
<td>123</td>
<td>123</td>
<td>123</td>
<td>123</td>
<td>123</td>
<td>123</td>
</tr>
<tr>
<td>ITS2</td>
<td>567</td>
<td>567</td>
<td>567</td>
<td>567</td>
<td>567</td>
<td>567</td>
<td>567</td>
<td>567</td>
<td>567</td>
<td>567</td>
</tr>
</tbody>
</table>

Table 2.5. Phylogenetic diversity of 16S OTUs from pooled OTU files from each treatment

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Phylogenetic Diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall Chisel</td>
<td>231.61</td>
</tr>
<tr>
<td>Fall Moldboard</td>
<td>179.55</td>
</tr>
<tr>
<td>No-till</td>
<td>205.39</td>
</tr>
<tr>
<td>Spring Moldboard</td>
<td>196.07</td>
</tr>
<tr>
<td>Spring Disc</td>
<td>196.14</td>
</tr>
</tbody>
</table>
Figure 2.8. Taxonomic results for 18S amplicons from four identical DNA templates sequenced on two separate MiSeq runs at different loading concentrations (R1= 12pM, R2= 6pM).

Significant differences in OTU richness within a taxonomic rank are marked with an asterisk via the default settings for a two sample comparison in STAMP (two sided G-test w/Yates’ + Fisher). All taxa with OTU richness too low to appear on this figure were discarded to make efficient use of space. There were no significant differences in these taxa. A) Spring moldboard B) Fall moldboard C) Spring moldboard D) Spring moldboard
Figure 2.9. Taxonomic results for 18S amplicons from eight within-plot replicates (blue or orange bars) sampled from four tillage treatment types. Significant differences in OTU richness within a taxonomic rank are marked with an asterisk via the default settings for a two sample comparison in STAMP (two sided G-test w/Yates' + Fisher). All taxa with OTU richness too low to appear on this figure were discarded to make efficient use of space. There were no significant differences in these taxa. A) No-till B) Spring disc C) Spring moldboard D) Fall chisel (fall moldboard omitted due to space limitations)
Figure 2.10. Taxonomic results for 16S amplicons from eight within plot replicates (blue or orange bars) sampled from four tillage treatment types. Significant differences in OTU richness within a taxonomic rank are marked with an asterisk via the default settings for a two sample comparison in STAMP (two sided G-test w/Yates’ + Fisher). All taxa with OTU richness too low to appear on this figure were discarded to make efficient use of space. There were no significant
differences in these taxa. A) No-till B) Spring disc C) Spring moldboard D) Fall chisel (fall moldboard omitted due to space limitations)
Figure 2.11. Taxonomic results for ITS2 amplicons from eight within plot replicates (blue or orange bars) sampled from four tillage treatment types. Significant differences are marked with an asterisk via the default settings for a two sample comparison in STAMP (two sided G-test w/Yates’ + Fisher). All taxa with OTU richness too low to appear on this figure were discarded to make efficient use of space. There were no significant differences in these taxa. A) No-till B) Spring disc C) Spring moldboard D) Fall chisel (fall moldboard omitted due to space limitations)

Figure 2.12. Principal component analysis results comparing taxonomic identifications and their OTU richness. Computed in STAMP using default settings. Comparisons were made using taxonomically identified OTUs from all treatment replicates from five tillage regimes. In the legend, the numbers in parentheses correspond to the number of replicates that were analyzed.
Figure 2.13. 16S taxonomic groups that yielded significant differences in OTU richness of OTUs (y-axis) between tillage treatments. ANOVAs followed by Games-Howell post-hoc tests were used to identify significant differences (p < 0.05) in the program STAMP. The p-values listed were generated from ANOVAs. OTUs were clustered at 99% similarity from 400bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures. The letters above boxes indicate the other box that they are significantly different from.
Figure 2.14. ITS2 taxonomic groups that yielded significant differences in OTU richness of OTUs (y-axis) between tillage treatments but no broad scale patterns were apparent. ANOVAs followed by Games-Howell post-hoc tests were used to identify significant differences (p < 0.05) in the program STAMP. The p-values listed were generated from ANOVAs. OTUs were
clustered at 99% similarity from 144bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures. The letters above boxes indicate the other box that they are significantly different from.

Figure 2.15. ITS2 taxonomic groups that yielded significant differences in OTU richness of OTUs (y-axis) between tillage treatments and demonstrated interesting patterns. ANOVAs followed by Games-Howell post-hoc tests were used to identify significant differences (p < 0.05)
in the program STAMP. The p-values listed were generated from ANOVAs. OTUs were clustered at 99% similarity from 144bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures. The letters above boxes indicate the other box that they are significantly different from.
Figure 2.16. Detailed Games-Howell post-hoc test results for the Dothideomycetes and Sordariomycetes data generated in STAMP. All parameters are identical to those described in Figure 2.15.

Figure 2.17. Genera within Sordariomycetes that were detected in the different tillage treatments from ITS2 reads. This figure was generated from pooled replicates in MEGAN 5. The legend is sorted from most diverse to least diverse.
Figure 2.18. Genera within Dothideomycetes that were detected in the different tillage treatments from ITS2 reads. This figure was generated from pooled replicates in MEGAN 5. The legend is sorted from most diverse to least diverse.

Table 2.6. Summary of all taxonomic identifications made for the 16S amplicons along with p-values from between treatment ANOVAs and Eta-squared effect sizes. Significant differences are boxed, see Figure 2.13 for post-hoc results.
<table>
<thead>
<tr>
<th>Taxonomic Id</th>
<th>p-values</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidobacteria</td>
<td>0.638639</td>
<td>0.146823</td>
</tr>
<tr>
<td>Acidobacteria</td>
<td>0.41629</td>
<td>0.218149</td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>0.082484</td>
<td>0.404704</td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>0.125343</td>
<td>0.360667</td>
</tr>
<tr>
<td>Armamonadates</td>
<td>0.685573</td>
<td>0.132801</td>
</tr>
<tr>
<td>Armamonadiales</td>
<td>0.044742</td>
<td>0.458212</td>
</tr>
<tr>
<td>Bacilli</td>
<td>0.736443</td>
<td>0.117586</td>
</tr>
<tr>
<td>Bacteria</td>
<td>0.007258</td>
<td>0.585835</td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>0.300509</td>
<td>0.263382</td>
</tr>
<tr>
<td>Bacteroidiales</td>
<td>0.579618</td>
<td>0.184701</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>0.446663</td>
<td>0.207542</td>
</tr>
<tr>
<td>cellular organisms</td>
<td>0.438001</td>
<td>0.210526</td>
</tr>
<tr>
<td>Chlamydiaceae/Verrucomicrobia</td>
<td>0.152653</td>
<td>0.343248</td>
</tr>
<tr>
<td>Chlamydia</td>
<td>0.83944</td>
<td>0.08554</td>
</tr>
<tr>
<td>Chlorobi</td>
<td>0.837962</td>
<td>0.086024</td>
</tr>
<tr>
<td>Chlorobi</td>
<td>0.002782</td>
<td>0.638869</td>
</tr>
<tr>
<td>Chloroflexa</td>
<td>0.446957</td>
<td>0.207441</td>
</tr>
<tr>
<td>Clostridia</td>
<td>0.096942</td>
<td>0.389401</td>
</tr>
<tr>
<td>Cyanobacteria</td>
<td>0.172334</td>
<td>0.330037</td>
</tr>
<tr>
<td>Cytophagia</td>
<td>0.961115</td>
<td>0.03793</td>
</tr>
<tr>
<td>delta/epsilon</td>
<td>0.438001</td>
<td>0.210526</td>
</tr>
<tr>
<td>Delta-proteobacteria</td>
<td>0.435998</td>
<td>0.211221</td>
</tr>
<tr>
<td>Elusimicrobiota</td>
<td>0.952943</td>
<td>0.042108</td>
</tr>
<tr>
<td>environmental_samples_Acidobacteria</td>
<td>0.129881</td>
<td>0.360239</td>
</tr>
<tr>
<td>environmental_samples_Bacteria</td>
<td>0.000623</td>
<td>0.707626</td>
</tr>
<tr>
<td>environmental_samples_Gemmataionadates</td>
<td>0.724274</td>
<td>0.121239</td>
</tr>
<tr>
<td>environmental_samples_Cyanobacteria</td>
<td>0.492788</td>
<td>0.192117</td>
</tr>
<tr>
<td>environmental_samples_Verrucomicrobia</td>
<td>0.093268</td>
<td>0.393111</td>
</tr>
<tr>
<td>environmental_samples_Verrucomicrobia</td>
<td>0.019247</td>
<td>0.522466</td>
</tr>
<tr>
<td>Epsilon-proteobacteria</td>
<td>0.375984</td>
<td>0.232536</td>
</tr>
<tr>
<td>Erythromicrobia</td>
<td>0.219217</td>
<td>0.302524</td>
</tr>
<tr>
<td>Fibrobacteria</td>
<td>0.822043</td>
<td>0.091176</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>0.076008</td>
<td>0.412252</td>
</tr>
<tr>
<td>Flavobacteria</td>
<td>0.731483</td>
<td>0.119076</td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>0.056112</td>
<td>0.439173</td>
</tr>
<tr>
<td>Gemmatimonadetes</td>
<td>0.094466</td>
<td>0.351889</td>
</tr>
<tr>
<td>Holophagae</td>
<td>0.051468</td>
<td>0.446343</td>
</tr>
<tr>
<td>Nitrosphaernae</td>
<td>0.736055</td>
<td>0.117702</td>
</tr>
<tr>
<td>Nitrospirae</td>
<td>0.018563</td>
<td>0.525008</td>
</tr>
<tr>
<td>No Hits</td>
<td>0.279033</td>
<td>0.275998</td>
</tr>
<tr>
<td>Physicisphaerae</td>
<td>0.84931</td>
<td>0.082278</td>
</tr>
<tr>
<td>Planctomyces</td>
<td>0.376598</td>
<td>0.232715</td>
</tr>
<tr>
<td>Planctomycetaceae</td>
<td>0.709914</td>
<td>0.125536</td>
</tr>
<tr>
<td>Proteobacteria</td>
<td>0.35074</td>
<td>0.24264</td>
</tr>
<tr>
<td>Solibacteres</td>
<td>0.797572</td>
<td>0.058808</td>
</tr>
<tr>
<td>Sphingobacteria</td>
<td>0.166122</td>
<td>0.33408</td>
</tr>
<tr>
<td>Unclassified</td>
<td>0.123622</td>
<td>0.365303</td>
</tr>
<tr>
<td>Verrucomicrobiota</td>
<td>0.369463</td>
<td>0.235336</td>
</tr>
</tbody>
</table>

Percent of taxa with significant p values = 12.5%
Table 2.7. Summary of all taxonomic identifications made for the 18S amplicons along with p-values from between treatment ANOVAs and Eta-squared effect sizes. Significant differences across all treatments are boxed, post-hoc results yielded no significant differences.
<table>
<thead>
<tr>
<th>Taxonomic Id</th>
<th>p-values</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alveolata</td>
<td>0.882717</td>
<td>0.07078</td>
</tr>
<tr>
<td>Amnion</td>
<td>0.160437</td>
<td>0.33788</td>
</tr>
<tr>
<td>Amoebozoa</td>
<td>0.039579</td>
<td>0.467758</td>
</tr>
<tr>
<td>Annelida</td>
<td>0.510795</td>
<td>0.186281</td>
</tr>
<tr>
<td>Arachnida</td>
<td>0.242787</td>
<td>0.290268</td>
</tr>
<tr>
<td>Arthropoda</td>
<td>0.438001</td>
<td>0.210526</td>
</tr>
<tr>
<td>Asciidaea</td>
<td>0.305613</td>
<td>0.261177</td>
</tr>
<tr>
<td>Bacillariophyta</td>
<td>0.052757</td>
<td>0.444444</td>
</tr>
<tr>
<td>Bacteria</td>
<td>0.757188</td>
<td>0.11132</td>
</tr>
<tr>
<td>Bdelloides</td>
<td>0.058302</td>
<td>0.444015</td>
</tr>
<tr>
<td>Bilateria</td>
<td>0.948915</td>
<td>0.044061</td>
</tr>
<tr>
<td>cellular_organisms</td>
<td>0.735863</td>
<td>0.11776</td>
</tr>
<tr>
<td>Cercozoa</td>
<td>0.076206</td>
<td>0.409635</td>
</tr>
<tr>
<td>Chelicerata</td>
<td>0.095502</td>
<td>0.38688</td>
</tr>
<tr>
<td>Choanoflagellida</td>
<td>0.052757</td>
<td>0.444444</td>
</tr>
<tr>
<td>Chromadorea</td>
<td>0.239835</td>
<td>0.291756</td>
</tr>
<tr>
<td>Ciliophora</td>
<td>0.547858</td>
<td>0.174534</td>
</tr>
<tr>
<td>Clitellata</td>
<td>0.478269</td>
<td>0.195833</td>
</tr>
<tr>
<td>Coccidia</td>
<td>0.47695</td>
<td>0.197331</td>
</tr>
<tr>
<td>Conocephalum</td>
<td>0.722534</td>
<td>0.121761</td>
</tr>
<tr>
<td>Dinophyceae</td>
<td>0.220732</td>
<td>0.301709</td>
</tr>
<tr>
<td>Ectosporidia</td>
<td>0.238081</td>
<td>0.292646</td>
</tr>
<tr>
<td>Ellipura</td>
<td>0.345825</td>
<td>0.244584</td>
</tr>
<tr>
<td>Eukarya</td>
<td>0.293434</td>
<td>0.266479</td>
</tr>
<tr>
<td>Eumetazoa</td>
<td>0.054307</td>
<td>0.41977</td>
</tr>
<tr>
<td>Eutardigrada</td>
<td>0.281827</td>
<td>0.27176</td>
</tr>
<tr>
<td>Euteleostomi</td>
<td>0.371971</td>
<td>0.234432</td>
</tr>
<tr>
<td>Foraminifera</td>
<td>0.192814</td>
<td>0.317423</td>
</tr>
<tr>
<td>Fungi</td>
<td>0.661294</td>
<td>0.140044</td>
</tr>
<tr>
<td>Gregarinasina</td>
<td>0.098037</td>
<td>0.383317</td>
</tr>
<tr>
<td>Hexapoda</td>
<td>0.214695</td>
<td>0.304981</td>
</tr>
<tr>
<td>insecta</td>
<td>0.258026</td>
<td>0.282781</td>
</tr>
<tr>
<td>Intramacronucleata</td>
<td>0.455894</td>
<td>0.203363</td>
</tr>
<tr>
<td>Lophotrochozoa</td>
<td>0.473794</td>
<td>0.198379</td>
</tr>
<tr>
<td>Mandibulata</td>
<td>0.758622</td>
<td>0.110884</td>
</tr>
<tr>
<td>Metazoa</td>
<td>0.031329</td>
<td>0.485579</td>
</tr>
<tr>
<td>Monogononta</td>
<td>0.002293</td>
<td>0.645625</td>
</tr>
<tr>
<td>Nematoda</td>
<td>0.481189</td>
<td>0.195928</td>
</tr>
<tr>
<td>Oligohymenophorea</td>
<td>0.906175</td>
<td>0.021867</td>
</tr>
<tr>
<td>Oomyctes</td>
<td>0.138756</td>
<td>0.353368</td>
</tr>
<tr>
<td>Opisthokonta</td>
<td>0.388603</td>
<td>0.228196</td>
</tr>
<tr>
<td>Panarthropoda</td>
<td>0.623966</td>
<td>0.151232</td>
</tr>
<tr>
<td>Pancrustacea</td>
<td>0.5086</td>
<td>0.180697</td>
</tr>
<tr>
<td>Protostomia</td>
<td>0.197662</td>
<td>0.314582</td>
</tr>
<tr>
<td>PX</td>
<td>0.962866</td>
<td>0.037088</td>
</tr>
<tr>
<td>Rhabditophora</td>
<td>0.527356</td>
<td>0.166913</td>
</tr>
<tr>
<td>Rhizaria</td>
<td>0.218775</td>
<td>0.302763</td>
</tr>
<tr>
<td>Tramenopiles</td>
<td>0.266905</td>
<td>0.278357</td>
</tr>
<tr>
<td>Trichomonadida</td>
<td>0.052757</td>
<td>0.444444</td>
</tr>
<tr>
<td>Unclassified</td>
<td>0.393385</td>
<td>0.225482</td>
</tr>
<tr>
<td>Viridiplantae</td>
<td>0.518216</td>
<td>0.183902</td>
</tr>
</tbody>
</table>

Percent of taxa with significant p values = 5%
Table 2.8. Summary of all taxonomic identifications made for the ITS2 amplicons along with p-values from between treatment ANOVAs and Eta-squared effect sizes. Significant differences across all treatments are boxed, see Figures 7,8,9 for post-hoc results.

<table>
<thead>
<tr>
<th>Taxonomic Id</th>
<th>p-values</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agaricomycetes</td>
<td>0.441342</td>
<td>0.209371</td>
</tr>
<tr>
<td>Agaricostilbomycetes</td>
<td>0.601925</td>
<td>0.157895</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>0.905058</td>
<td>0.062545</td>
</tr>
<tr>
<td>Atractiellomycetes</td>
<td>0.803697</td>
<td>0.077423</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>0.36146</td>
<td>0.238459</td>
</tr>
<tr>
<td>Chytridiomycetes</td>
<td>0.715497</td>
<td>0.123867</td>
</tr>
<tr>
<td>Chytridiomycota</td>
<td><strong>0.004485</strong></td>
<td><strong>0.613522</strong></td>
</tr>
<tr>
<td>Cystobasidiomycetes</td>
<td><strong>0.000865</strong></td>
<td><strong>0.693842</strong></td>
</tr>
<tr>
<td>Dothideomycetes</td>
<td><strong>5.54E-05</strong></td>
<td><strong>0.79088</strong></td>
</tr>
<tr>
<td>Eurotiomycetes</td>
<td><strong>0.006033</strong></td>
<td><strong>0.596715</strong></td>
</tr>
<tr>
<td>Fungi</td>
<td>0.524642</td>
<td>0.181853</td>
</tr>
<tr>
<td>Glomeromycetes</td>
<td><strong>0.048108</strong></td>
<td><strong>0.452212</strong></td>
</tr>
<tr>
<td>Glomeromycota</td>
<td>0.493532</td>
<td>0.191874</td>
</tr>
<tr>
<td>Insecta</td>
<td>0.238143</td>
<td>0.292615</td>
</tr>
<tr>
<td>Ileatiomyceta</td>
<td>0.772831</td>
<td>0.106552</td>
</tr>
<tr>
<td>Leotiomycetes</td>
<td>0.843969</td>
<td>0.08405</td>
</tr>
<tr>
<td>Microbotryomycetes</td>
<td><strong>0.003407</strong></td>
<td><strong>0.6284</strong></td>
</tr>
<tr>
<td>Mitosporic Ascomycota</td>
<td><strong>0.035158</strong></td>
<td><strong>0.477609</strong></td>
</tr>
<tr>
<td>Monoblepharidomycetes</td>
<td>0.260014</td>
<td>0.281827</td>
</tr>
<tr>
<td>No Hits</td>
<td>0.624598</td>
<td>0.151042</td>
</tr>
<tr>
<td>Opisthokonta</td>
<td>0.459959</td>
<td>0.203018</td>
</tr>
<tr>
<td>Orbiliomycetes</td>
<td>0.783484</td>
<td>0.103278</td>
</tr>
<tr>
<td>Pezizomycetes</td>
<td>0.256166</td>
<td>0.283678</td>
</tr>
<tr>
<td>Pucciniomycetes</td>
<td>0.762895</td>
<td>0.109585</td>
</tr>
<tr>
<td>Pucciniomycotina</td>
<td>0.742596</td>
<td>0.115733</td>
</tr>
<tr>
<td>root</td>
<td>0.337525</td>
<td>0.247904</td>
</tr>
<tr>
<td>Saccharomycetes</td>
<td>0.909588</td>
<td>0.060802</td>
</tr>
<tr>
<td>Sordariomycetes</td>
<td><strong>0.003117</strong></td>
<td><strong>0.633084</strong></td>
</tr>
<tr>
<td>Tremellomycetes</td>
<td>0.224417</td>
<td>0.299743</td>
</tr>
<tr>
<td>Unclassified</td>
<td>0.268018</td>
<td>0.278035</td>
</tr>
<tr>
<td>Ustilaginomycetes</td>
<td>0.358724</td>
<td>0.239519</td>
</tr>
</tbody>
</table>

Percent of taxa with significant p values = 26%
Table 2.9. Summary of all Taxonomic identifications made for the 18S-AMF amplicons along with p-values from between treatment ANOVAs and Eta-squared effect sizes.

<table>
<thead>
<tr>
<th>Taxonomic Id</th>
<th>p-values</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diversispora</td>
<td>0.418098</td>
<td>0.217506</td>
</tr>
<tr>
<td>Fungi</td>
<td>0.614131</td>
<td>0.154158</td>
</tr>
<tr>
<td>Glomerales</td>
<td>0.30265</td>
<td>0.262454</td>
</tr>
<tr>
<td>Glomeromycetes</td>
<td>0.60588</td>
<td>0.156695</td>
</tr>
<tr>
<td>Glomeromycota</td>
<td>0.986171</td>
<td>0.02184</td>
</tr>
<tr>
<td>Glomus</td>
<td>0.889113</td>
<td>0.068477</td>
</tr>
<tr>
<td>Opisthokonta</td>
<td>0.601925</td>
<td>0.157895</td>
</tr>
<tr>
<td>Paraglomus</td>
<td>0.438001</td>
<td>0.210526</td>
</tr>
<tr>
<td>Unclassified</td>
<td>0.910083</td>
<td>0.06061</td>
</tr>
</tbody>
</table>
Figure 2.19. Heat-map plot detailing the number of 16S sequences (OTUs) that were assigned to certain taxonomic groups using the SILVA 119 database and the SINA classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.
Figure 2.20. Heat-map plot detailing the number of 18S sequences (OTUs) that were assigned to certain taxonomic groups using the SILVA 119 database and the SINA classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.
Figure 2.21. Heat-map plot detailing the number of ITS2 sequences (OTUs) that were assigned to certain taxonomic groups using the UNITE database and the BLASTn classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.
Figure 2.22. Heat-map plot detailing the number of 18S-AMF sequences (OTUs) that were assigned to certain taxonomic groups using a custom database and the BLASTn classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.
Figure 2.23. Rarefaction curves generated in UPARSE from the pooled sample libraries for each tillage treatment. The x-axis indicates the size (%) of the sub-sampled dataset while the y-axis indicates OTUs.

2.19 MICROBIAL DIVERSITY DISCUSSION

The taxonomic results and alpha-diversity metrics from this study indicate that the bacterial and protozoan communities under the four different tillage treatments and no-till are very similar (Tables 2.4, 2.6, 2.7). However, when examining bacterial communities using the phylogenetic diversity of 16S OTUs, which is independent of a taxonomic framework, it is apparent that these treatments do have unique fingerprints of diversity (Table 2.5). The phylogenetic diversity of fall chisel plow was identified as the most diverse (231.61) followed by
The two spring tillage treatments, moldboard and tandem disc, were identical from this perspective (196.07 vs 196.14). The least diverse was fall moldboard (179.55). These results suggest that while these treatments have similar richness within coarse level taxonomic groups, the identities of the OTUs themselves may be different. This will be a fruitful metric to pursue for the other genes sequenced in this study, especially 18S. It is also important to consider that the methods employed here cannot detect differences in OTU abundances or biomass and this may indeed differ between treatments. Significant differences in OTU richness were found in the fungal community, but for the majority of taxonomic identifications no significant differences were observed between treatments.

At the time of writing, there are few studies examining the microbial communities under different tillage regimes from a metagenomic perspective (Souza et al. 2013, Souza et al. 2015; Carbonetto et al. 2014; Navarro-Noya et al. 2013; Lauber et al. 2013; Lienhard et al. 2013). Three of these studies found that tilled soils contained a more diverse and rich microbial community (See section 1.5 of the introduction). In all cases, significant differences were limited to a small proportion of taxonomic groups, and different sampling and analytical methodologies make comparisons difficult. Furthermore, all of these studies only examined one form of tillage. In a simple tilled versus no-till comparison, random biological variation could be interpreted as biologically significant. However, in this study when viewed in the context of the multiple tillage treatments, the significant differences observed in the bacterial communities are more likely to be explained by random biological variation and may disappear when examined temporally (Figure 2.13). This notion is further supported by the lack of substantial effect sizes in this study, and none of the reviewed studies reported effect sizes with their p-values (Table 2.6). In this study, clear depictions of the within-plot replicate variability can be seen in Figures 2.9-2.11.
is reasonable to assume that the majority of this variability is biological and not an artifact of laboratory or analytical practices. When sequencing four identical DNA templates on two separate MiSeq runs, the majority of taxonomic identifications were not significantly different (Figure 2.8). They were not, however, sequenced under identical conditions, with one set of replicates being loaded onto the MiSeq with double the library concentration that was later determined to be required (reads from these two runs were not included in the final analysis). This likely explains the significant differences and variability that was observed, but given that there were few significant differences, it may be assumed that there would be even less variability under identical loading conditions. This assumption is supported by one of the first proof of concept microbiome publications for the new Illumina chemistry (Caporaso et al. 2012). They sequenced identical human microbiome communities on the Illumina MiSeq and HiSeq and found no significant differences between the identical replicates. This is, however, the only reviewed study that used such a control for non-biological artifacts. Although it will reduce sequencing depth, this form of control should be common practice to ensure laboratory or analytical procedures are not introducing significant variability. It is worth re-iterating here that the results obtained for this study were garnered from the latter MiSeq runs that were loaded with identical library concentrations.

As with the bacterial communities, the fungal communities of these soils also displayed significant differences in six of the taxonomic identifications. Of these, four do not display a treatment specific pattern (Figure 2.14). The remaining two however, the Sordariomycetes and the Dothideomycetes show an interesting pattern (Figure 2.15). In these soils, the most diverse taxonomic group is the Sordariomycetes (103±19 mean OTU richness across treatments); however, no-till soils possess the lowest OTU diversity (86±10 OTUs; p= 3.12e⁻³). This is in
agreement with the results obtained by Souza et al. (2013). However, the most interesting result of this study is that no-till soils significantly possess the highest diversity of Dothideomycetes (43±1 vs 29±6; p= 5.54e⁻⁵) supported by low between-replicate variability and consequently a significant effect size (Eta-squared = 0.79).

Anecdotal experience from sampling these soils suggests that the no-till plots are suffering from compaction. Soils under this treatment type were much denser and retrieving samples was made more difficult to a degree that no-till soils could be indentified in a blinded manner. Our suspicions are supported by previous soil structure research in these plots by Mueller et al. (2009) who were able to determine with visual structure assessments that the no-till soils were compacted and "unfavourable" in general. This makes perfect sense given that under a lack of tillage, there is no force to counteract the compressive forces of heavy farm equipment i.e. planters, sprayers and combines (Pastorelli et al. 2013). Additionally, the differences observed in the Sordariomycetes and the Dothideomycetes could be related to the well documented differences in organic matter and nutrient concentrations of tilled versus no-till soils (Six et al. 1999; Carr et al. 2013; Panettieri et al. 2013).

Explaining the biogeography of microorganisms is a fascinating field as it integrates knowledge of ecological, taxonomic, evolutionary, spatial and environmental factors that act across a hierarchy of scales. The most popular although highly debated, is the notion that "everything is everywhere, but the environment selects" i.e. that the biogeographic patterns of microorganisms are not dispersal limited and their diminutive size places them under fewer physiological constraints than large organisms (Fontaneto and Brodie 2011). Therefore, the environment is responsible for their distributions. This notion of boundary free dispersal certainly holds true in these experimental tillage plots. Not only are they tightly linked spatially,
but the agricultural vehicles provide a means of cross inoculation. Therefore, differences in community structure are likely to be due to environmental factors since there are no obvious barriers to dispersal between plots. However, this ease of dispersal between plots could also explain the non-significant differences in microbial richness of this study as it is possible that inactive spores or cysts are distributed across the treatments.

Being soil fungi, much is unknown about the life histories of the Sordariomycetes. Most of the information available for them is in regards to their phylogeny and morphological characters (Zhang et al. 2006; Hsieh et al. 2007; Phadeng et al. 2009). From a functional perspective, the Sordariomycetes are known mostly as human and plant pathogens with *Fusarium* being the most commonly studied, but this is surely a skewed view of their full life history repertoire (Parry et al. 1995; Mukherjee et al. 2012). The *Fusarium* community is much more diverse in the fall chisel plow treatments, largely contributing to the high level of Sordariomycetes diversity in those plots (Figure 2.17). In addition, fall chisel and fall moldboard plow also contain a more diverse community of *Cylindrocarpon* and *Metarhizium* which are also largely known as plant pathogens. It is not clear why the fall treatments would lead to these differences, but it is worth noting that the spring moldboard treatment does not follow this pattern.

With respect to the Dothideomycetes, the significantly higher OTU richness detected under no-till can be attributed to a diverse community in the genus *Sporormiella*, which aside from one OTU does not appear in any of the tilled soils (Figure 2.18). The *Sporormiella* are largely known as obligate coprophilous decomposers of herbivorous animal dung and have been used to trace ancient megafaunal populations (Davis and Shafer 2006; Raper and Bush 2009; Rule et al. 2012). Crucially, over the 30 years that the Elora experimental tillage plots have been
maintained, there is no record of manure applications. The aforementioned studies all used increased Sporormiella spore counts as diagnostic characters for increased herbivore densities with the assumption that they are only associated with dung. The results from this study suggest that the Sporormiella may indeed have other unknown life history strategies within the soil. It is worth noting that metagenomic techniques cannot ascertain whether this Sporormiella community is actually growing in no-till soils, but it is highly unlikely that this pattern is due to a diverse community of inactive spores.

With regards to the AMF communities of these soils, evidence from both the ITS2 primers and the 18S primers designed for AMF, suggest that no significant differences in OTU richness of taxonomic ranks exist between treatments. The other 18S universal PCR primers that were used in this study, did not detect any significant differences in the Eukaryotic communities of these soils. It is worth noting that the vast majority of 18S sequences could only be identified as Eukaryota by MEGAN 5. This could be due to different names being assigned to the same sequences in the SILVA 119 database and as a result MEGAN 5 selects Eukaryota as the 'least common ancestor'. It is also likely that many of the Eukaryotes are not present in the SILVA 119 database. While the SILVA database is actively curated, quality checked, aligned and operates under a single taxonomic framework, it is still difficult to trace the origins of library sequence data in a next generation bioinformatic pipeline where thousands of identifications are obtained. The same is true for the fungal UNITE database, which like SILVA mines all of its sequences from International Nucleotide Sequence Database Collaboration (INSDC) databases (Koljalg et al. 2013; Quast et al. 2013). In both databases, accession numbers are provided and UNITE has an additional accession number identifying species hypotheses, but it is still difficult to examine data sources in a high throughput manner.
This study along with the majority of reviewed metagenomic studies, did not investigate changes in soil microbial community dynamics with time. One of the few soil microbiome studies to address the temporal component in the context of different soil management practices is Lauber et al. (2013). Their results from bacterial 16S amplicons suggest that OTU richness changes significantly over a growing season and was most heavily influenced by soil moisture and temperature. However, they found that from a community composition perspective, β-diversity is more consistent through time, with the significant differences that were detected being smaller in magnitude. Finally, their results suggest OTU richness under different land use types exhibited different patterns of change despite being located in close proximity. Another study found contradictory results although they used "16S rRNA gene finger-printing, based on terminal restriction fragment length polymorphisms" and a qPCR assay of several genes involved in the nitrogen cycle (Uksa et al. 2014). They found that microbial communities exhibited greater variation based on spatial sampling (topsoil, rhizopshere, drilloosphere) when compared with temporally collected samples. A study from the same experimental tillage plots in Elora, Ontario yielded interesting results (Smith et al. 2009). PCR–denaturing gradient gel electrophoresis (DGGE) targeted the amoA, nirK, nirS and genes involved in the nitrogen cycle in addition to measuring N₂O production and found that seasonal climactic variations had the largest influence on nitrifier and denitrifier bacterial populations. In addition, they found that the nitrifier and denitrifier community composition differed between tillage treatments after the spring thaw in March, but were unable to make claims about the functional significance of this result. Research from environments other than soil, demonstrates that many bacterial taxa exhibit conditional rarity whereby they fluctuate from being rare to abundant within a community through time and act as a "seedbank" of diversity (Caporaso et al. 2012; Shade et al. 2014).
This study is also unable to make claims about the abundance or biomass of microorganisms in these soils. This is due to two biases that are present in amplicon library preparation, the first being a PCR amplification bias whereby primer mismatches reduce the efficiency of amplification for certain taxonomic groups (Schloss et al. 2011; Cai et al. 2013). The second major bias is differential amplification efficiency of indexing primer tags during index PCR and sequencing (Deagle et al. 2013). Both of these biases mean that quantitative measures of biodiversity using amplicon sequencing are highly confounded as amplicon concentration will not accurately reflect true proportions of organisms within a community. Other factors that prohibit quantitative analyses are differential gene copy numbers across different taxonomic groups and artifacts introduced by bioinformatic processing (Schloss et al. 2011). A full discussion of the latter is beyond the scope of this study due to different the pipelines required for different sequencing platforms, the multitude of pipelines that are available and the range of different experimental goals that require unique analyses. In general, however, adapter trimming, quality filtering, clustering and chimera filtering can all influence community proportions (Schloss et al. 2011; Deagle et al. 2013).

The application of fertilizers in these plots is likely to alter the microbial community composition when compared with an undisturbed environment. Based upon a meta-analysis of 107 datasets from 64 global field trials, Geisseler and Scow (2014) concluded that mineral fertilizer applications result in a 15.1% increase in microbial biomass. However, this only occurs in soils with a pH above 5. In highly acidic soils, fertilizers actually reduce microbial biomass. The impact of fertilizer applications on microbial biodiversity are less clear. Ramirez et al. (2010) determined that increasing nitrogen concentrations lead to lineage specific changes in sequence diversity, but there was no change in the diversity of the lineages themselves. Other
studies did not detect any significant impacts of N applications on functional diversity, but again, there were differences in richness within lineages (Fierer et al. 2012; Lupwayi et al. 2012).

Other studies have investigated microbial biomass and abundance under different tillage practices by using methodologies that do not confound quantitative measurements. A common method for assessing microbial biomass involves characterizing phospholipid fatty acid (PLFA) concentrations and their chemical diversity. Sun et al. (2010) used this technique to examine bacterial and fungal biomass under no-till, minimum till, deep ploughing and conventional tillage across a depth gradient. They found that bacterial biomass was higher in the zero till treatment, but only in the upper 5cm. Fungal biomass followed a similar trend. However, across the entire depth profile (60cm) no significant differences were detected. Using PLFAs, van Groenigen et al. (2010) arrived at a similar conclusion when comparing conventional or shallow non-inversion tillage from samples collected at 0-5cm and 5-20cm. They found that both fungal and bacterial biomass was higher under reduced tillage in the upper 5cm. Babujia et al. (2010) found that the total microbial biomass was 35% greater under no-till in addition to increases in basal CO$_2$ respiration rates.

With respect to the crop rotation in use at the Elora plots, it is conceivable that this would influence microbial diversity as well. Certainly within the rhizosphere, an intimate relationship between plants and soil microbiota has been documented (reviewed by Gaiaro et al. 2013). However the influence of crop rotations on the microbial diversity of bulk soils appears to be subtle or nil in some cases (Navarro-Noya et al. 2013; Souza et al. 2013). However, in other cases, rotations that include legumes have been measured with more diverse bacterial communities when compared with grain monocultures (Xuan et al. 2011).
2.20 CONCLUSIONS AND RECOMMENDATIONS

The microbial diversity of these tillage treatment was similar on a broad scale. However, the fungi, Sordariomycetes and Dothideomycetes yielded interesting and contradictory results. The tilled soils were all comparable in terms of their diversity of these two classes, but the no-till soils were significantly different. For the Sordariomycetes, no-till had a significantly lower diversity while Dothideomycetes were significantly more diverse under no-till. Much of the increased diversity of Sordariomycetes was in the two fall tillage treatments, and was attributed to pathogenic fungi. Future work should examine pathogen loads of crops during the time that these soil samples were collected to corroborate this evidence. For the Dothideomycetes, the increase in diversity under no-till could be attributed to the Sporormiella - a supposedly obligate decomposer of dung. However, no manure has ever been applied to these plots. This could suggest a new life history strategy for Sporormiella but future work is needed to confirm this. In addition, future work should focus on temporal analyses of soil biodiversity under different tillage practices and address the issue of compaction in certain soil types.

At the outset of this study, it was hypothesized and predicted that:

**H1.** Agricultural fields under no-till have an increased retention of organic matter, pore connectivity and decreased physical disturbance compared to tilled fields. These environmental characteristics are more congenial for a more diverse community due differential requirements among taxa.

**P1.** No-till fields harbor a more diverse community of bacteria, protozoa, arthropods and fungi and high pore connectivity.
**H2.** Microbial diversity is fostered by the grazing of secondary saprotrophs.

**P2.** Fields with higher microbial diversity will also contain a high diversity of nematodes and protozoa that graze on bacteria (causation in this case will require manipulation in future experiments).

**H3.** Tillage is traumatic for microorganisms due to shearing forces and other physical disturbances and there is differential sensitivity to disturbances among taxa.

**P3.** No-till soils will harbor a greater microbial diversity.

In the context of the microbial communities H1 and H2 may be too simplistic as the bacterial and protozoan communities were very similar across treatments in a taxonomic framework. However, their phylogenetic diversities suggest the 16S OTU membership of these coarse level taxonomic groups may indeed differ between treatments. With regards to H3, tilled soils in this study have a richer community of unicellular Sordariomycetes, while no-till has a richer community of Dothideomycetes which create hyphal networks. This may indeed be due to the physical disturbances in tilled soils and the compacted nature of no-till.

**CHAPTER III: SOIL ARTHROPOD DIVERSITY UNDER DIFFERENT TILLAGE PRACTICES**

**3.1 SOIL SAMPLING**

Bulk soil samples were collected to a depth of 15cm using a trowel and placed in 2L plastic bags. All tillage treatment samples were collected on September 26, 2013 from the minimum tillage experimental plots at the Elora Agricultural Research Station, Ontario, Canada (43°38'38.0"N 80°24'20.5"W). Three replicates were collected for each of the five treatment
types from separate replicate plots. Three samples were also collected from a maple-dominated hardwood woodlot that is 500m away from the tillage plots (43°38'53.4"N 80°24'23.8"W) and is 0.035 km$^2$ in area. In order to estimate the maturity of the woodlot, the entire woodlot was surveyed and circumferences of the 20 largest trees were measured at a height of 145cm. Then, tree diameters were calculated and multiplied by species specific growth factors calculated by the International Society of Arboriculture. In addition two large maple stumps were found and the number of rings per 10cm were counted. It was not possible to count the total number of rings as the stumps were in an advanced state of decomposition.

3.2 SOIL SAMPLE PROCESSING

Samples were placed in BioQuip Collapsible Berlese funnels with the only modification being an addition of a two-ply layer of cheese cloth to the metal grate (Figure 3.0). This was done to prevent large amounts of soil from crumbling through the grate as drying progressed. 60W light bulbs were used to dry the soil samples until they felt completely desiccated (~1 week). Emerging arthropods were collected in 70% ethanol. After fixing, 100% ethanol was added to increase the concentration to 95% and the samples were stored at -20°C. Since only eight Berlese funnels were available, half of the bulk soil samples (randomly selected) were therefore stored at 4°C for ~1 week. Soils were weighed after drying in the Berlese funnels. It must be noted, that the sampling strategy employed here was designed to target soil arthropods and will largely miss lumbricid and nematode worms.
Figure 3.0 Stylized depiction of the Berlese funnel set-up that was used to coax soil arthropods out of bulk soil samples and into ethanol for storage and processing.

### 3.3 PHYSICAL SORTING OF ARTHROPOD SPECIMENS

Arthropods were visually sorted under a dissecting microscope and identified to the taxonomic levels of class or order by a non-expert undergraduate student (Danielle Bourque) using Marshall (2006). The student was blinded in terms of treatment type. Larger arthropods such as Coleoptera or Diplopoda were sub-sampled by removing either a leg or a few body segments. For smaller Arthropods such as mites and Collembola, entire organisms were used. All organisms from a sample were placed together in a 1.5mL Eppendorf tube in 95% ethanol and stored at -20°C until DNA extraction. Unfortunately, voucher specimens are not available due to the DNA extraction method used.
3.4 DNA EXTRACTION AND PURIFICATION

Ethanol was removed from all tubes using a heated SpeedVac centrifugal evaporator. The following DNA extraction protocol was modified from Ivanova et al. (2006) based on personal communication with Natalia Ivanova.

1. For each tube of bulk arthropods, 50μL of Insect Lysis Buffer and 5μL of Proteinase K (20 mg/mL) were mixed in a sterile container. 50μL of Lysis Mix was then added to each tube.
2. Tubes were incubated at 56°C in a shaking incubator at 800rpm for 16 hours.
3. Tubes were centrifuged at 6500g for 15 seconds to remove any condensate from the cap strips.
4. 100μL of Binding Mix was added to each sample.
5. Samples were mixed by pipetting and were transferred from the tubes into spin columns. Epoch Biolab spin columns with attached lids were used.
6. Spin columns were centrifuged at 5500g for 2 minutes to bind DNA to the GF membrane.
7. **First wash step:** 180μL of Protein Wash Buffer (PWB) was added to the spin column. Columns were centrifuged at 5500g for 2 minutes. Contents in collection tube were discarded.
8. **Second wash step:** 700μL of Wash Buffer (WB) was added to the spin column and was centrifuged at 5500g for 4 minutes.
9. Flow-through in the collection tube was discarded. Remaining WB was removed from the filter by centrifugation at 10,000g for 4 minutes.
10. Collection tubes were replaced with 1.5mL tubes with open lids. Spin column lids were opened and incubated at 56°C for 20 minutes.
11. 50μL of ddH20 (pre-warmed to 56°C) was added directly onto the membrane of each column and incubated at room temperature for 1 minute.

12. DNA eluate was collected by centrifugation at 10000 g for 5 minutes.

3.5 PCR INHIBITOR REMOVAL

Initial PCR tests failed and it was surmised that PCR inhibitors were present in the DNA extracts. The MO BIO PowerMax Soil DNA isolation kit was chosen as it contains two inhibitor precipitation reagents and was used for the microbial soil samples from the study. In order to purify the bulk arthropod DNA extracts, the following modifications were made to the manufacturers protocols: all reagent volumes were scaled down by a factor of 0.03 so the entire volume of DNA extract could be purified using Epoch Biolab spin columns and all centrifuge spin times and speeds were adopted from the bulk Arthropod DNA extraction protocols previously outlined. All samples were eluted in 60μL of solution C6 (10mM Tris-HCL).

3.6 TAXON SPECIFIC PCR

The following CO1 mini-barcode primers were ordered from Integrated DNA Technologies with Illumina adapter tails and mixed in equal parts to yield a cocktail:

RonIIdeg_R 5’-GRRGGRTARAYAGTTCATCCWGTWCC-3’
AMR1deg_R 5’-CAWCCWGTWCCKRMNCCWKCAT-3’
MLepF1 - 5’-GCTTTCCCACGAATAAATAATA-3’
HCO2198 5’-TAAACTTCAGGGTGACCAAAAAATCA-3’
LepR1 5’-TAAACTTCCTGGATGTCCAAAAATCA-3’
These primers were chosen as they have been used in previous mite and Collembola biodiversity studies and were therefore used to generate many of the sequences deposited in the Barcode of Life Datasystems (BOLD) library of CO1 sequences (Hogg et al. 2004; Porco et al. 2012; Young et al. 2012). Also, CO1 has also been demonstrated to be far superior to 18S in the context of unexplored marine invertebrate biodiversity (Tang et al. 2012). In addition, the total amplicon length of 407bp is compatible with the 250bp paired end read chemistry of the Illumina MiSeq. As with the microbial portion of this study, the high fidelity polymerase KAPA HiFi HotStart in the ReadyMix format was used. Annealing temperature gradient experiments were used to select 52°C, 55°C and 58°C as "low, medium and high" annealing temperatures to be used for all samples as recommended by Schmidt et al. (2013). In accordance with the microbial portion of this study, three replicate 12.5μL PCR reactions were set up for each sample to be run at the three different annealing temperatures. KAPA’s reaction setup protocols were adhered to with 6.25μL of ReadyMix, 4μL of template, 0.75μL primer cocktail (0.6μM) and 1.5μL ddH2O being used. Eppendorf Mastercycler Pro S thermal cyclers were used to run the following protocol: 95°C for 5 minutes followed by 30 cycles of denaturation at 98°C for 20 seconds, primer annealing for 15 seconds and extension at 72°C for 35 seconds. A final extension at 72°C for 5 minutes was used. PCR success was confirmed using an Invitrogen E-Gel and then amplicons from the three different annealing temperatures for each sample were pooled. All amplicons were purified with Aline PCRClean DX.

3.7 INDEX PCR
The same index primers that were used for the microbial portion of this study were used to index each pooled set of amplicons. KAPA HiFi HotStart ReadyMix reactions were set up in 50 μL total volumes as follows: 25μL ReadyMix, 1.5μL (0. μM) of both index primers, 8μL template amplicons and 14μL of ddH2O. The index PCR cycling program that was used for the microbial portion of this study was also used here. PCR success was verified using an Invitrogen E-Gel. Amplicons were purified using PCRClean DX at a ratio of 1.8 (beads): 1 (sample). Purification success was verified using an Invitrogen E-Gel and amplicon concentration was measured using a Thermo Fisher Scientific Nano-Drop (Simbolo et al. 2013). All samples contained >250ng of amplicons. This permitted the use of the Invitrogen SequalPrep Normalization Plate (96) Kit to normalize all samples prior to Illumina MiSeq sequencing using the 250bp paired end read chemistry. Amplicons were loaded into the MiSeq as a 6pM denatured library which was 15% PhiX.

### 3.8 BIOINFORMATIC PIPELINE

In general, the arthropod sequences were processed using the same setting as the microbial sequences. Any differences are highlighted below. Primers were trimmed from all reads using cutadapt (Martin 2011). UPARSE was used to merge all read pairs, trim all sequences to a universal length of 400bp and quality filter with a maximum expected error threshold of 1 (Edgar 2013). Singletons and doubletons were discarded and all sequences were clustered at an OTU radius of 1 (99% clustering). All OTUs were aligned and visually inspected in MEGA6 (Tamura et al. 2013) to ensure the absence of pseudogenes by checking for stop codons. The same alpha-diversity metrics as for the microbial OTUs were calculated in UPARSE. BLASTn was used in the NCBI web portal to identify all OTUs with
"uncultured/environmental sample sequences" being excluded from the NCBI database. BOLD was not used as MEGAN5 cannot parse the BOLD ID result files to select the "most correct" ID. All BLAST results were downloaded in .xml format and the LCA method was used to select the most correct taxonomic names for each OTU in MEGAN 5 (Huson et al. 2011).

3.9 ARTHROPOD RESULTS

The mature woodlot was determined to be at least 150 years old based on tree diameters and ring counts. Based on the visual taxonomic identifications of the arthropod specimens, all of these soil communities are dominated by mites (Arachnida) and Collembola. When examining the principal component analysis results for all arthropod OTUs there is a high degree of overlap between treatment clusters (Figure 3.1A). However, when examining genus-level results, clusters emerge with a high degree of both coalescence and distinctness for the spring moldboard, fall moldboard and woodlot treatments. From an alpha- diversity perspective, all agricultural treatments are similar and the woodlot samples are the most diverse (Table 3.1).

Soil samples with higher dry masses do not correlate with higher OTU discovery (Figure 3.2). In fact, the weakly correlated trend ($r^2$=0.21) suggests the exact opposite, but this is skewed by the fact that the diverse woodlot samples all had significantly lower masses via a two sample t-test assuming unequal variances ($p = 5.88e-4$).

When examining OTUs from all samples, it is apparent that the primers and or the chosen PCR cycling conditions did not amplify DNA from the Diplopoda, Chilopoda or Symphyla that were identified visually. However, these taxa were never found in high abundance and their presence was highly variable between samples.
When examining the Collembola OTU diversity, a distinct pattern emerges (Figure 3.3). The mature woodlot that is 500m away from the experimental tillage plots has a much higher OTU richness compared with the tilled and no till soil. There is, however, a high degree of variability, and this result is not statistically significant. It is, however, supported by the raw number of specimens that were collected (Figure 3.4). In both OTU richness and raw abundance (counted specimens), the spring tandem disc treatment has the lowest mean. This too, however, is not statistically significant. Within the Collembola, some interesting significant differences were found, the woodlot has a much higher diversity of genera that are not found in any of the tilled systems (Figure 3.7). A high proportion of the woodlot Collembola community contains the Poduroidea and the Isotomidae, which are not found in any of the agricultural plots (Figure 3.7). Conversely, the agricultural plots are dominated by other OTUs assigned to Entomobryomorpha, which did not appear in any of the woodlot samples (Figure 3.7).

The OTU richness of the mite community between treatments displays a different pattern. There is a high degree of variability with all means being comparable except for the spring tandem disc treatment, that only contains 1 OTU (Figure 3.5), identified as Opiella. This pattern is not statistically significant but is again supported by the raw specimen abundance data (Figure 3.6). As with the Collembola, there are significant differences in mite taxa that only appear in the woodlot treatments such as the Ceratozetes, Ascidae, and Oppiidae (Figure 3.8). The Pygmephoroida appeared in all agricultural treatments expect for spring tandem disc and did not appear in the woodlot (Figure 3.8).

Some infrequent but strange identifications were obtained, as seen in Table 3.2. They are Decapoda, Onychophora and Aphonopelma. The latter matched 5 OTUs, and the highest sequence similarity was 80%, with multiple hits in the NCBI database. Similar results were seen
following individual queries in BOLD. These sequences could be sequencing or PCR chimeras, despite making it past the UPARSE chimera filtering step. Interestingly, these only occurred in the spring disc treatment. The Onychophora (*Euperipatoides rowelli*) identification was only for one OTU sequence and the vast majority of the other hits were to *Anopheles*. It is not clear why the LCA algorithm ignored these. This sequence yielded no hits in any of the BOLD databases. Finally the Decapoda identification also hit Collembola and Coleoptera sequences in NCBI and had three 100% matches with Hymenoptera sequences in BOLD.

Based on the rarefaction curves generated for all of these treatments, it is reasonable to assume that the majority of the mite and Collembola community has been sequenced (Figure 3.8).

### 3.10 ARTHROPOD FIGURES

Table 3.1. Alpha-diversity metrics computed for CO1 OTUs in UPARSE. N dictates the number of quality filtered sequences that contributed to the OTUs.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>OTU Richness</th>
<th>Shannon Entropy</th>
<th>Renyi Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO1 fall chisel</td>
<td>9004</td>
<td>23</td>
<td>1.2</td>
<td>3.1</td>
</tr>
<tr>
<td>CO1 fall mold</td>
<td>8353</td>
<td>27</td>
<td>2.2</td>
<td>3.3</td>
</tr>
<tr>
<td>CO1 no-till</td>
<td>10221</td>
<td>24</td>
<td>1.8</td>
<td>3.2</td>
</tr>
<tr>
<td>CO1 spring disc</td>
<td>3220</td>
<td>16</td>
<td>1.6</td>
<td>2.8</td>
</tr>
<tr>
<td>CO1 spring mold</td>
<td>5629</td>
<td>25</td>
<td>2.4</td>
<td>3.2</td>
</tr>
<tr>
<td>CO1 woodlot</td>
<td>4862</td>
<td>51</td>
<td>2.8</td>
<td>3.9</td>
</tr>
</tbody>
</table>
Figure 3.1. Principal component analysis results from STAMP acquired using default settings. Comparisons were made using taxonomically identified CO1 OTUs from all treatment replicates from five tillage regimes and a mature woodlot. In the legend, the numbers in parentheses
correspond to the number of replicates that were analyzed. Plot A includes all taxonomic identifications that were made for all OTUs. Plot B only includes OTUs that were identified as either Collembola or Mites (Arachnida) both of which comprised the majority of these soil communities.

Figure 3.2. Relationship between bulk soil sample dry mass and the number of CO1 OTUs obtained from those samples. Samples marked in red were collected from the mature woodlot while samples marked in blue were collected from the tillage plots.
Figure 3.3. CO1 OTU richness of Collembola OTUs (y-axis) between tillage treatments and a mature woodlot. An ANOVA did not identify any significant differences (p > 0.05). OTUs were clustered at 99% similarity from 400bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures.
Figure 3.4. Raw specimen abundances for Collembola sorted from all soil samples and visually identified.
Figure 3.5. CO1 OTU richness of Mite (Arachnida) OTUs (y-axis) between tillage treatments and a mature woodlot. An ANOVA did not identify any significant differences (p > 0.05). OTUs were clustered at 99% similarity from 400bp Illumina MiSeq reads that were generated using three different PCR annealing temperatures.
Figure 3.6. Raw specimen abundances for Mites (Arachnida) sorted from all soil samples and visually identified.
Figure 3.7. Heat-map plot detailing the number of CO1 sequences (OTUs) that were assigned to Collembola using the NCBI database (excluding environmental samples) and the BLASTn classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.
Figure 3.8. Heat-map plot detailing the number of CO1 sequences (OTUs) that were assigned to mites using the NCBI database (excluding environmental samples) and the BLASTn classifier. Treatment replicates (bottom of plot) were clustered using the nearest neighbour method in STAMP and the results are displayed at the top of the plot.
Table 3.2. Summary of all Taxonomic identifications at the level of order or higher made for CO1 amplicons along with p-values from between treatment ANOVAs and Eta-squared effect sizes comparing OTU richness within taxonomic ranks.

<table>
<thead>
<tr>
<th>Taxonomic Id</th>
<th>p-values</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aphonopelma</td>
<td>0.458197</td>
<td>0.294118</td>
</tr>
<tr>
<td>Coleoptera</td>
<td>0.700732</td>
<td>0.200206</td>
</tr>
<tr>
<td>Collembola</td>
<td>0.105465</td>
<td>0.493885</td>
</tr>
<tr>
<td>Decapoda</td>
<td>0.458197</td>
<td>0.294118</td>
</tr>
<tr>
<td>Diptera</td>
<td>0.524593</td>
<td>0.267734</td>
</tr>
<tr>
<td>Gastropoda</td>
<td>0.458197</td>
<td>0.294118</td>
</tr>
<tr>
<td>Hemiptera</td>
<td>0.514456</td>
<td>0.271694</td>
</tr>
<tr>
<td>Isopoda</td>
<td>0.458197</td>
<td>0.294118</td>
</tr>
<tr>
<td>Ixodes</td>
<td>0.139</td>
<td>0.463966</td>
</tr>
<tr>
<td>Lepidoptera</td>
<td>0.401963</td>
<td>0.317576</td>
</tr>
<tr>
<td>Mite (Arachnida)</td>
<td>0.122229</td>
<td>0.478207</td>
</tr>
<tr>
<td>Onychophora</td>
<td>0.458197</td>
<td>0.294118</td>
</tr>
</tbody>
</table>
When examining the principal component analysis results for the Collembola and mites, very distinct clusters emerge which are far more coherent than any observed for the microbes (Figure 3.1[B]). The mature woodlot samples contained the highest diversity and abundance of Collembola, however these results were not significant, due to high between-replicate variability.
There were significant differences upon examining the identifications within the Collembola, the Poduroidea and the Isotomidae dominate the woodlot samples and are not found in any of the agricultural plots (Figure 3.7). Conversely, the agricultural plots are dominated by certain Entomobryomorpha OTUs which did not appear in any of the woodlot samples (Figure 3.7). With respect to tillage treatments, spring tandem disc tillage had the lowest mean Collembola abundance and diversity, although the latter was very slight. These differences may be attributable to differences in soil density, treatment intensity, litter abundance and composition. The fact that all agricultural plots did not differ significantly in terms of Collembola richness is also supported by Brennan et al. (2006).

It is possible that local climates influence the response of soil invertebrates to changes in soil structure and nutrient availability. The decomposition rate of identical oak leaves (*Quercus prinus*) was compared in two humid tropical forests and one temperate forest, and leaf mass loss was significantly faster in the tropics (Heneghan et al. 1999). Additionally, arthropods in the temperate forest had a more reduced role in decomposition. These results are further supported by Six et al. (2002) who measured a more rapid carbon turnover in tropical versus temperate soils. They also determined that no-till systems in both tropical and temperate environments gradually accumulate organic carbon when compared with tilled systems.

A standard volume of 2L was collected for all samples, and based on dry soil mass, the forest soils were significantly less dense ($p = 5.88e-4$) (Figure 2). Larsen et al. (2004) experimentally manipulated soil bulk densities and inoculated them with three different Collembola and found that two of the three species had significantly lower abundances under increasing compaction. Interestingly, the two species that were sensitive to compaction were within the Poduroidea while the one that was not was within the Entomobryomorpha, thus
matching the observations of this study. However, another study examined the impact of organic matter removal and compaction on wild forest Collembola and found that only organic matter removal decreased their abundances (Eaton et al. 2004). In the context of management practices, tilled soils have been observed with lower overall arthropod abundances compared with no-till and "old fields" (Wickings et al. 2013). In that study, arthropods were more abundant under decomposing grass when compared to corn, but were less abundant in litter types with high lignin contents.

With respect to the mite diversity of these soils, there is a high degree of variability with all means being comparable except for the spring tandem disc treatment which only contains 1 OTU (Figure 3.5) which was identified as *Opiella*. This pattern was not statistically significant but was supported by the raw specimen abundance data (Figure 3.6). As with the Collembola, there are significant differences in mite OTUs that only appear in the woodlot treatments from taxa such as the Ceratozetes (Oribatida), Ascidae (Mesostigmata), and Oppiiidae (Oribatida) (Figure 3.8). The Pygmephoroida appeared in all agricultural treatments expect for spring tandem disc and did not appear in the woodlot (Figure 3.8). Experimental manipulations of woodlot vegetative cover and soil compaction have demonstrated that arthropod diversity decreases with increasing disturbance intensity and oribatid mites seem particularly sensitive. Battigelli et al. (2004) found that clear cutting and soil compaction reduced mesofaunal diversity by 93% and the diversity of rare orbibatids by 40% in a forest in British Columbia, Canada. They also found that Prostigmata and Mesostigmata diversities increased with treatment intensity presumably to fill empty niches. Other researchers have also suggested that the Oribatida are useful bioindicator species (Crossley 1992; Sanchez-Morino et al. 2009; Cluzeau et al. 2012). Therefore, the greater orbibatid diversity in the woodlot compared with all agricultural plots
suggests that the soil environment in all agricultural plots, including no-till, is in a negatively disturbed state. It is worth noting that other soil disturbances such as bioturbation are actually beneficial to the soil environment as tunneling by various invertebrates allows for the movement and water, nutrients and gases (Wilkinson et al. 2009).

3.12 CONCLUSIONS AND RECOMMENDATIONS

At the outset of this study it was hypothesized and predicted that:

**H4.** Tillage disturbances are harmful for soil arthropods due to shearing forces and other physical disturbances and there is differential sensitivity to disturbances among taxa.

**P4.** No-till soils will contain a greater richness of arthropods when compared to tilled soils.

The impact of tillage on the arthropod community is not as dramatic as it was hypothesized to be. The arthropod communities are quite similar across the five agricultural management practices and the woodlot samples are the most diverse in terms of community composition and richness.

In the context of previous research, the Collembola and mite results from this study suggest that even the no-till soils are in a more disturbed state than woodlot soils and of the four tillage treatments, spring tandem disc tillage is the most intense in terms of disturbance to the arthropod community. However, a high degree of within-replicate diversity resulted in a lack of significant differences in OTU richness for the Collembola and mites. However, within these taxonomic groups there were significant differences in the richness of compaction sensitive taxa, many of which were only found in the woodlot.
4. FINAL SUMMARY AND CONCLUSIONS

This study was the first to examine the influence of agricultural tillage practices on the entire soil ecosystem using metagenomic techniques. Four genes were chosen to investigate bacterial, fungal and protozoan diversity, while one gene was used for arthropod diversity with all being sequenced at a very high read depth per sample (259,331 merged reads ± 78,953). Four tillage treatments were compared with a no-till treatment, using samples collected from the long term management plots in Elora, Ontario. Overall, the microbial communities of these soils were very similar (p>0.05), but some interesting differences were found with regards to the fungi. Overall, the most diverse fungal group was the Sordariomycetes (103±19 mean OTU richness across treatments), and no-till had a significantly lower diversity compared with the tillage treatments (86±10 OTUs; p= 3.12e⁻³). The fall tillage treatments were the most diverse in terms of the Sordariomycetes (113±16 OTUs), but a large proportion of this diversity were pathogenic genera upon plants. Conversely, no-till soils possessed a significantly more diverse community of Dothideomycetes (43±1 vs 29±6; p= 5.54e⁻⁵) with the majority of this diversity belonging to the Sporormiella. This genus was essentially absent from the tillage treatments, and even more interestingly, it is known as an obligate coprophilovore. However, manure has never been intentionally applied to these soils. With regards to the arthropod communities, additional mature woodlot samples yielded a more diverse community of Collembola and other compaction sensitive taxa, suggesting that even the no-till soils suffer from anthropogenic disturbances. However, from an OTU perspective, a high degree of between-replicate variability resulted in a lack of statistically significant differences between treatments for the Collembola and mites. However, within these taxonomic groups there were significant differences in the richness of compaction sensitive taxa, many of which were only found in the woodlot. Overall, no-till or
conservation tillage practices are recommended, but the results generated here suggest that they cannot be solely relied upon to increase soil biodiversity.

REFERENCES


college


